User Manual

Version 5.9

© 2007 to 2019 Experian

The Client Interface

The Experian Pandora client is split into the following areas:

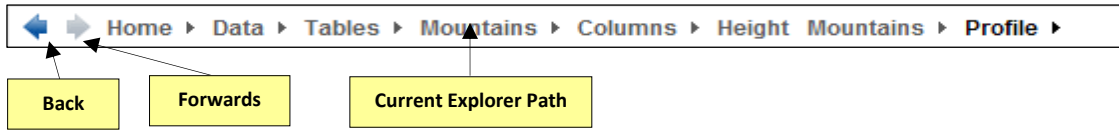- The **navigation bar** and **quick bar**
- The **explorer**
- The **desktop**
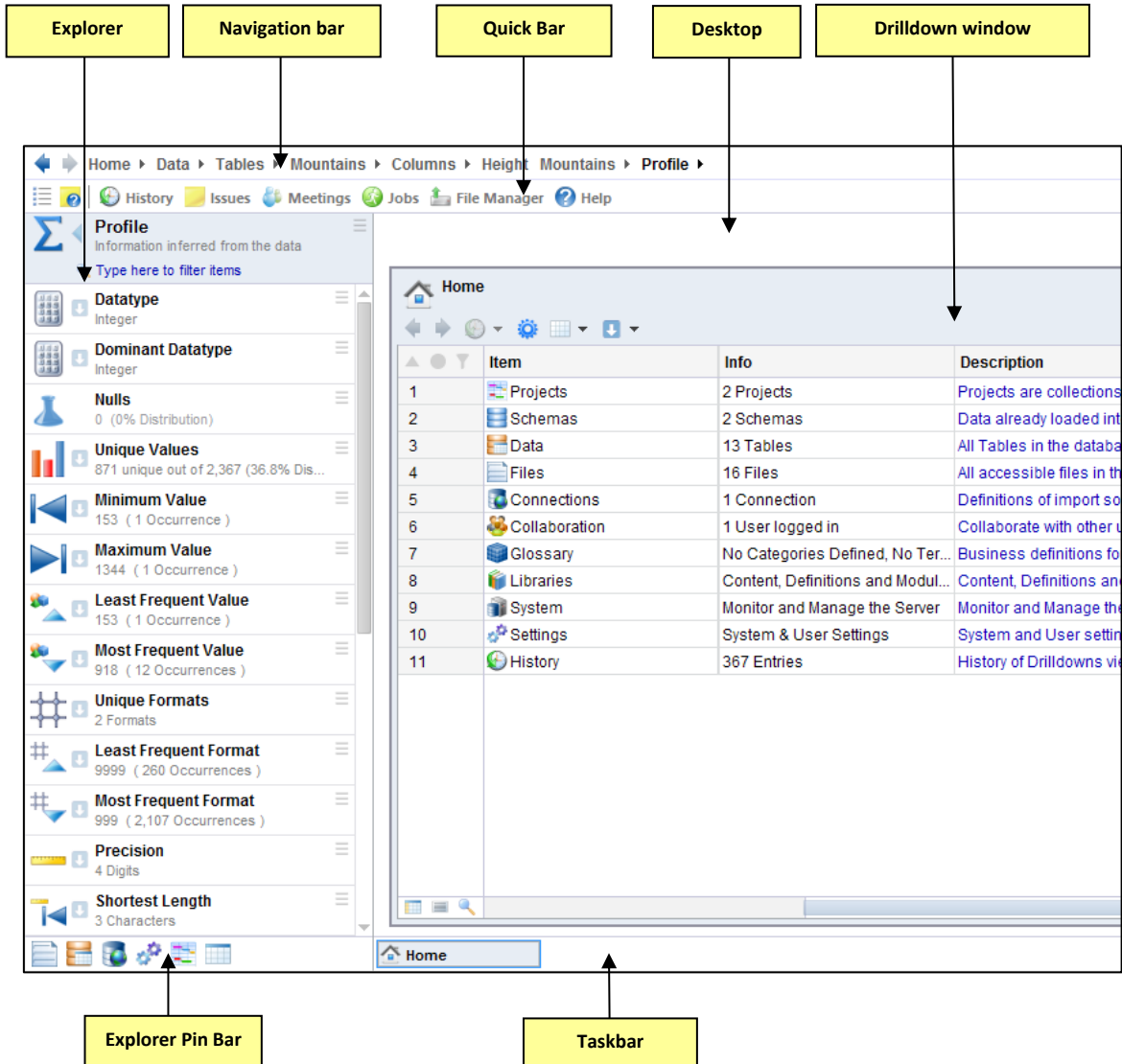
When you exit the Experian Pandora application, the state of your session will be saved on the server. This will be remembered regardless of which physical machine you next log in from.

Navigation Bar

Back and Forward arrows navigate back and forward in the history of items you have navigated in the **explorer**.

The **navigation bar** path acts like a *bread-crumb* bar showing you each of the levels in the hierarchy down to your current navigated position in the **explorer**. This is shown as a series of item names, separated by right-hand arrows.

Left click any item name in the navigation bar and the explorer will navigate to that location.

Left click the arrow next to the item name in the navigation bar to get a drop down menu to select the location you want the explorer to navigate to.

While the **explorer is visible**, the movements you make using the navigation bar will be reflected in the explorer.

If you have **hidden** the explorer, the movements you make using the navigation bar will be reflected in the drilldown area, with a drilldown containing the menu items or objects on that level.

If you right-click on one of the navigation bar items, you will get the same menu that you get if you right click on that header item in the **Explorer**.

Quick Bar

The **Quick Bar** at the top of your application window is a useful tool for quick navigation to some of the most common tasks used in Experian Pandora .

The first tool is to Show/Hide the Explorer. The second is a short cut to your user preferences.

When a number appears with the icon click the icon to view the drilldown and information, e.g. if you have a job loading the Jobs icon will have a number against it while the job is loading. Open the drilldown to view the status of the scheduled job and all relevant information.

The 'Action' button is used to list and run any global scripts that have been defined in Experian Pandora. Global scripts run without any specific context, and therefore are listed in a single central location.   The button will be disabled if there are no global scripts. If a global script is added, the button will become enabled, and clicking on it will open a drilldown showing all the scripts. Clicking on the script will execute it. Please refer to the 'Scripting and REST Guide' for more information.

Pin Bar

From the Explorer you can drag and pin the icons of your choice for quick access. Hover over the icon and a tool tip will pop up to tell you what the icon is for. If you pin individual tables, hover over the icon to see which table it is. Clicking on an icon will take you directly to the location in the Explorer. A right-click menu is available on each icon for quick access to the related areas of Experian Pandora. To unpin an icon simply right-click and select **Unpin**. Experian Pandora ships with a set of icons which can be changed by the user.
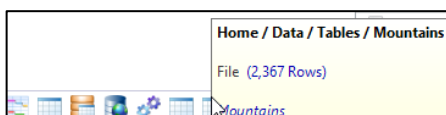
Explorer

The **explorer** is used to navigate multiple levels of data, metadata, information and statistics that form the Experian Pandora ecosphere. The explorer shows you a single level at a time, with your overall path displayed in the **navigation** bar. The **explorer** allows left/right navigation through the path, whereas the navigation bar allows instant access to any point.

The large icon to the left of an Explorer item is used for navigation or action, with the smaller moniker icon to the right of the larger icon detailing the type of action available:

- Move to next level of the Explorer
- Move to previous level of the Explorer (only appears on the Explorer header
- Open the associated drilldown for this Explorer item
- Toggle the Explorer item (used for on/off settings)
- Edit the Explorer item (used for settings with a dialog box)

The name of the item can be clicked on to perform the associated activity, which will depend on the moniker icon visible to the left of the item name. Click the item name to invoke the activity. In addition to a moniker being displayed, the item name will change colour and become underlined when you hover over it to denote that an activity is available if clicked.

Explorer folder items with no activity available will appear light grey (for example an empty folder of Tables). Explorer items that are not folders but do not have an associated activity such as a drilldown will still be appear black for ease of reading. These will remain black when you hover over them so you can see that there is not further drilldown available.

Two additional icons may appear to the far right of the explorer item:

- Displays the menu for the Explorer Item. This can also be accessed by right-clicking on the explorer item

- Denotes that one or more notes are attached to object described by the Explorer item.
  Click to view the attached notes.

The **arrow** or **icon** will take you back a level in the explorer.

The **filter** can be used to search the explorer menu items. It supports wildcard searches (*) and regular expressions.

If the **item name** in the explorer item is underlined on hover over, clicking on it will open the drilldown.

The **icon** represents the explorer item, and in some cases, information about its current state.

This is the moniker icon which denotes what activity this element will perform when clicked on.

When this is an **arrow**, clicking it will go to the previous or next level in the explorer depending on the direction of the arrow.

If this is a **blue box with a down arrow**, clicking it will display an associated drilldown.

When the icon is a **finger on a button,** clicking this will toggle a setting on or off.

If the icon is a **white page with pen**, clicking it will open an associated edit dialog.

This is the name of the current level in the **explorer** .

More information about the item which either explains its use or provides additional metadata about it.

The **sticky pad** icon signifies that this item has Notes attached to it. Single-click on the icon to view the Notes in a drilldown window.

The icon of **three lines** seen at the edge of the explorer item will provide the sub-menu for an item when clicked. This is also available by right-clicking anywhere on the item.

The **pin bar** can be used by dragging explorer icons onto it to create a shortcut, giving the user easy access to their frequently used explorer items.

Columns
13 Columns
Type here to filter items

Customer Id
33,360 (99.71% Unique)

Discount Code
8 (0.02% Unique)

Forename
3,386 (10.12% Unique)

Surname
9,991 (29.86% Unique)

Email
33,360 (99.71% Unique)

Telephone
33,190 (99.2% Unique)

Company Name
27,639 (82.61% Unique)

First Order Date
13,500 (40.35% Unique)

Sales Last Year
324 (0.97% Unique)

Addressline
25,822 (77.18% Unique)

City
2,269 (6.78% Unique)

Country
7 (0.02% Unique)

Post Code
27,511 (82.23% Unique)

Desktop

The desktop is your workspace for viewing **drilldown** windows for data and metadata. This may be the result of user-requested queries of data and metadata from the Experian Pandora Server, but also for application information such as the schedule of tasks and system information. Some drilldowns are real-time, updating without any input from the user. Experian Pandora endeavours to always show you the latest up-to-date information, refreshing drilldowns automatically.

Many of the items in the explorer will produce a drilldown, either by single-clicking on the item name or by selecting one of the available drilldown options that appears on the item menu. All menu items that produce a drilldown are in bold type where as if it is light-grey colour there is no drilldown available.

At the bottom of the desktop you can see the taskbar.

The desktop supports **multiple drilldown windows** as well as the standard **single drilldown mode**. This can be changed in **My Preferences** under **Settings**. A quick way to add another window to the task bar is to right click on any tab in the task bar and select **Clone.** You will now have two tabs which will work independently.

Drilldown windows can be horizontally split by clicking on the Drilldown Action Menu and selecting **Split** (see section on How to Use Drilldowns for further details).
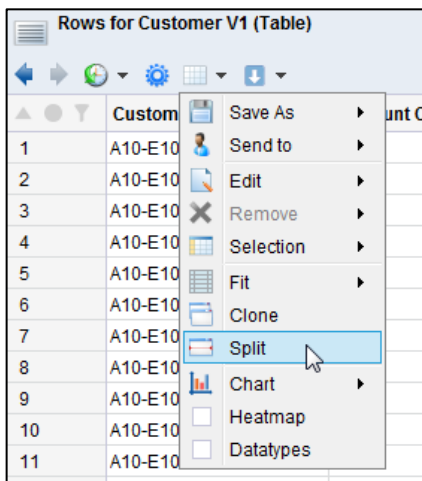
Drag and Drop

Items on the explorer can have their associated drilldowns invoked by dragging their icons onto the desktop.

Dropping onto the Desktop

If you are not in Single Drilldown mode, when dragging the explorer icon, Experian Pandora will open a new drilldown window when you drop the explorer icon over an area of blank desktop.

Dropping onto a Drilldown

Dropping the explorer icon onto an existing drilldown window whether in Single Drilldown mode or not will result in the new drilldown being added to the existing drilldown window as part of its history.

If the CTRL key is held as you drop the icon then Experian Pandora will open a new window for the drilldown.

If you are looking at the rows of a table and dragging another table's rows on top of it, the Linking section of the Drilldown Manager will appear (see the section Table Linking for more information).

Dropping onto the Taskbar

Experian Pandora will open a new window for the drilldown or if you are in single drilldown mode it will open it in the drilldown window. To return to your previous drilldown use the back arrow in the drilldown window.

Loading Data

All data loaded into Experian Pandora is automatically analysed, producing a rich set of statistics about the quality of the data using all of the source data without any sampling or estimation.

Tables

The Experian Pandora database has **Tables** of data, which are made up of one or more **Columns**. A Table has multiple **Rows** or **Records** of data. Table names can be configured prior to load or specified at load time and must be unique. Experian Pandora permits renaming of a Table, and auto-generates Table names to ensure unique names where necessary. Experian Pandora provides a full list of Tables in the **Data > Tables** drilldown.

Tables can support multiple versions. The latest version is shown in the Tables drilldown, with the corresponding version number. Previous versions are found under the **Versions** folder of a given Table in the Explorer or the drilldown. Tables which have multiple versions enabled all share a common name (that of the Version Group), but retain the memory of their original names in-case multiple versions are disabled at a later stage.

Columns

A Column is a single field in a table. Columns contain **Values** which can be one of several **Data types**. Column names must be unique within a Table and Experian Pandora handles renaming of Columns and auto-generation of Column names automatically. In the case where no Column names were supplied, Experian Pandora will highlight that fact to the user.

Experian Pandora calculates statistics for the Column both at an overall Column level and per datatype discovered within the Column.

Schemas

A Schema is a sub-division of the set of Tables loaded in Experian Pandora . **A Table can only belong to a single Schema.** It is best used to model source or target systems or databases.

Creating a Schema

On the Explorer, select **Home > Schemas**, right-click and select **New** to create a new Schema and assign Tables to it.

Deleting a Schema

To delete a Schema you must first remove any Tables from the Schema, and then select the **Delete** menu option on the Schema you wish to delete.

Default Schema

Tables always belong to a Schema. If none have been created, they will belong to the **Default** Schema.

Values

Experian Pandora is not **strictly typed**, and will **never** reject data when loading. For example, a Column may be assumed to be **Integer** when, actually some records contained **Alpha Numeric** values. Experian Pandora identifies the correct datatype of each value when loading data in accordance with the user-selected loading rules on a per Column basis. The default recommended set of rules is normally suitable for all data. Each Column can support **multiple** data types concurrently, each with its own full set of statistics calculated independently per datatype and collectively for the entire Column.

Experian Pandora automatically standardises values where needed and keeps track of how data was changed. This frees the analyst from worrying about potentially different data formats and gives the data a single business meaning regardless of technical storage differences. For example, the numbers **0**, **000**, **0.00** and **00.00** all represent **'0'** (zero). Experian Pandora will store such values as **0**. There are a wide number of standardisation operations that it performs, including, but not limited to, Date and Money. All leading and trailing white space is normally removed from Values when loading.

When using high precision calculations resulting in Decimals there are 3 settings:

High	128 bits	IEEE 754R Decimal128 format which is 34 decimal digits and uses the IEEE754R rounding mode of half-even
Medium	64 bits	IEEE 754R Decimal64 format which is 16 decimal digits and uses the IEEE754R rounding mode of half-even
Low	32 bits	IEEE 754R Decimal32 format which is 7 decimal digits and uses the IEEE754R rounding mode of half-even

Values are only ever held once within the Experian Pandora Database, regardless of how many times they occur across all tables and columns. Experian Pandora tracks the usages of all Values with links to the Columns that use them for efficient navigation and searching. All values can be up to 32,767 bytes in length.

Datatypes

Decimal

Numbers with fractional parts. Experian Pandora stores Decimal values of any precision and scale and with no loss of precision. Experian Pandora standardises all decimals, removing excess leading and trailing zeroes, and standardising decimal whole values such as 32.0 into integers.

Integer

Whole integer numbers of any precision. Experian Pandora standardises all numbers and the rules are configurable to avoid, for example, telephone numbers with leading zeroes being erroneously loaded as Integer Values. Experian Pandora handles scientific notion and leading and trailing signs.

Money

Monetary amounts in any currency using the language setting for each column. The amounts themselves are treated much like Decimal values, allowing full calculation capabilities.

Date/Time

Date and time values are recognised on load and can then be standardised or modified post-load using functions. The language of a column determines localisation rules applied with the date parsing and allows for named months such as May, Jun etc. in multiple languages.

Alphanumeric

Values that do not fit into any of the other categories, i.e. strings made of alphabetic, numeric or other characters. A typical example of an Alpha Numeric value would be a product name.

Automatic

Whilst not strictly a datatype, you will see this symbol often in Experian Pandora. This denotes that the datatype is automatic, and the data will be inferred into whatever types match the content of the data, based on the loading rules applied.

Format Pattern

For every Value loaded, Experian Pandora generates the Format Pattern, which is a template description of what data is contained in a Value. For example, the Value **FRED** is made up of four alphabetic characters, and its Format Pattern would be **AAAA** - as **A** is used to denote an alphabetic character

Format Pattern Code	Meaning
S	a white space character such as a space, linefeed, carriage return or form feed
-	a dash
_	an underscore
9	a digit
A	a letter
€	a monetary symbol such as £, $, €
.	a period
,	a comma
+	a plus sign
/	a forward slash
\	a backward slash
"	a single, double or grave quote
#	a hash (pound) sign
(	an open bracket either (,[, { or <
)	a close bracket either),], } or >
@	an at sign
?	any other character not fitting in the above list

Experian Pandora also represents a simple version of a Format, where multiple characters are reduced to a single and multiple definitions are also reduced, to help find Values with similar, but not identical, Formats. The reduced pattern definitions are:

Simple Format Pattern Code	Meaning
S	one or more white space characters
#	one or more punctuation characters
9	one or more digits
A	one or more letters
?	one or more unprintable characters

Phonetic Pattern

For Alpha Numeric Values, Experian Pandora automatically determines up to two Phonetic Patterns which is a *phonetic encoding*, or a *sound alike* representation of the original value. These group together Alpha Numeric Values that may have been spelled differently or have differing case. Due to its mixed roots, English is well known for having irrational spelling practices. These days, most countries have also accumulated names from all over the world. Experian Pandora returns up to two Phonetic Patterns for each Value analysed, the **English Phonetic** and the **Non-English Phonetic.** In the **English Phonetic** you will see the pronunciation used most widely by English speaking people. The **Non-English Phonetic** will return the pronunciation for the origin of the word. A much wider array of Phonetic patterns are also available including **Soundex**, **Caverphone**, **Metaphone** and much more.

Sources of Data

Experian Pandora supports the loading of data from:

- **Delimited files** (e.g. comma separated, tab separated etc.), with or without headers
- **Excel Spreadsheets** (both Excel 2003 **xls** format and Excel 2007-2010 **xlsx** format)
- Any data source that has a type IV **JDBC driver** – which covers all popular database management systems and file types such as Oracle, SQL Server and MySQL.
- Files with fixed-length, non-variable, non-delimited records

Experian Pandora also accepts **Metro 2®** format files, which are converted into a CSV file format prior to loading.

Experian Pandora provides a great deal of automation in the identification of data content, particularly with delimited files where Experian Pandora will attempt to automatically identify:

- Headings being present on the first line
- The correct delimiter between fields
- The correct end-of-line characters

Loading of data supports any character set including double-byte Unicode character sets and mainframe character sets. It will attempt to identify the correct character set for delimited files by looking for BOM markers and sampling data. It also supports multiple localisations (even individually per Column) which are critical in assessing the minimum and maximum values during analysis.

Values are standardised according to user defined rules and automatically given the correct data type.
Default processing is:

- Numeric data has non-significant zeros stripped off
- Dates are automatically recognised and converted
- Character strings are trimmed of leading and trailing spaces
- Decimal data has leading integer and trailing decimal zeros removed

Standardisation rules used can be changed on a per-column basis, and form part of the metadata for the Column (essentially as a pre-loading transformation rule).

The integrated **Scheduler** provides priority-controlled scheduling of data loads as well as other processes, and includes controls for scheduling out-of-hours and restricting times when activities can occur.

Data File Preparation

If you are extracting data from an existing database into a delimited file format for loading, it is a good idea to ensure that whoever is extracting that data for you does not make the common mistakes that can occur when you export data to a delimited file, such as:

1. Forgetting to put headers on the first line if they were available as an export option
2. Not quoting fields that have delimiters in the data fields themselves.
3. Ensuring a print format for number has not been used, i.e. ensuring numbers do not have separators.
4. Blank lines left in the data. These will appear as rows of nulls.
5. Leaving garbage in the export file. This commonly happens when a DBA writes a script to export a Data file, and leaves the commands or response from the export program in the output file.
6. The Export character set.

Preparing Metro 2® format files

Experian Pandora has the ability to convert Metro 2® format files to CSV format, ready to load into the database.

The header, base portion of the data record and trailer can be in either 366 packed format or 426 character format. Where multiple J1 or J2 segments are present, only the last one will be included in the converted file.

Note: Multiple J1 and J2 segments are supported in Experian Pandora v5.9.6 and later.

The Metro 2® file needs to be copied into a 'metro2' directory, where Experian Pandora will automatically detect it. By default, this directory is in the database root's **\data** directory, for example **C:\Pandora\data\metro2**. You can change this through the **Metro2 Autoconvert Directory** server setting:
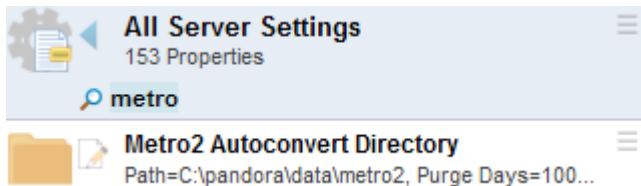
When the Metro 2® files are dropped into the above directory, Experian Pandora will convert them into a CSV file format. The original Metro 2® file will be deleted, and an equivalent CSV file with the same file name and with visibility set to "Everyone" will appear in the import directory, and in Files:
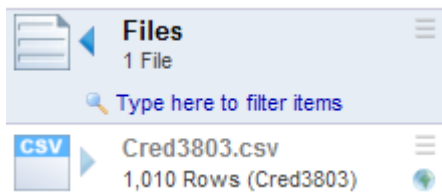

The log will show:

```
2017-08-07 19:29:32 Pandora: Successfully converted \pandora\data\metro2\Cred3801.dat to
\pandora\import\Cred3803.csv - Records Written: 1539, File Format: Character
```

If the file is not valid, Experian Pandora will try to convert it three times, and then fail (it will be blacklisted). The file will remain in the **\metro2** folder. Overwriting the file with another file of the same name will remove the blacklisting, and Experian Pandora will re-try the conversion.

The log will show:

```
2017-08-07 19:17:58 Pandora: The file \pandora\data\metro2\Cred3800.dat is not a valid Metro2 file.
2017-08-07 19:17:59 Pandora: The file \pandora\data\metro2\Cred3800.dat is not a valid Metro2 file.
2017-08-07 19:18:00 Pandora: The file \pandora\data\metro2\Cred3800.dat is not a valid Metro2 file.
2017-08-07 19:18:00 Pandora: The Metro2 file Cred3800.dat has been blacklisted as it has exceeded the
maximum of 3 attempts at conversion
```

Once you have uploaded the Metro2® file, you can configure how the date fields should be loaded by Pandora. By default, date fields in Metro2® files are maintained in raw format, so that the validation rules function as expected. You can toggle this behaviour using the Metro2® Raw Processing server setting. If the setting is set to 'No', dates appear in human-readable format:
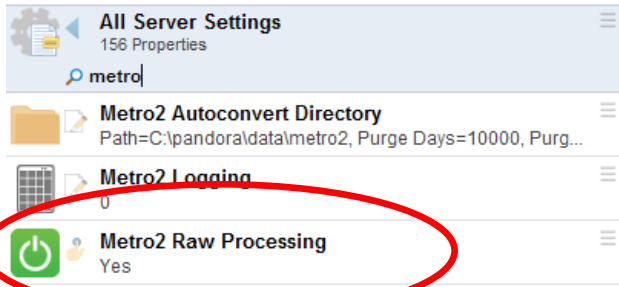
File Upload

Experian Pandora provides the option to upload files from the client workstation and any available network location. This will transfer any selected files to make them available on the server to load. You can select one or more files to be transferred, including associated DDL metadata files if available.

Press the **Upload** button on the Quick Bar, or right-click **Files > Upload Files…** in the Explorer, to open the Upload Files dialog box. Find the files you want to upload:
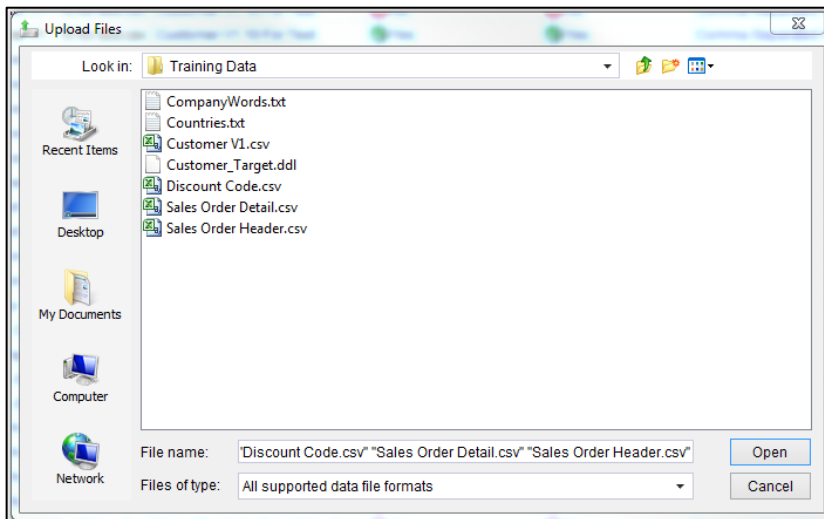

Select **Open** and the Upload Files Visibility dialog will appear.

You can also drag files directly into the Experian Pandora window, bypassing the Upload Files dialog and going straight to the Upload Files Visibility dialog.
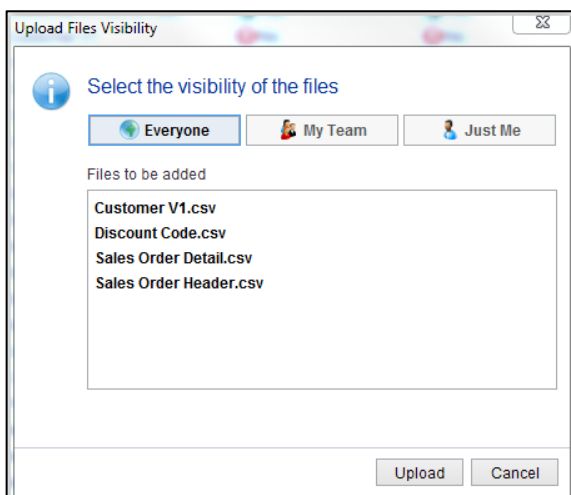
Here you can choose who the files are visible to: Everyone; My Team; Just Me. Select the option required and click Upload to continue. **Note:** Once a file is uploaded the visibility can be changed. On a file you want to change, **Right click > Change Accessibility**. There you can change the visibility.

If the files you are uploading have associated DDL metadata files they will automatically be picked up without you having to select them explicitly for uploading.
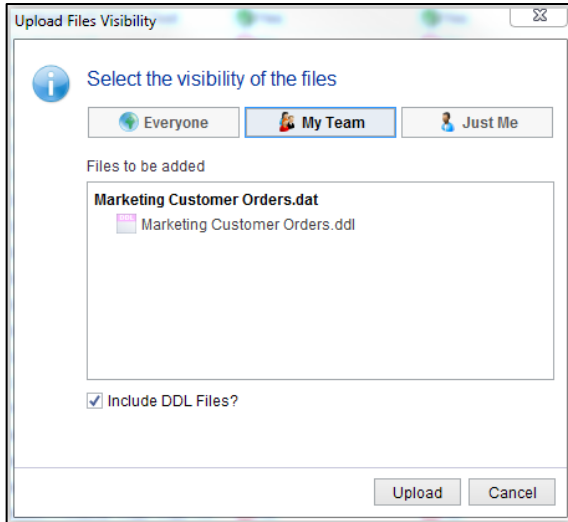

Here only Marketing Customer Orders.dat was selected, but the DDL file was identified and will be uploaded unless you uncheck the Include DDL Files box.

As soon as you press Upload the files will begin being copied to the server. If the files are small you will immediately see popup notification windows informing you that they have been transferred.
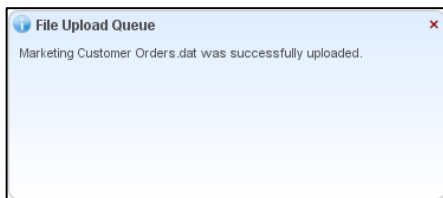
For larger files that take longer to transfer, you will see a red dot showing the count of files on the **Upload** button in the Quick Bar, and the button with the label acts as a progress bar for each file.
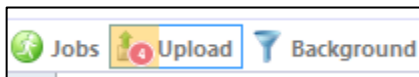
While the **Upload** button is in this state, pressing it brings up the File Upload Queue dialog box instead of the Upload Files dialog.
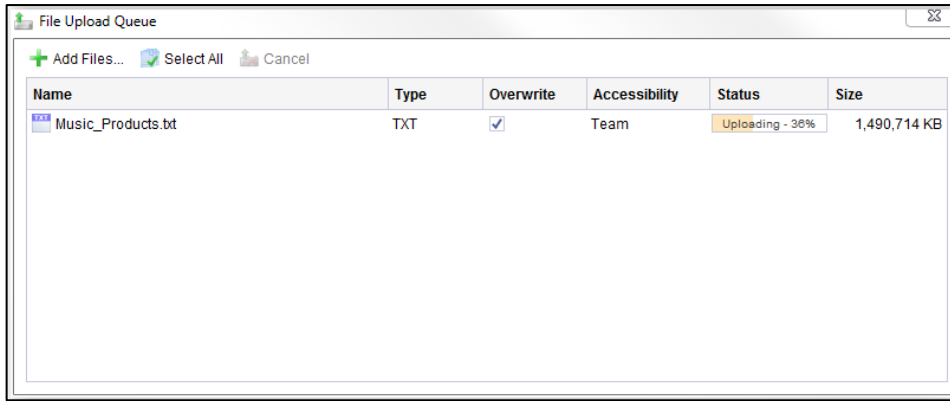
Here you can see the remaining files that are queued to be uploaded. You can select and cancel files if you decide that you don't want to continue uploading them. You can also select **Add Files** to choose additional files to upload.

Open the **Files** drilldown to view the details of the uploaded files:

When Excel files are imported they are split into Worksheets so they are available to Load individually if required. Each sheet will have a separate row in the Files Explorer and Drilldown and show the Sheet name for reference.

.ddl files will be imported with the relevant file and show with a Yes in the column **Has DDL?**

Scroll along the Files Drilldown for information immediately available about the data you are planning to load, or view it in the Explorer under the **File name > Info**.

Modified is a timestamp derived from Files which tells the user when the file being loaded was last changed or modified.

The files are all stored on the Experian Pandora server and are divided into 3 access categories:

- **Everyone** – Visible by all users
- **My Team** – Visible by members of your team
- **Just Me** – Visible only by you

The location of these files is in the *IMPORT* area for Experian Pandora. Public files are in the root of the import directory. *My Files* are in sub directories named after the user, prefixed with "**u_**", and *Team Files* are in sub directories named after the Team, prefixed with "**g_**". In both cases the directory names are suffixed with a "**.dir**" extension.

Loading Files

Files may be loaded from either a local file system or a remote file system. The remote file systems currently supported are Hadoop HDFS, or you can load from Cloud based external file systems, such as Azure Blob and Amazon S3.

Prior to Loading you can preview the data by selecting **Rows (Preview)** on the file you want to Load, or click the file name in the Explorer:
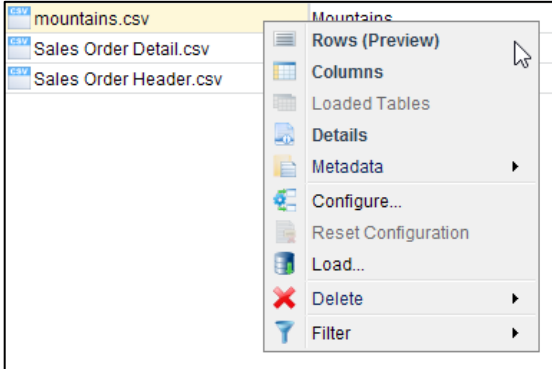
If you wish to modify the column widths to show as much data as possible, select **Fit > Contents**:

Now select load either from the Drilldown Toolbar or from the Explorer and right-click for the menu to open the Load dialog box:

Load as

This is the name the table will be given when it is loaded. If there is already a table with this name, a pop-up will advise you that there is already a table loaded with the same name and give you the option to load as a new version or load as a standalone table. If you opt to load as a new version the file will be loaded as a new version of the existing table with an automatically managed version number assigned. If you opt to load as a standalone table it will be given a generated name composed of the ta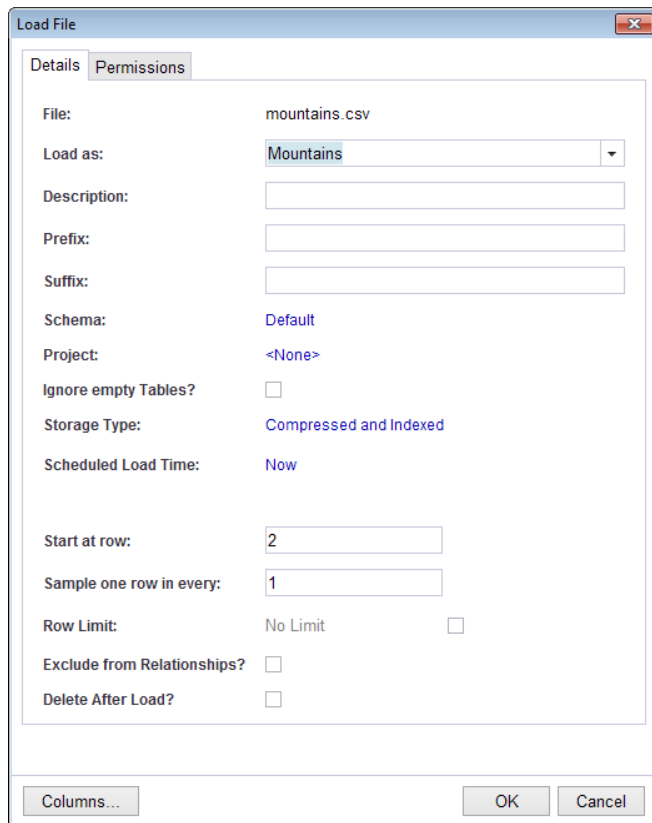ble name and a number. Alternatively you can cancel the load, change the name of the table and it will no longer load as a new version of the table previously loaded but as a completely separate new table with the name you have specified.

Description

This is a user-defined description of the table. You can leave this blank if you do not wish to enter a description.

Prefix

This is text that will be added at the front of the table name when loading.

Suffix

This is text that will be added at the end of the table name when loading.

Schema

Tables will be loaded to the Default Schema unless you select another schema. If a table already had a schema configured, the configured schema will be overridden with the schema that is selected at this point.
A Schema is a sub-division of the Experian Pandora database. A Table can only be a member of a single Schema. If you want to create a new schema, select Create New and enter the relevant details. The table will then load directly to the new schema if selected.

Project

This is the Project into which the loaded tables will be placed. If a table already had a project configured, the configured project will be overridden with the project that is selected at this point.
If <none> is selected the table will not be put into a Project. A Project is an arbitrary sub-division of the Experian Pandora database. A Table can be a member of a many Projects for a multitude of purposes. If you want to create a new project, select Create New and enter the relevant details. The table will then load directly to the new project if selected.

Ignore Empty Tables

This will ensure that the data loader does not create empty tables. You may want to load an empty table, for example, to obtain its structure for other purposes (e.g. Data Mapping).

Storage Type

Select either Compressed and Indexed or Uncompressed and Indexed. Loading as Compressed and Indexed enables the table data that is saved on the server to be compressed after loading, which significantly saves disk space.

Scheduled Load Time

This is the time you want the data load to occur, which defaults to "Now". Even if the load is scheduled immediately, or for a specific time, it may not run at that time if there are outstanding jobs that take priority over it or if the Administrator has specified a limited time during which they can run.

Start at row

This is the row from which to start loading. If the **First Line Is Headings** File Option is set then this will always start at 2 by default (e.g. after the first row).

Sample one row in every

This allows data to be sampled at load time by selecting one row in every *n* rows, either evenly or randomly. Defaults to **1** and **evenly** which will load everything.

Row Limit

This defaults to No Limit. This option is useful if you want to test load a file before committing to load it in its entirety. This limits the load to a certain number of rows only. No more rows than this number will be loaded. If you tick the box on the right hand side, an additional entry box appears allowing the number of rows required for load to be entered.

Exclude from Relationships

This will stop the automatic analysis and discovery of relationships for this table at the load stage. If you wish to discover relationships at a later stage please refer to the section about Relationships and Manual Discovery.

Delete After Load

This is an option to delete delimited files or spreadsheets from the import area after it has been successfully loaded, releasing disk space and preventing the build-up of unwanted files on the server.

Setting Permissions
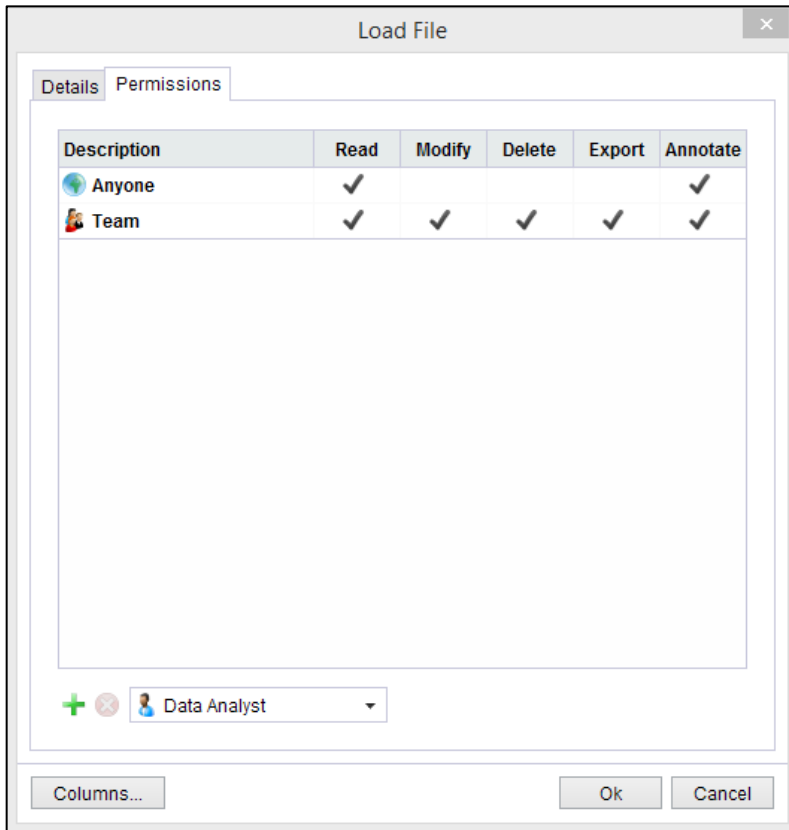
All loads will be given the default system-wide permissions (configured by the Administrator). The default is that the owner and all Users in the Team of the owner will have full access to the loaded table, whilst other users will only have Read and Annotate permissions. Clicking on the **Permissions** allows you to change the permissions given to the Table that is created by the load and allocate specific permissions to user roles.

Load Multiple Files

You can Load multiple tables all together by selecting them in the Files Drilldown, right-click Load Files. The options available are more limited than when loading individual tables. If you wish to change the **Load as** name you will need to do it BEFORE selecting the tables you want to load together and then leave the default to multiple values:
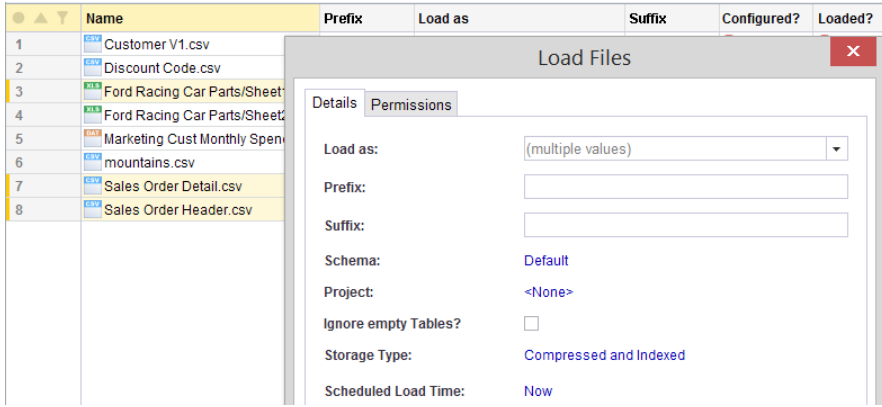
The **Jobs** icon in the Quick Bar will indicate when you have a new Job running. To view the Scheduled Jobs, click on the Jobs icon in the Quick Bar. Alternatively, from the explorer select **System** > **Jobs**. This will open the **Jobs** drilldown.

The user can view the loaded Tables by navigating in the explorer **Data > Tables**. You can view tables that are in a Schema or Project by selecting your Schema/ Project and selecting Tables.

Configuring a File before Loading

As delimited files and Excel spreadsheets are not strongly typed data sources, they may need additional configuration prior to loading. Configuring a file prior to Loading is optional.

Standardisation Options

You can apply standardisation options at load time to determine how Experian Pandora will treat the data in the input source when loading and interpreting data content. Setting the expected datatype of a Column automatically sets the preferred standardisation options for that datatype. You can use **Automatic** as a datatype if you do not know for sure what content is in a Column and Experian Pandora will work it out for you based on the actual values in conjunction with the standardisation options. Loading will be considerably faster if you do know what datatypes are in use and elect to specify the file at load time.

System-wide default Standardisation Rules can be configured in the **Server Settings > Data Loader.**

Configure Table Options: Details Tab

Right-click on the file name in the Explorer or select the Configure icon in the Drilldown Window to open the Configure File dialog.
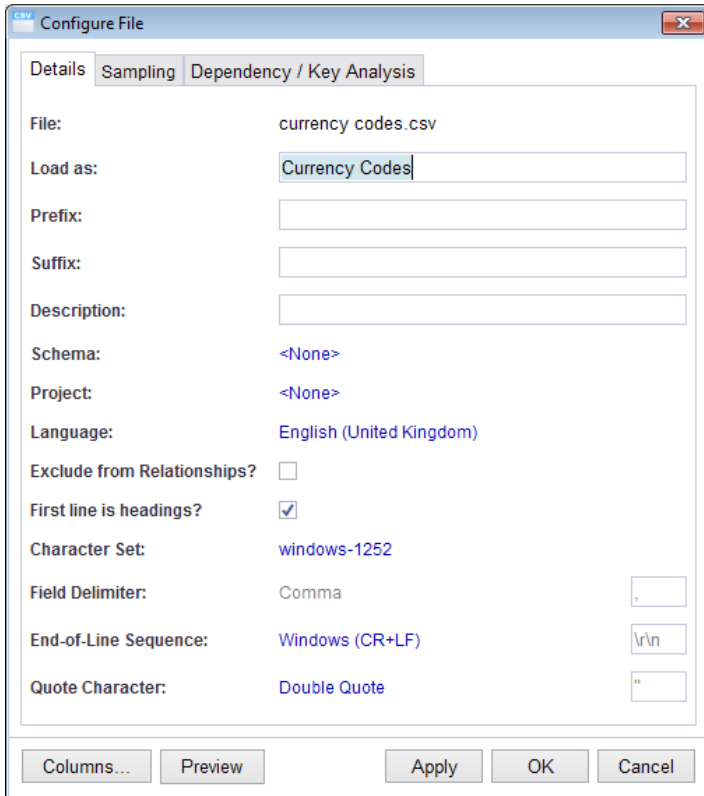
Load as

This is the name the table will be given when it is loaded. If there is already a table with this name, it is assumed that this is a newer version of the data. The file will be loaded as a new version of the existing table with an automatically managed version number assigned. You can change the name of the table at this point and it will no longer load as a new version of the table previously loaded but as a completely separate new table.

Prefix

This is text that will be added at the front of the table name when loading.

Suffix

This is text that will be added at the end of the table name when loading.

Description

This is a user-defined description of the table. You can leave this blank if you do not wish to enter a description.

Schema

Select the Schema you want the table to be placed in. If you have already selected a schema the setting will be carried to the final load. If you have not already created a schema, selecting **New…** will open the Schema creation dialog. When you have saved the new Schema, it will be selected for the table you are loading and you will return to the Load Data dialog. If you change the schema here, it will override any previous selection.

Project

If you have already selected a project the setting will be carried to the final load. As with Schema you can create a new project at this point and add the table to it. If you change the project here, it will override any previous selection.

Language

This is the language to use when interpreting the data, and is applied to all columns (unless overridden for an individual column). Experian Pandora can support multiple languages in a single import file, with a different language per column. This is manually-configured on either a Table or Column basis and ensures that data is profiled and sorted for those columns in accordance with the rules of the language. An example would be ensuring that values starting with umlauts appear next to the letter U, and not after Z. This affects the inferred minimum and maximum values in analysis, as well as for sorting values in a column.

Exclude from Relationships

This will stop the automatic analysis and discovery of relationships for this table at the load stage. If you wish to discover relationships at a later stage please refer to the section about Relationships and Manual Discovery.

First line is heading

The first line of the file optionally contains heading information, i.e. the names of the columns in the file. This is typical of comma-separated-value files, for example those generated by Microsoft Excel.

Character Set

Any character set can be used for loading data, including double-byte character sets, for example UTF-16. It is recommended that the most appropriate character set be used for the data in question. For example, ISO-8859-1 will be appropriate for most US/UK data. UTF-8 will allow most Unicode data to be loaded without resorting to double import file sizes, as UTF-8 will compress to single bytes in most cases. UTF-16 is best left for data sources that require extended character sets, such as Asian data. It is also possible to load UTF-32 data. Experian Pandora tracks all encoding errors on a per Column basis that occur when loading data, which is usually as a result of either corrupted data or the wrong character set being used. It is possible to drill down to any rows containing encoding errors by navigating to the Column in question in the Explorer and selecting **Profile > Encoding Errors**.

Field delimiter

The character used to delimit fields being loaded. Changing this will change the preview display. To preview the updated delimiter in the header row, right-click **Files** in the Explorer and select **Refresh Files**. Multiple options are provided, or a custom delimiter can be entered. Experian Pandora can use any delimiter, defined as a single character, including a character code of 0 (as in the ASCII code 0, **not** the number 0).

End-of-line sequence

A string of characters used to identify the end of each record. Typical options are provided, or a custom string can be entered.

Quote Character

Experian Pandora can use any single character as both a start and end quote for fields. Quoting is necessary if values contain either the field delimiter in use or the end-of-line delimiter to avoid the file reader incorrectly detecting a new field/new record. It is not necessary for all fields to be quoted, just the ones that contain the field or end-of-line delimiters, although Experian Pandora does not mind if all fields are quoted. For fields that actually contain quote characters and are being quoted then the quote character used must be escaped. This is achieved by doubling up the quote character. For example, using the delimiter **'** to surround fields, and one of the fields contains the value **O'Connor**, then the single quote in the value must be escaped to avoid being interpreted as the end of a field – giving **"O'Connor"**.

Configure Table Options: Sampling Tab

Start at row

This is the row from which to start loading. If the **First Line Is Headings** File Option is set then this will always start at 2 by default (e.g. after the first row).

Sample one row in every

This allows data to be sampled at load time by selecting one row in every *n* rows, either evenly or randomly. Defaults to **1** and **evenly** which will load everything.

Row Limit

This defaults to No Limit. This option is useful if you want to test load a file before committing to load it in its entirety. This limits the load to a certain number of rows only. No more rows than this number will be loaded. If you tick the box on the right hand side, an additional entry box appears allowing the number of rows required for load to be entered.

Dependency / Key Analysis Tab

Experian Pandora gives you the option to discover Dependencies and Keys at the point of loading the data. Please see the chapter Dependency and Key Analysis for information about Dependencies and Keys.
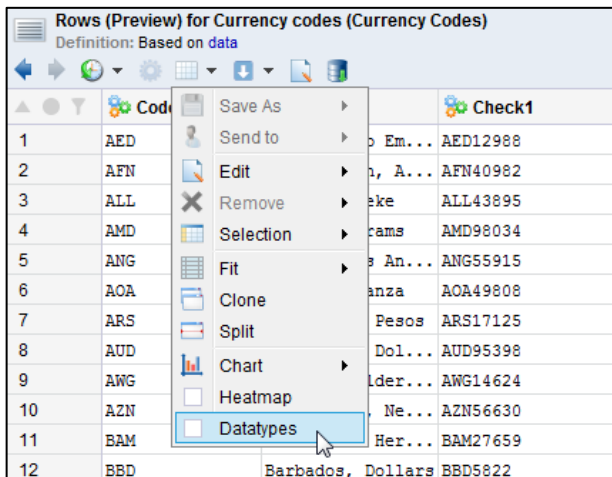
Column Options

The Column Options allows the options for each column to be picked individually. Experian Pandora automatically analyses the data based on rows at the start of the file to ascertain the appropriate data type for each column. This is not necessarily correct as it is only based on a small sample of data, but usually is correct. By setting the expected data type, the data is parsed appropriately on screen, and an appropriate icon is placed on the column heading. Standardisation rules are automatically chosen for the selected data type. If you are unsure what data is in the column, set the data type to **Automatic**, and ensure the standardisation rules are selected for all types you wish to detect. Experian Pandora allows mixed datatypes within a single Column.

System-wide default Standardisation Rules can be configured in the **Server Settings > Data Loader.**

Configuring Column Details

Each column has an icon to signify the data type that Experian Pandora thinks best suits the data from a small initial sample of the file. This is used to simply set default standardisation options applied to every column, which can be overridden at the individual column level. It is good practice to check the preview to see if it looks right.

To view the datatypes, click on the Drilldown Action Menu on the Preview Drilldown window toolbar and select **Datatypes**:

From the Configure dialog for the File on the bottom left hand side select **Columns** to open the Configure Columns dialog:
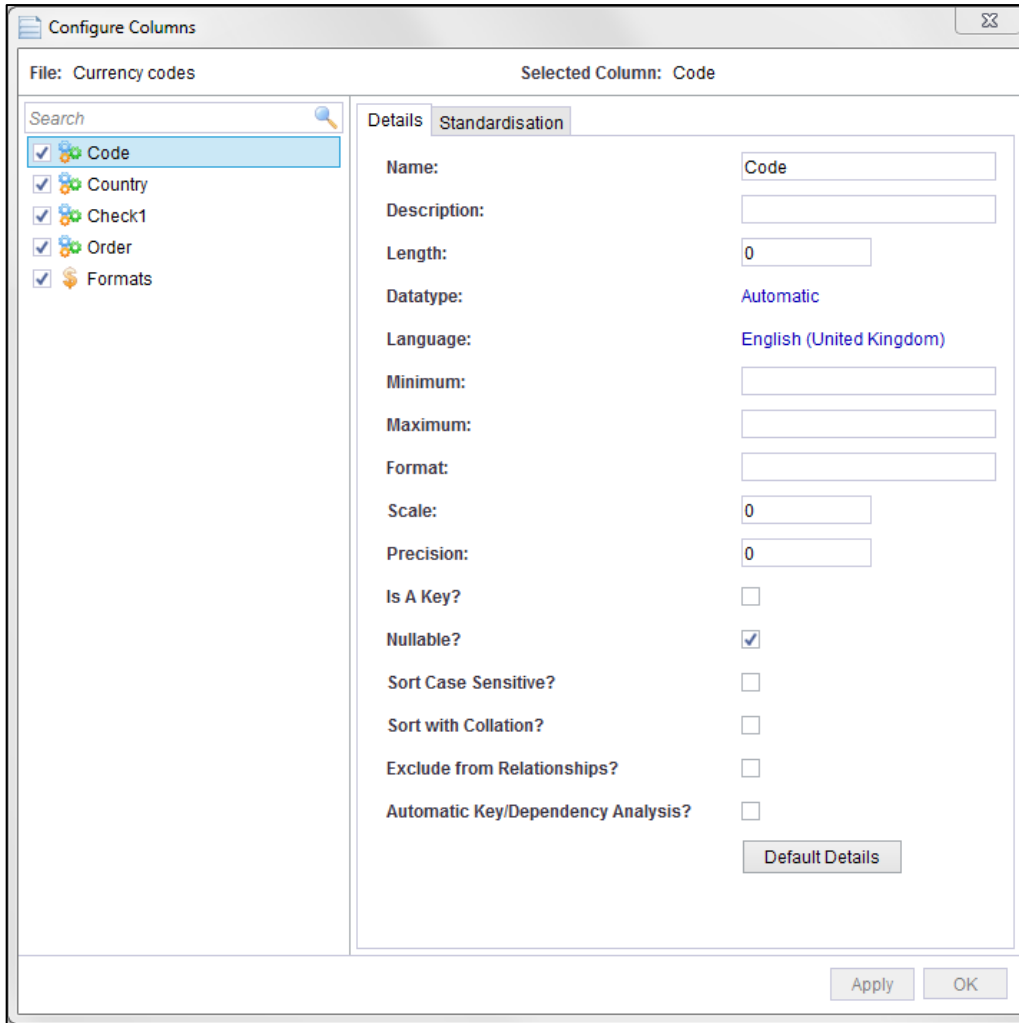
In the column Configure dialog you can choose which columns you want to load. Any that are un-ticked will not load in the table. They will disappear from the preview when you hit the apply button. The data is still available but will not load.

Each column can be individually configured or you can select two or more columns and apply settings. Those that are greyed out cannot be altered en-mass e.g. changing the column name and description.

Name
This is the name of the column, which must be unique within a Table. Experian Pandora will ensure this on load by modifying name clashes accordingly with a sequence number.

Description
The user-defined description of the column. This is optional and will appear in the Explorer with the column name.

Length
The maximum length in characters of data in the Column

Datatype
The expected data type describing the content of the Column. This is purely documentary and has no effect on how the data will be loaded, apart from pre-selecting the Standardisation options according to the selected data type, all of which can be changed at will.

Language
The language used for processing values in the Column. Ensures regional specifics are applied to comparison and sorting.

Minimum

The smallest expected value in the column

Maximum

The largest expected value in the column

Format

The expected format pattern for values in the Column

Scale

The documented scale of decimal values.

Precision

The documented precision of decimal or integer values

IS a Key?

Indicates whether key checking should occur by looking at how unique the values in the column are.

Nullable

Indicates whether the expected content can contain nulls or not. This is documentary and has no effect on loading.

Sort Case Sensitive

Whether case sensitivity should be ignored when sorting values in the Column

Sort with Collation

Whether collation should be used, i.e. the rules of the defined language, when sorting values in the Column

Exclude from Relationships

This will stop the automatic analysis and discovery of relationships at the load stage. You need to select each column that you want to stop the automatic analysis and tick the check box. If you wish to discover relationships at a later stage please refer to the section about Relationships and Manual Discovery.

Automatic Key/Dependency Analysis

If ticked, this column will be included in the automatic analysis of keys and dependencies when the table is loaded. Changing this will be reflected in the Dependency/Key Analysis tab of configuring the complete file.

Standardisation
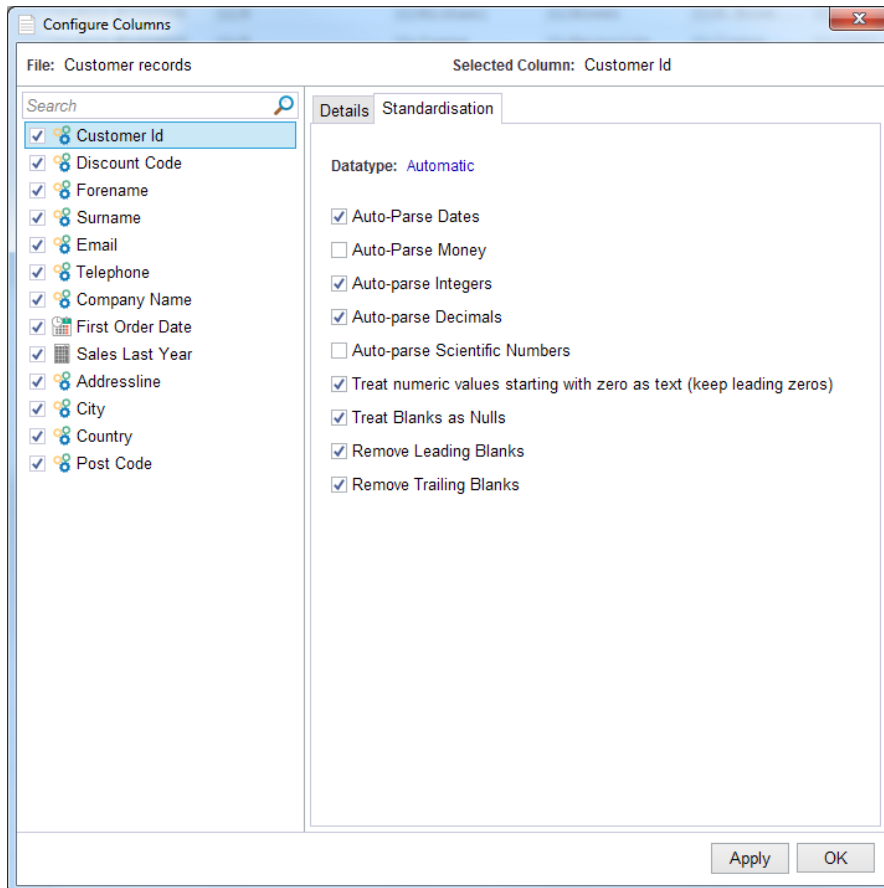
Auto-parse Dates
Automatically parses dates using the language for the Table/Column and the default date parsing rules.

Auto-parse Money
Automatically parses monetary values identifying any valid currency code or symbol.

Auto-parse Integers
Automatically converts numeric values with no fractional part to Integers. It is not recommended to disable this option if the data in question is actually integer numbers.

Auto-parse Decimals
Automatically converts numeric values with fractional parts to Decimals. It is not recommended to disable this option if the data in question is actually decimal numbers.

Auto-parse Scientific Numbers
Converts values in Scientific Notation, such as "1E47", into numbers.

Treat numeric values starting with zero as text (keep leading zeros)
Treats values with leading zeros as a number, removing all leading zeros in the process. Switch this off if loading telephone numbers, bank sort codes or account numbers etc.

Treat Blanks as Nulls
Treat all blank values as NULL (e.g. empty). If not, blank values will be loaded verbatim. It is recommended this option be enabled and that blanks are treated as NULL.

Remove Leading Blanks
Remove leading white space from values. It is recommended this be always enabled.

Remove Trailing Blanks

Remove trailing white space from values. Note that if **Treat Blanks as Nulls** is turned off, and both **Remove Leading Blanks** and **Remove Trailing Blanks** are turned on, then Experian Pandora will ensure that at least a single space is retained for entirely blank values. We generally recommend that if it is important to distinguish between NULL and BLANK, that you ensure that trailing and leading blanks are removed, otherwise Experian Pandora will load values with varying spaces, which from a business point of view is generally meaningless, and will treat them as different values.

Once all your configurations have been completed, hit the OK button. The file will now show as Configured in the Files Drilldown and is ready to Load.

In the Files Drilldown we can now see that Mountains was Loaded without being Configured and Currency Codes was Configured prior to Load.

Name	Prefix	Load as	Suffix	Configured?	Lo
Currency codes.csv		Customer V1		❌No	❌
Ford Racing Car Parts/Sheet1.xls		Discount Code		❌No	❌
Ford Racing Car Parts/Sheet2.xls		Sheet1		❌No	❌
Marketing Customer Orders.dat		Sheet2		❌No	❌
mountains.csv		Marketing Cust Monthly Spend		❌No	❌
Sales Order Detail.csv		Mountains		❌No	❌
Sales Order Header.csv		Sales Order Detail		❌No	❌

If you have made an error in your Configuration you can right-click on the file name and select **Reset Configuration.** This will remove everything you have done to configure the data.

Name	Prefix	Load as	Suffix	Configured?	Lc
Currency codes.csv		Customer V1		❌No	❌
Ford Racing Car Parts/Shee		t Code		❌No	❌
Ford Racing Car Parts/Shee				❌No	❌
Marketing Customer Orders				❌No	❌
mountains.csv		g Cust Monthly Spend		❌No	❌
Sales Order Detail.csv		ns		❌No	❌
Sales Order Header.csv		rder Detail		❌No	❌

Context menu:
- Rows (Preview)
- Columns
- Loaded Tables
- Details
- Metadata ▶
- Configure...
- Reset Configuration
- Load...
- Delete ▶
- Filter ▶

If you realise you have made an error after having loaded the table, you can delete the table from Tables in the Explorer (if you have delete permissions) then return to Files, reconfigure and reload.

Individual and Additional Column Configuration

Each column can be individually configured in the Explorer or in the Rows (Preview). Open the **Rows (Preview)** in a drilldown window. You can make simple changes to the columns using the right-click menu option in the Explorer on the Column name icon, name or menu icon in the Columns of the table and see the changes in real time in the drilldown window.

Select **Configure** to open the dialog box for the column:

Click Apply and check the details in the Rows (Preview).

Let's change **County** to Alphanumeric, exclude **Code,** write a description for **Check1** and move the column name and description **Check 1** down. The column name **Code** now shows a red cross and is no longer in the preview and will not be loaded and the Column name for **Check 1** has moved one to the right and the column name for **Order** has moved one to the left. This function is for use where column names are in the wrong place – it <u>does not</u> move the contents of the column. **New** and **Discard** work in the same way.

Once you have made your changes to the configuration you are ready to Load the data.

Augmenting Metadata with DDL

Data can be optionally and automatically augmented with column names, documented lengths and datatypes using ANSI DDL (Data Definition Language) files if they are present in the import directory. These must have the same name as the data file being loaded with an extension of **ddl**. To see if a file has a .ddl associated with it, scroll along the Files Drilldown and check the **Has DDL?** Column. Care should be taken using this mechanism to load data.

To view the metadata right-click menu and select **Metadata > DDL** or **Copybook**.
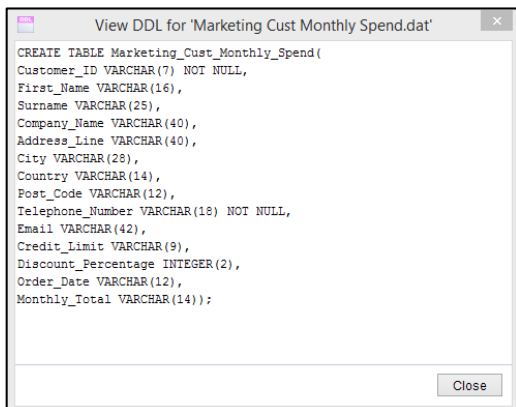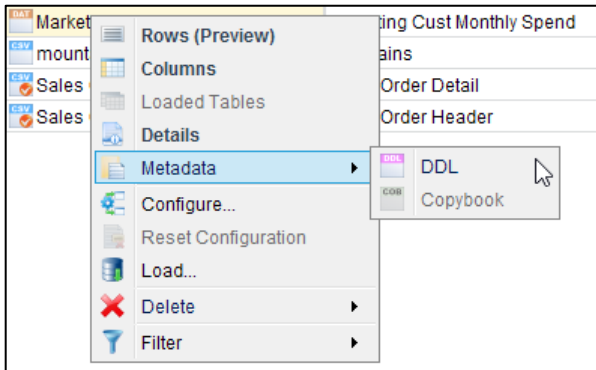
Documented Metadata

The documented metadata allows an analyst to set the expected characteristics of a Column. This information can be used in Validation Rules to automatically highlight Columns where the actual characteristics of the data do not match the documented ones.

Null and Space values

Experian Pandora has multiple options for the handling of blank values and null values as follows:-

1. Trailing spaces can be optionally automatically removed from values
2. Leading spaces can be optionally automatically removed from values
3. Values composed entirely of spaces can be compressed to no spaces (NULL), a single space, or left as they are as a string of variable space.

When you are loading values that consist of multiple spaces, if all of the above options are turned off then this will result in storage of values composed of one or more spaces. These will always be displayed as blank values in the various drilldown views, but hovering over them in the view will indicate their length (i.e. how many spaces). Null values can be displayed as *<Null>* or as blank depending on user preferences. Post load transformation operators to convert null to space and vice-versa can also be used. Experian Pandora tracks null counts automatically and space counts by virtue of the fact that it counts value usages automatically (i.e. it treats a space value as a value unless you instruct Experian Pandora to treat blank values as NULL values on load).

It is **not** recommended that varying space values are loaded as there is unlikely to be a sensible business or technical reason for storing such data, and you will lose the ability to easily relate fields of varying space together. It is recommended that if space-filled values mean something different to NULL values that these are loaded and compressed to a single space. If varying space values have some business meaning they are best translated into something more meaningful post-load using the transformation capabilities in Experian Pandora.

Uploading Metadata Files

As well as uploading metadata together with the data files they are associated with, Experian Pandora can upload metadata files separately. These metadata files can then be used to describe the target layout for Data Mapping.

To upload metadata files right-click **Files > Upload Metadata…** in the Explorer to open the Upload Metadata Files dialog box.
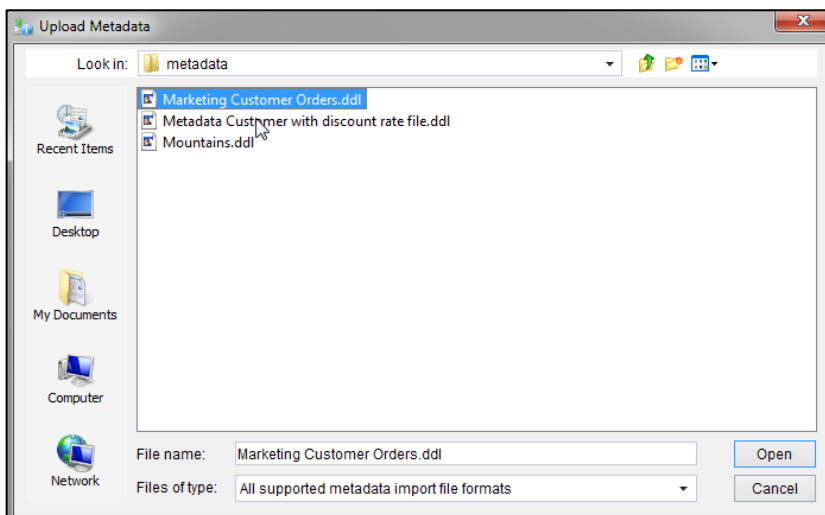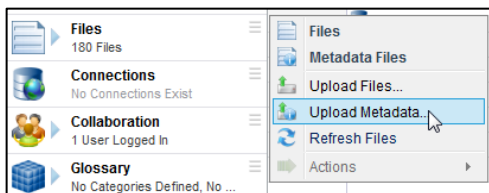
Select **Open** and the selected files will be uploaded. If Experian Pandora currently has a queue of files being uploaded, the metadata files will be added to the end of the queue. This can be seen by pressing the Quick Bar **Upload** button if it is showing a red dot and process bar.

When the upload is complete, the metadata files can be viewed in the **Metadata Files** drilldown. This is accessed by right clicking **Files > Metadata Files** in the Explorer.

Metadata Files
Last Updated on 29 July 2015 17:31:41 BST

4 Rows

	Filename	Type	Size	Created	Last Modified
1	Marketing Customer Orders.ddl	DDL files (*.ddl)	497	29-Jul-2015 15:49:56	29-Jul-2015 16:31:41
2	Metadata Customer with discount rate file.ddl	DDL files (*.ddl)	439	29-Jul-2015 15:49:56	29-Jul-2015 16:31:41
3	Mountains.ddl	DDL files (*.ddl)	177	29-Jul-2015 15:49:56	29-Jul-2015 16:31:41
4	remesas.cob	COBOL files (*.cob)	1174	29-Jul-2015 15:49:56	29-Jul-2015 15:49:56

Loading data from Excel Spreadsheets

This is very similar to loading data from Delimited Files, except that Experian Pandora splits the individual Worksheets in an Excel Workbook so that they can be configured and loaded individually if required.

The sheets are easily identifiable as the **Filename** will be the same for multiple differently named sheets. The sheet name is displayed with the File and unless you Configure the file it will load as the Sheet Name :

Name	Prefix	Load as	Suffix	Configured?	Loaded?	Type
BIG_FLAT.dat		Exa01 Domic Pers Int		❌ No	❌ No	Fixed-Length File (.dat)
Currency codes.csv		Currency Codes		✅ Yes	❌ No	Comma Separated Values File (.csv)
Customer V1.csv		Customer V1		❌ No	❌ No	Comma Separated Values File (.csv)
Discount Code.csv		Discount Code		❌ No	❌ No	Comma Separated Values File (.csv)
Ford Racing Car Parts/Sheet1.xls		Sheet1		❌ No	❌ No	Excel 2003 File (.xls)
Ford Racing Car Parts/Sheet2.xls		Sheet2		❌ No	❌ No	Excel 2003 File (.xls)
Marketing Cust Monthly Spend.dat		Marketing Cust Monthly Spend		❌ No	❌ No	Fixed-Length File (.dat)

Follow through the standard steps for configuring and/or Loading.

If you delete any of the sheets relating to an Excel Workbook from **Files**, ALL related sheets will be deleted.

Hadoop

Hadoop Overview

Experian Pandora supports files stored in a remote HDFS file system. These are made accessible by configuring a Hadoop connection and are treated as Read-Only files.

Supported File Types

- Comma Separated Values files with an extension of .csv
- Fixed Length files with an extension of .dat
- Delimited Files with an extension of .txt

Creating a Hadoop Connection

Navigate to **Home -> Connections** and right-click on **Connections**, select New… to add a new connection.

In the Create Connection dialog select the **Hadoop HDFS** Adapter from the drop down list of available Adapters.

Not all fields are mandatory, however you must supply at least a **Hostname** and **Port**. The Hostname is the IP address or DNS name of the Primary NameNode of your Hadoop cluster. The port is a site specific Hadoop property. Typically you will find this in the Hadoop configuration file core-site.xml

If you are not familiar with Hadoop configurations ask your Hadoop administrator to provide this information.
Below is a minimal example of the core-site.xml. You will find the HDFS port number under the property fs.default.name

```
<configuration>
 <property>
  <name>fs.default.name</name>
  <value>hdfs://0.0.0.0:9000</value>
 </property>
</configuration>
```

You will usually require a **Username** and **Password** to access the files stored on the Hadoop cluster.
This username and password are the Linux username and password on the Hadoop host machine. File access permissions for your username will determine which files are visible to Experian Pandora. Only those to which you have read access and are supported file types will be visible.

The **Root Directory** is the starting directory from which to start the file discovery. If it is not specified it will default to the root of the HDFS file system ("/"). You should choose the starting directory carefully since HDFS file systems may contain a huge number of files. You also have the choice of recursing through all the directories below your starting directory. If you wish to include all other directories below your starting directory then tick the box **Flatten.**

When Experian Pandora connects to Hadoop it will list the available files in the Explorer pane. You may configure these files individually prior to loading them.

Create Connection

Details	**Properties**	**Permissions**

Name:	Hadoop cluster
Description:	Hadoop cluster
Adapter:	Hadoop HDFS
Username:	pandora
Password:	•••••••
Hostname:	192.168.0.99
Port:	9000
Root Directory:	/
Flatten:	☑

Save   Cancel

Connections
2 Connections
🔍 Type here to filter items

Hadoop
23 Hdfs Files

Hadoop
14 JDBC Tables

➤ Drop items here to pin them

Connections
2 Connections
🔍 Type here to filter items

Files
Last Updated on 10 June 2015 09:04:24 BST

32 Rows

Hadoop
23 Hdfs Files

Hadoop
14 JDBC Tables

▲ ● ▼	Name	Load as	Configured?	Loaded?
1	Applications.csv	Applications	❌ No	❌ No
2	CCustomerV1.csv	Ccustomerv1	❌ No	❌ No
3	Customer V1.csv	Customer V1	✅ Yes	✅ Yes

File Upload from external File Systems

You can load files from external cloud-hosted File Systems, by configuring settings in Server Settings. Configuring these settings will allow you to load from Azure Blob and Amazon S3 external file systems.

Press the **Upload** button on the Quick Bar, or right-click **Files > Upload Files…** in the Explorer, to open the Upload Files dialog box. If you have more than one File System to choose from, the dropdown menu will look like this:
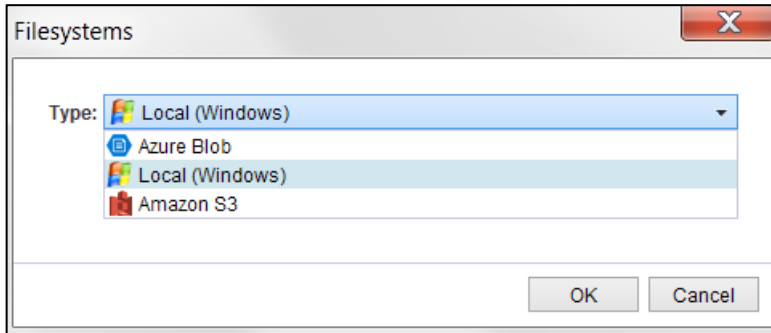
Select a File System from the drop down menu, and it will display a list of available files to upload. Note that the same set of file types (.csv, .txt, .dat etc.) are supported in all file systems, just like with Local (Windows).
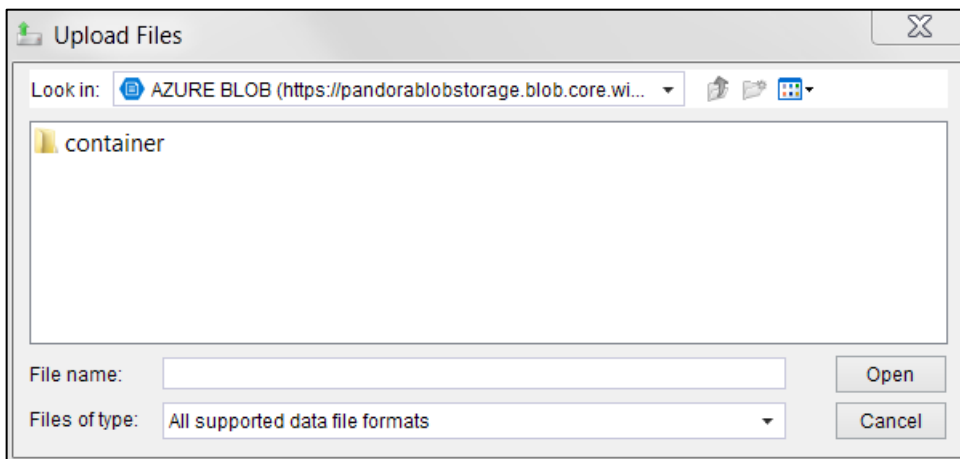
After choosing files to upload, loading the files into Pandora is exactly the same as loading locally into Pandora.

Configuring the external file systems

To be able to use these external file systems, navigate to **Home > Settings > Server Settings > Filesystems**. Here you will find two settings, one called Amazon S3 and another called Azure Blob. **Right click > edit** on the setting you wish to configure.
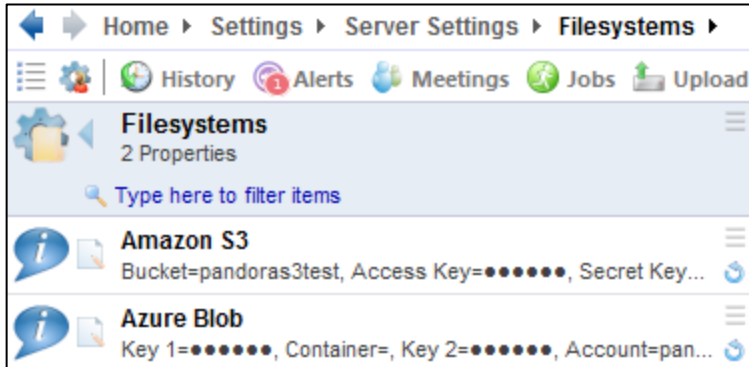
Amazon S3 storage

The following items are required in the S3 configuration:

- An S3 region (this should be the Region, e.g. eu-west-1, rather than the region name, e.g. Ireland).
- The accounts Access key
- The accounts Secret key
- An optional Bucket name

You can find the above information in your AWS Portal
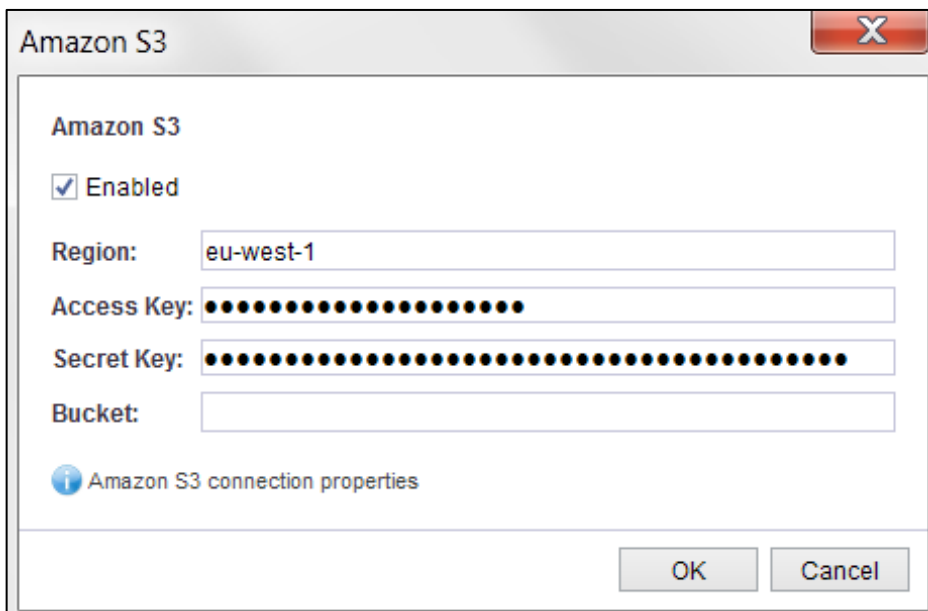
Azure Blob storage

The following items are required in the configuration:

- The Account name
- The primary Key
- The secondary key
- An optional Container name

You can find the above information in your Azure Management Portal

Loading data from an external database (JDBC)

Loading data from an external JDBC compliant database such as Oracle or SQL*Server or for access to a particular database, there must be a JDBC driver installed for that type of database, and the Data source appropriately configured by the Administrator to access the remote database.

Configuring JDBC Connections

Most database vendors provide a JDBC (Java Database Connectivity) driver for their database products. These drivers are usually available on the vendor's support websites. Follow the steps in the sections below to prepare Experian Pandora for loading from a database.
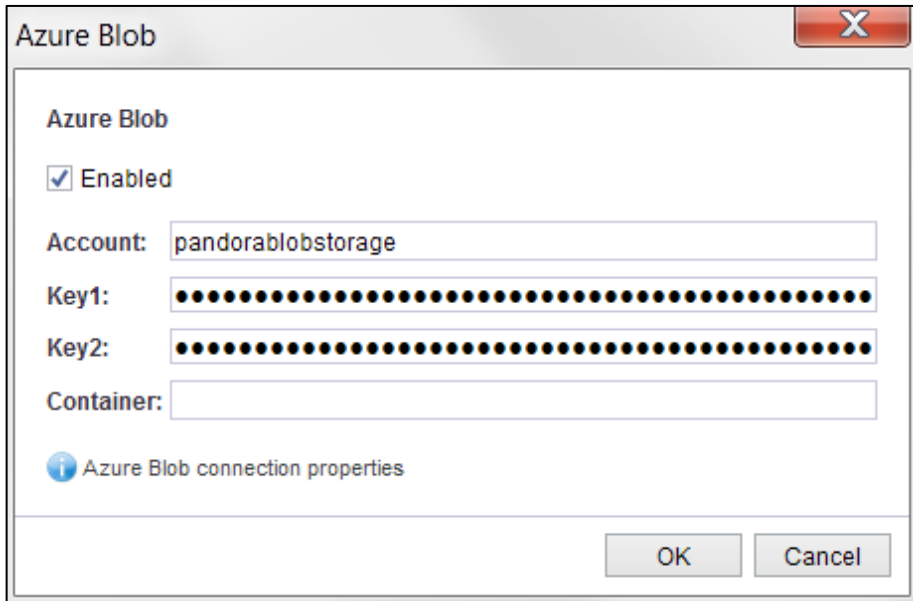
Install the Required JDBC Driver

As above, obtain the required JDBC driver for your database. Ensure that it is compatible with the database version you are using. The database vendor will usually make the version Dependencies extremely clear. For use with Experian Pandora the driver must be compatible with JDK version 6.

After you have downloaded the driver file (these usually have a file extension of **.jar**) the System Administrator will need to install it by copying it to the Experian Pandora database subdirectory "**drivers/jdbc**".

Defining a JDBC Data source Connection

You must first be logged into Experian Pandora as a user with Administrator capabilities. Right-click on the **Connections** item from the explorer, select **New…** and you will be presented with the **Create Connection** dialog. The fields present in the **Connection** tab of the dialog will vary with the different types of Adapter chosen.

Name
This is the name of the data source as used within Experian Pandora.

Description
This is a general free-text description about this data source.

Adapter
This is the driver definition type to use. Experian Pandora provides some standard types, such as SQL Server.

Username
This field is used to specify the username used to log into the remote database.

Password
This field specifies the password for the user to authenticate log on to the external database.

Hostname
This field specifies the IP address or hostname for the server running the JDBC database.

Port
The port number for the TCP/IP service to connect to on the remove JDBC database server

Schema
The name of the database Schema to use

Force Schema Name
Forces the server to always specify the schema name in any SQL statements sent to the database. This should not normally be required. Contact your support representative for assistance on this.

Table Pattern

The pattern to use when listing tables in the remote database catalogue. When using Table Patterns, '**%**' indicates a multiple character wildcard and '**_**' for single character wildcard

Filter Regular Expression

A regular expression to filter tables viewable in the Connection

Filter is Case Sensitive

Indicates whether the filter is case sensitive

Include Tables

Retrieve Table information from the remote database catalogue

Include Views

Retrieve View information from the remote database catalogue

Instance Name

The name of the database instance on the remote JDBC database server
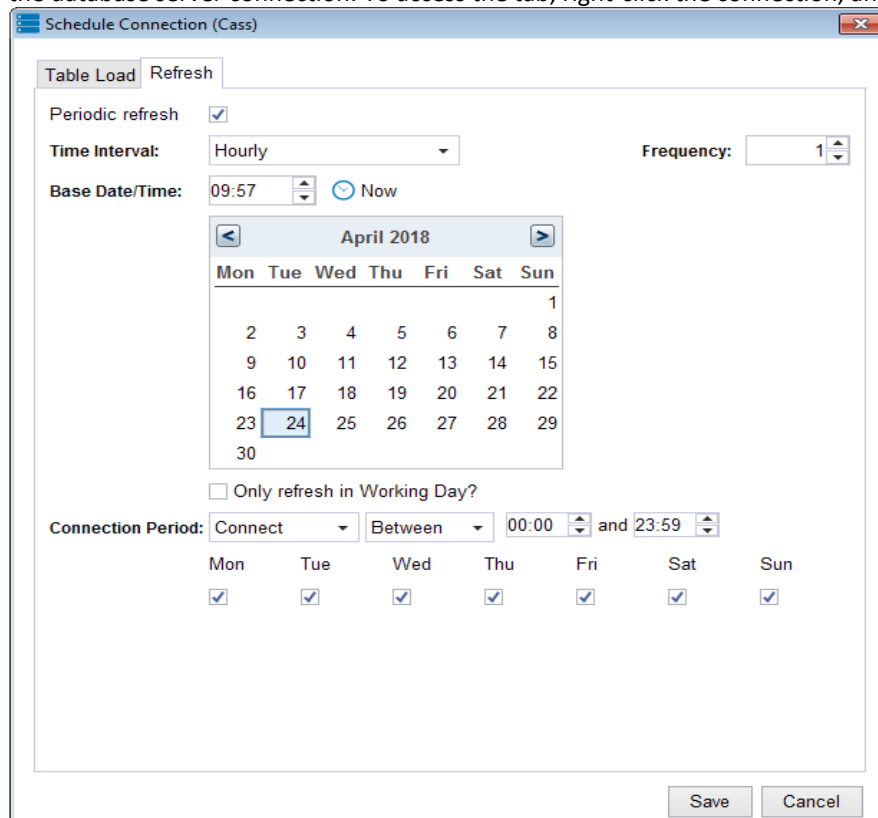
Database

The name of the database on the remote JDBC database server

IMPORTANT NOTE: If the list of tables presented is empty, then it may well be that a username prefix is needed or explicit permissions need to be given to the database user to list the catalogue by the Database Administrator for the remote database being connected to.

Specifying the Data Source Refresh Period

The **Refresh** tab on the **Schedule Connection** dialog allows you to specify how often the data source will refresh its list of tables from the database server connection. To access the tab, right-click the connection, and choose **Schedule Load**…

The fields above are described in more detail below:

Periodic refresh

Allows you to specify whether the Connection is refreshed manually or periodically. When the checkbox is selected, the remaining fields define how often the data source will refresh its list of tables. When the checkbox is not selected, the data source will only refresh its list of tables via the **Refresh Now** option (available on the Explorer as a right-click menu action or from the Connections drilldown).

Time Interval

The time interval by which refreshes will be attempted. Can be one of:

Time Interval	Description
In intervals of minutes	Every n minutes
Hourly	Every n hours
Daily	Every n days
Weekly	Every n weeks

Frequency

The number of Time Intervals between refreshing.

Base Date & Time

The starting date and time from when to start checking.

Only refresh in Working Day

Disables refreshing outside working day hours.

Connection Period

Allows you to specify the time when to connect to or disconnect from the remote database. When set to **Always** it will attempt to maintain a permanent connection/disconnection to the remote database. When set to **Between** you will be able to specify the days of the week and the start and end time for when the Connection is connected or disconnected. The start time specifies when it will attempt to connect/disconnect and the end time specifies when it will try to disconnect/connect (which means that it is possible to specify a start time that is later in the day than the end time).

Automating the Loading of Tables

The **Table Load** tab (which is only available when editing an existing JDBC, not when creating a new one) allows you to specify rules for automatically re-loading one or more database tables at regular intervals, and contains the following fields:

Schedule Definition File

(Previously known as Trigger File) The name of an XML schedule definition file that defines the list of tables to be automatically reloaded. If the filename is not absolute (e.g. not preceded by a drive letter or file separator) the file is assumed to exist in the main import directory.

To the right of the Schedule Definition File field is a **Configure** button that, when clicked, will present you with the following dialog which will allow you to add or add or remove table load entries from the schedule definition file:
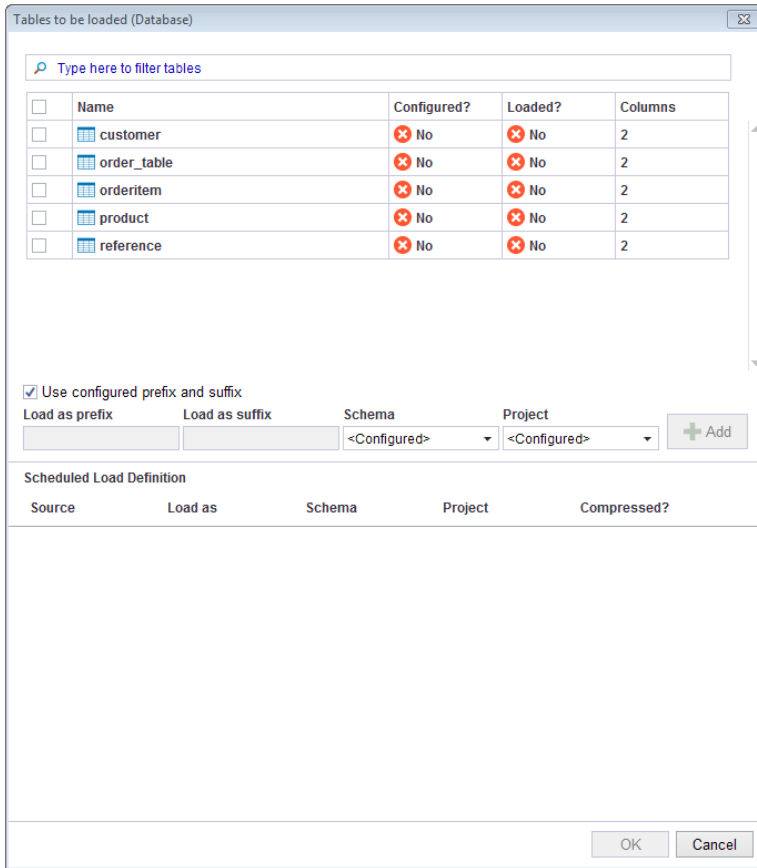
To add new tables to the schedule definition file so that they will be loaded periodically:

1. Select the tables you wish to add to the schedule definition file by ticking the checkbox for those particular tables. You can select as many or as few as you'd like.
   If the connection has a large number of tables, it may be helpful to filter the list by entering keywords in the 'Type here to filter tables' search box.
2. Once you have selected your tables, you can choose to use the prefix and suffix already configured on the table, or override the prefix and suffix by unticking the 'Use configured prefix and suffix' checkbox. You will then be able to enter your own prefix and / or suffix which will apply to the "Load as" name of all selected tables.
3. Select the schema and project for the tables you have selected or alternatively, leave this as "<Configured>" in order to use the currently configured schema and project for each table.
4. Once you are happy with your selection of tables and the configuration, press the "Add" button to add the tables to the schedule definition file. Note that the changes will not take effect until you have pressed the OK button on the dialog.

The following example shows a schedule definition file definition that would, at the specified interval (see below):

5. Load the table **customer** and give it the name **Customers**
6. Load the table **reference** and place it in the **Ref** schema
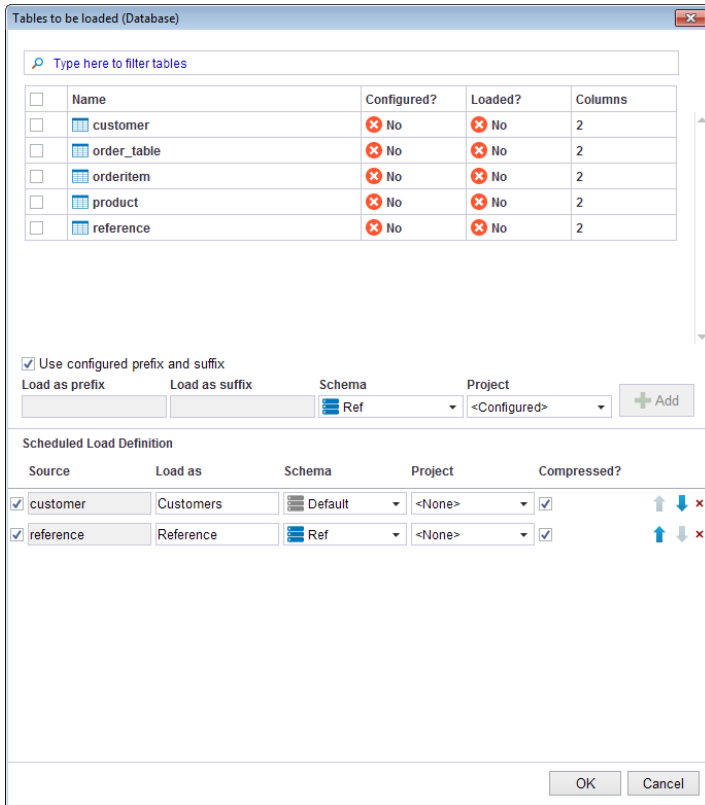
The **Source** field specifies the name of the database table to reload.

The **Load as** field allows you to optionally specify the name by which the table should be loaded and the **Schema** and **Project** fields specify the Schema/Project into which the table should be loaded.

Compressed? When this is ticked, the table when loaded will take up less memory. If un-ticked, the table will not be compressed. The advantage of this is that it is slightly quicker to access the data in the table, at the expense of disk space.

Tables will be loaded in the order that they are specified in the definition. This order can be modified using the up and down buttons on each table load entry.

It is also possible to specify the schedule definition file in **JSON** (JavaScript Object Notation) format. Please contact your Support representative for further details.

Only reload when Schedule Definition File has changed

If you enable this checkbox, reloading will only occur if the last modified time of the schedule definition file is later than the last time reloading occurred.

Time Interval

The interval for which reloads will be attempted. Can be one of:

Time Interval	Description
In intervals of minutes	Every n minutes
Hourly	Every n hours
Daily	Every n days
Weekly	Every n weeks
Monthly	Every n months

Frequency

The number of Time Intervals between reloading.

Base Date & Time

The starting date and time from when to start checking.

Only reload in Working Day

Ignores any time intervals that arise on non-working days to ensure activity only occurs on a working days.

Create the Data source

After all the required fields have been populated in the **Create Connection** dialog, the final step is to save the new connection.

Previewing a table

The number of rows that is displayed when previewing an external table depends on the time it has taken to load the table preview.

You can change the maximum amount of time in milliseconds that is spent loading the table preview using the **JDBC Maximum Preview Load Time** setting. Any rows loaded within this time are displayed.

You can also configure the maximum number of rows displayed in the preview before the maximum preview load time elapses using the **JDBC Maximum Preview Row Count** setting**.** The default is zero, which represents the maximum number of rows obtainable within the maximum preview load time.

Directly accessing external data

All available external data sources are accessed through the **Connections** explorer item.
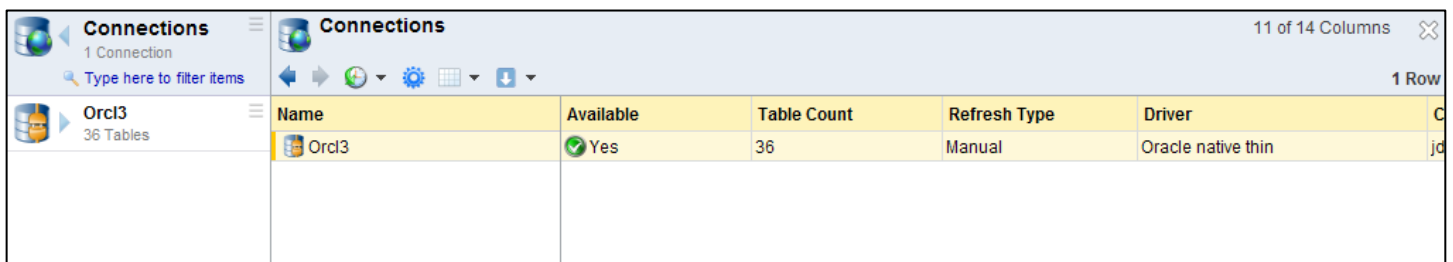
The drilldown provides additional information such about the data source, including the number of files or tables available, and connection information or status in the case of Database Tables. The **Available** column relates to only those data sources that are remote and require monitoring by Experian Pandora. Select one of the available connections to see the list of all available tables in the remote database. Select one or more of them to load.

Connection Monitoring

Particularly important for remote data sources, such as JDBC, Experian Pandora provides background processing for all data sources and background connection management and catalogue querying. This happens asynchronously and incrementally for JDBC databases.

For example, if you are connecting to a database with 10,000 tables over a wide area network, it could take over an hour to retrieve the catalogue information alone! You can simplify this problem by using multiple data sources that have different Table Patterns, to limit which tables appear in which data source.

Experian Pandora connects asynchronously and constantly monitors the connection. If it is deemed available it reads the catalogue of tables, and regularly updates it in the background.

The data source is shown as connected, by virtue of the green tick symbol. A red cross on a data source indicates it is not connected.

A remote data source may be unavailable for several reasons:-

- The network is down
- You have configured the data source connection parameters wrongly
- The user and/or password are incorrect
- The schema does not exist
- The database is not currently running
- Incorrect or no JDBC driver installed

Every remote data source has an associated monitor, which is a background process in Experian Pandora that monitors the state of the remote source.

To view these monitors, right-click on the **Home** item, select **System> Services.**

Alternatively, you can click on **System > Services** from the explorer

The monitors record all of their activity to the system log, provided logging is active (which is the default behaviour). To disable or enable logging, right-click on the Connection and select **Logging > Disable** or **Logging > Enable** respectively. There is also an option to view the system log filtered just for the messages from the monitor. This is achieved by right-clicking on the Connection and selecting **Logging > View**.
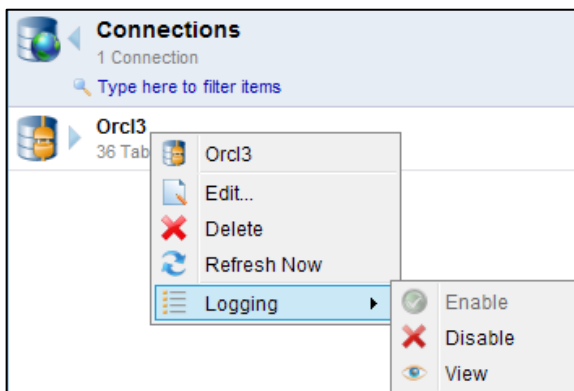
Viewing table and column source

For GDPR purposes, it is important to view the original data source for your tables and columns.

To view this information for your tables and columns, click **Home** > **Data**, and in the Explorer click the Tables or Columns link as appropriate. In the Drilldown view on the right, scroll until you see the following column headings.

Table External Name (Columns Drilldown) / External Name (Tables Drilldown)

For a flat file, this is the original location of the file.
For a JDBC database, this is the JDBC table name.

Table Connection (Columns Drilldown) / Connection (Tables Drilldown)

For a flat file, this is the location of the imported file.
For a JDBC database, this is the name of the JDBC connection the table was loaded from.

If these columns are not visible, use the Drilldown Manager to move the columns from Hidden to Visible.

Using the Explorer

Home

This is the starting point of navigation. On the associated menu you will find access to useful, commonly used actions or drilldowns. These menus can also be accessed by right-clicking on the **Home** icon in the **Explorer** or **navigation bar.**
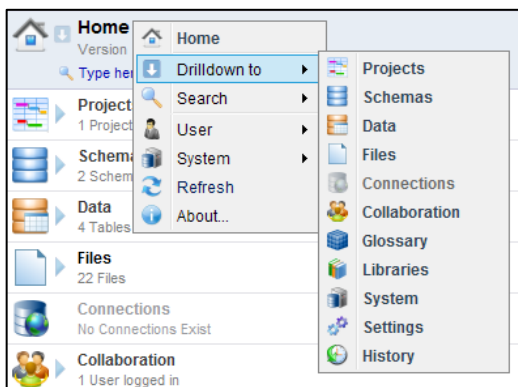
Search

This provides access to enterprise-wide search facilities for data and recorded notes.

Quick Search

Provides quick searching of all data in the server.

Ranked Search

Provides sophisticated searching of all data in the server with results categorised and ranked in order of relevance.

Search Notes

Enables searching of all user defined documentation notes in the server.

User

This menu provides you access to:
- Set your status for other users to see
- Edit your user details, e.g. email address, telephone number
- Change your password and picture
- Change your preferences via the **My Preferences** drilldown

System

Quick access to common Server related activities such as:-

- View running Jobs, Processes and Services and information about configuration, and storage
- Get log file evidence for support or audit purposes
- Enter a new license key
- Password Policy
- Restart or shutdown the server
- Rediscover intra-table relationships from scratch
- Refresh your view of server information
- View the About information box which tells you which client and server version you are running, your license key and various information about your product. A short cut to this information is also found by clicking on the application icon top right of the tool bar.

Projects

Projects are groups of Tables which can be created for any task, e.g. particular analysis or a data migration task. Expanding this item reveals all currently defined Projects. To create a new Project, select **New** from the item menu. The **Tables** tab of the dialog allows adding and removing of Tables to the project. A Table be can in as many projects as is desired. The **Outliers** tab gives you the opportunity to choose which data you would like to see in the Project-level Outliers drilldown. Projects can also contain other Projects (known as Sub-Projects). To create a Sub-Project, access the Sub-Project item beneath the Project in question and select **New** from the menu.

Schemas

Schemas are subdivisions of the tables in the Experian Pandora database. Expanding this item reveals all currently defined Schemas. To create a new Schema, select **New** from the item menu. The **Tables** tab of the dialog allows adding and removing of tables to the Schema. A Table can only belong to a single Schema. The **Outliers** tab gives you the opportunity to choose which data you would like to see in the Schema-level Outliers drilldown. The **Default** Schema cannot be removed and contains all Tables loaded without a defined Schema.

Data

Provides global access to all Tables loaded in Experian Pandora, regardless of what Schema they belong to. Additionally it provides global access to Columns, Saved Drilldowns, Rules, Views and Relationships. In addition to this, it also provides access to the global Value drilldowns as values are only held once across all Tables.

Files

Details all files that are currently available for loading into Experian Pandora. These are **Delimited files** (e.g. comma separated, tab separated etc.), with or without headers, **Excel Spreadsheets** (both Excel 2003 **xls** format and Excel 2007-2010 **xlsx** format), Files with fixed-length, non-variable, non-delimited records.

Any file or table that has an orange tick on its icon in the Explorer has already been loaded to Experian Pandora. You can view the status of the file, as to whether it has been configured before Loading, or Loaded without configuring, whether it has an associated DDL or Copybook etc. Additional files can be obtained via the File Manager and will then be listed in this area.

Connections

Provides direct access to any data source that has a type IV **JDBC driver** – which covers all popular database management systems and file types such as Oracle, SQL Server and MySQL. Expanding a connection will show the available files or tables in that connection. You can usually expand again to see the Columns within that file or Table. The menus provide facilities to configure files or tables for loading, preview their data content, delete them if allowed and of course initiate loading of them. See the section on Loading Data for further details.

Collaboration

Provides the ability to collaborate with other users, via Meetings, instant Messaging and Note creation. See section on Collaboration for further details.

Glossary

The Glossary is where the users of Experian Pandora can configure and define Business Terms and create Business Categories. These allow you to categorise sets of Columns in multiple different Tables with a common set of Business definitions as opposed to relying on technical Column names that mean little in the real world.

Libraries

This item provides access to a variety of reference data sources that can be used during analysis and transformation of data:

- **Domains** are a source of reference data that is dynamically loaded into the memory of the Experian Pandora server and changes on demand. A wide range of functions and capabilities are provided that relate specifically to Domains. These can be used for existence checking, standardisation and translation to name but a few uses.

- **Business Constants** are simply values with a business name. This is useful where complex technical values need to be utilised simply – a great example is handy regular expressions to validate data which visually are gobbledygook but via a Business Constant are simple English. Business Constants can also be used for enterprise wide definitions that may be used in multiple Rules or Drilldowns and need to be changed in one place. A Tax rate is a good example of this.

- **Standard Functions** are the built-in functions in Experian Pandora which are used when creating Expressions.

- **User Defined Functions** are the user-defined functions built from existing standard functions to create custom transformation or validation functionality.

- **ISO Countries** is the reference source used for the **Convert ISO Country** function. The associated Drilldown is complete with current flags for all countries.

- **Reference Table Vendors** is the top-level grouping for **Reference Tables**, which are compact, portable tables of reference data, which can be imported or packaged and licensed for use in other instances of Experian Pandora. A Reference Table can be created from any Drilldown.

System

This provides information about various facets of the Experian Pandora server. Further details are under the System Administration section.

Settings

Enables you to edit your personal preferences as well as those of the server in general, if you have permissions to do so.

It also provides the ability to create, modify or delete:
- Roles for Users
- Teams of Users
- Users themselves
- Categorisations for Notes
- Defined Impacts for Notes
- Defined Statuses for Notes
- Common categories for clustering Rules

History

This provides access to your Drilldown History accumulated automatically, much like the history on an Internet browser.

How to use Drilldowns

Drilldowns are the primary means of visualising data and metadata in a tabular format in Experian Pandora. There are many different drilldowns available including mirror drilldowns for levels in the Explorer, allowing you to navigate Experian Pandora with just a single drilldown window. Many of the items in the Explorer produce a drilldown view, either by single-clicking on the name of the explorer item or by accessing the item menu and selecting on of the emboldened menu items.

Drilldowns usually provide the means to manipulate the viewed data and create new derived drilldowns dynamically, without the need for any technical or programming skills. The following screenshot shows a drilldown of the rows of data in a loaded Table.

Title Bar

The title bar displays key information about the Drilldown:

- An icon, indicating the Drilldown type
- The description of the Drilldown and the name of the object it applies to (if relevant).
- The number of rows in the drilldown (or groupings if it is an aggregated/grouped drilldown)

Column Count

Displays the number of columns which are showing in the drilldown and the number that are available. See the Drilldown Manager section below for more information.

Toolbar

The drilldown toolbar buttons are summarised in the diagram below. Clicking on any of the yellow tabs will take you directly to the explanations in detail.
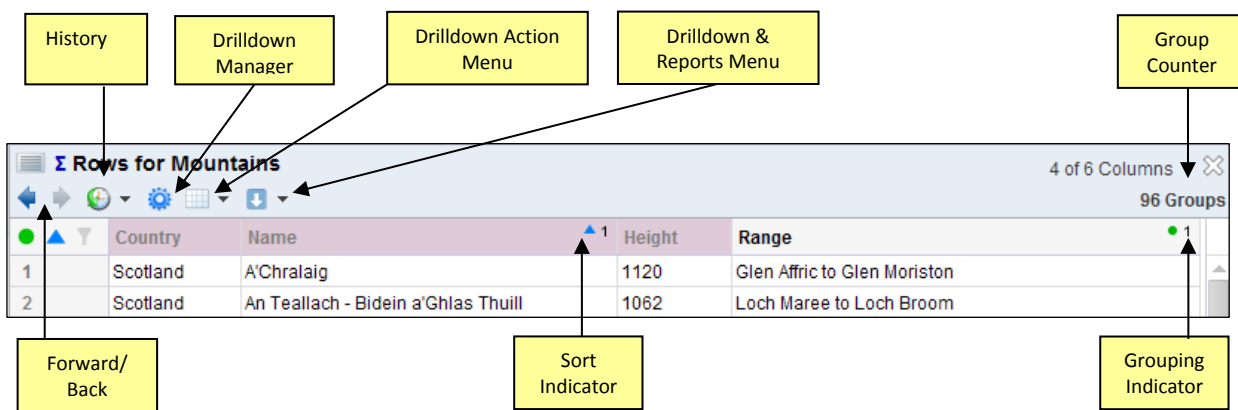
Forward & Back Buttons

Every Drilldown window has a history of activity. For example, if you sort by a column, this results in a new Drilldown within the same window. Simply clicking the **back** button will return you to the unsorted Drilldown. This provides an effective undo/redo mechanism when you are manipulating data or metadata with filters, aggregations and transformations. Note that this window history is **independent** of the overall Drilldown History for a user that is accessible from the **History** item in the Explorer.

History

The **history** dropdown allows you to jump to a specific previous drilldown in the history for this Drilldown only. The last place you accessed will be shown in **bold.**

Drilldown Manager Button

This button displays the **drilldown manager** in a panel that appears below the toolbar. This is discussed in more detail below in the section Drilldown Manager.

Drilldown Action Menu

Save As

Drilldowns can be output in a variety of ways such as:

Drilldown

Saves the Drilldown as a Saved Drilldown, which is a like a shortcut.

View

Saves the Drilldown as a View, which behaves like a Table but is really an overlay in real-time over a Drilldown.

Table Filter

Saves the Drilldown as a Table Filter which can be used to filter out records for Validation and Quality Rules on Tables.

Table

Saves the Drilldown as a new Table, fully analysed

Reference Table

Saves the data as a compact, portable Reference Table.

File

Saves the Drilldown as a file on your client PC or on the server in a variety of formats:

- HTML
- CSV (Comma-Separated-Values)
- Tab-Delimited text
- Pipe-Delimited text
- PDF (Adobe Portable Document Format)
- Excel Spreadsheet (both XLS and XLSX variants)

Target

For exporting your data to a JDBC connection, in your chosen format.

Web Link

A live URL to a row-limited view of the content in the current Drilldown. This is not a point-in-time capture so may change when viewed later in time.

Clipboard Image

Captures a bitmap image of the Drilldown and place it in the clipboard for pasting into other applications.

Send to

You can share a Drilldown by sending it to another user in an immediate interactive chat session. The recipient will then be able to interact with the same information while you chat with them regardless of where they are in the world. When you select the **Send to** item you are presented with a list of currently-available users. If a user is not logged on but does have an email address it will instead use your default Email application (if defined) to construct an email to that user with an attached live Web link URL.

Edit

Provides capabilities to edit aspects of the Drilldown.

Details

Edit the name and description of the Drilldown.

Remap

The Remap Dialog provides an exact and fuzzy automatic mapping for a Drilldown to a previously defined structure. This can be used, for example, to rapidly remap a loaded table to a new previously defined target structure. This is only applicable for drilldowns of data rows such as the rows of a Table or a View.

SQL

Allows editing of the Drilldown by direct editing of the SQL, if you have the user ability to do this. We do not recommend the manual editing of SQL as it is complex and error prone.

Remove

Remove provides the options to remove Sorting, Grouping or Filtering or All of the Above.

Selection

Provides options related to any data selected in the Drilldown:

Copy

Copies the selection to the clipboard so that you may paste it into another application, provided your user has the permission to do so.

Select All

Selects the whole drilldown.

Clear Selection

Clears any current selection of data in the Drilldown.

Fit

Provides options to fit the data in the drilldown to:

- Contents – will fit the data to the width of the contents that are currently displayed in the drilldown.
- Headers – fits the contents to the width of the Headers for each column.
- Packed – shows the first few characters of the header and data for each column. Move the mouse over the column headers to see the name of the column. You can expand individual columns to see more data by double clicking on the right hand divider in the column header. Packed is a useful option when moving the column order within the drilldown. See below.

Clone

Creates a new Drilldown window with an identical Drilldown inside it.

Split

Divides the window into two halves, stacked vertically, to display two Drilldowns within a single window. The top Drilldown is linked to the bottom so that any Drilldowns invoked from the top half will appear in the lower half with both halves retaining their own history. Drilldowns invoked from the bottom will stay in the bottom. **Split** acts like a toggle, turning splitting on and off.

Heatmap

Some drilldowns provide the ability to view data as a heat-map where the cells are coloured according to the relevance of values as a *heat*. This ranges from blue, through to green, yellow and finally red. Blue is the coldest, red is the hottest. An example of this is the rows of a loaded Table. Experian Pandora will colour the values according to how often they occur in the column they are in. Red would be the most highly occurring values.

Datatypes

Some drilldowns provide the ability to view the datatype of the data in the table. The column name will show the icon for the Datatype it has loaded as and against each entry in the table the icon will show the actual datatype. The datatype that does not agree with the datatype loaded will show in **BOLD** with the icon of its actual datatype next to it.

Drilldown and Reports Menu

Provides access to related information or alternate representations. This will vary depending on what you are currently viewing. In most cases you will see:

Info

Displays a new information Drilldown with detail about the columns of the current Drilldown, including all of the columns that are not visible. The information is in positional order, so simply sorting the list will provide you with an alphabetic list of columns.

Summarise

Displays a summary Column Profile drilldown with detail about the columns of the current drilldown with commonly calculated profiling statistics, including counts of passes and fails for columns with validation expressions, and errors and warnings for failed transformations. This is a very quick way to get an overall profile of the quality and results in all the visible columns of a Drilldown without resorting to a full analytical load. It only analyses what is necessary and on repeat viewing produces the result instantly.

Reports

Certain Drilldowns will have instantly available reports, for example the Table Quality report and the Quick Mapping Report.

Row Number

The Row Number is an identifier for the rows that are displayed in the drilldown. The Row Number column remains visible regardless of the horizontal scroll position of the window.

Indicator Panel

This is a quick reference panel to see when grouping, sorting and filters have been applied to the drilldown. When they are applied they will change from ghosted to coloured.

Click on the Indicator Panel to open the menu options:
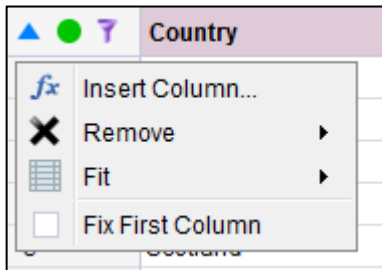
Insert Column

Opens the Expression Editor. See the section on Data Transformation.

Remove

Remove provides the options to remove Sorting, Grouping or Filtering or All of the Above.

Fit

Provides options to fit the data in the drilldown to:
- Contents – will fit the data to the width of the contents that are currently displayed in the drilldown.
- Headers – fits the contents to the width of the Headers for each column.
- Packed – shows the first few characters of the header and data for each column.   Move the mouse over the column headers to see the name of the column. You can expand individual columns to see more data by double clicking on the right hand divider in the column header.   Packed is a useful option when moving the column order within the drilldown. See below.

Fix First Column

Will lock the first column in place so when scrolling through the table it is always visible. Any column can be moved into the first column position and locked. To unlock it, open the menu and un-tick the box.

Action Panel

The action panel contains three icons.   The first is the column selector, the second the Go To Row and the third the Find options.

Column Selector

Click this to popup a searchable list of visible columns (sorted by name). This allows fast location of columns in Drilldowns where the column count is high. Clicking on a column in the list will position the window to that column if it is not already visible.

Go To Row

To move quickly to a row in the drilldown click the Go To icon, or use the keyboard short cut: Ctrl + G, enter the number of the row, click the blue arrow or hit Enter. The drilldown will navigate to position the required row at the top.

To close the panel click the red cross or use the keyboard short cut Esc.

Find

Using the Find tool you can search for a value in the drilldown. The results will be highlighted by row and individually in the drilldown Using the blue back and forward arrows you can navigate between the results or the keyboard shortcuts.

The criteria choices are:

- Case Sensitive   the search will be case sensitive search if selected
- Regular Expression   type in a regular expression and the results will have to match the search pattern
- Wrap Searches   if selected the result will scroll through the data repeatedly, if not selected the search will start from the top of the table and stop at the end
- Match all Terms   all words / terms must be in the row for a result to be valid
- Limit to Selected Column select a cell in the column you wish to search and the results will be limited to that column

Keyboard short cuts:

Ctrl + F	opens the Find panel
Return Key	next result
F3	next result; opens the panel if closed
Shift + F3	previous result; opens panel if closed
Esc	once will close the panel (if open), twice will exit the search

Scroll Bar & Column Navigator

The scroll bar is combined with an alternative to the column selector button. Click on the blue scroll bar to move along the columns. The name of the column will change as you move over them. For column navigation move your mouse along the bar without clicking on it and the names of the columns will appear in the order that they appear in the drilldown. Click on the desired column to position the window accordingly.

Column Headings

As well as showing the name of the column, column headings show the following additional information:

Sorting Indicator

Experian Pandora will show a blue up arrow for ascending sort, or a green down arrow for descending sort and the sort index of the column, i.e. the sort position of the column relative to other sorted columns. The image below shows the **Name** column sorted ascending with sort index 2 and the **Height** column sorted descending with sort index 1, or in other words the data is first sorted by **Height** in reverse order, and then by **Name** in ascending order. The sort icon in the Indicator Panel will change to a solid colour when sort is applied.

To create a sort, left click in the column header for the column you want to sort. The first sort will be ascending and if you click again, it will sort descending. Click a third time to remove the sort. To add a second or subsequent sort in another column, hold the shift key down and left click on the relevant column. Alternatively, you can right-click on the desired column and select one of the available Sort options.

Grouping Indicator

Experian Pandora will show if a column is grouped, by displaying a green dot and the group index of the column relative to other grouped columns. The group icon in the Indicator Panel will change to a solid colour. The screenshot extract below shows how the Country column header appears when grouped. Note the grey texted, pastel red background column headings, which signify that the remaining columns do not have an aggregation function selected or cannot be exactly determined. In this case, the first value encountered in that grouping is displayed – equivalent to the **First** aggregation function.

This grouping is done by holding down the Ctrl key and left clicking on the Country column or by right-clicking on the Country column header and selecting **Group > Only By This**. Note that when a drilldown is grouped instead of displaying the number of rows, it changes the display to the number of **groups**.

Filter indication

An icon will display on the name of a column if it is involved in a filter and the Filter icon in the Indicator Panel will change to a solid colour. A filter is applied via the **Drilldown Manager** which is accessed by selecting the **drilldown manager** button on the toolbar.

Aggregation and transformation indication

The following indicators may also appear on a Column heading in-front of the Column name:

Indicator	Meaning
ƒ	The column has been created as a result of a transformation function
§	The column has been formatted and may be appear different to the actual data content
♯	A Count aggregation. Bold and dark indicate it is a Distinct Count (unique values only)
∑	A Summation. Bold and dark indicates it is a Distinct Sum (sum of only the unique values)
A	First value in a grouping
Ω	Last value in a grouping
Λ	Smallest value in a grouping
V	Largest value in a grouping
σ	Result of a standard deviation aggregation. Bold and dark indicates Distinct Standard Deviation (unique values only)
σ2	Result of a variance aggregation. Bold and dark indicates Distinct Variance (unique values only)
{}	Value grouping aggregation. This is the set of values as a list from the column the aggregation was applied to. Bold and dark indicates Distinct Group Values (unique values only, sorted)

Moving and Hiding Columns

The order of the columns can be changed by dragging and dropping them to the desired location in the drilldown. There are two places where this can be done. The first is on the Column Headers. Select the column you wish to move by clicking on the Column Header and drag it to the required location.

The second is on the Scroll Bar and Column Navigator. Again click on the Column Header and move to the required location.

To hide a column, select the Column Header name of the column to hide and drag it up to the tool bar area of the drilldown. A message will appear saying "+Hide". Release the column and it is now hidden. To recover hidden columns use the Drilldown Manager. See the section below Using the Drilldown Manager.

Drilldown column heading menu

Right-clicking on a column header will display the column menu:

Edit

Edit produces a dialog which allows you to edit everything about the column details and metadata and its appearance.

Delete

Only applicable to columns created as a result of the **Insert Column** menu item, from an expression. This will remove that created column entirely from the Drilldown.

Hide

Hides a column in the Drilldown. Hidden columns are still present and available but are simply not shown by default. You can display the column again by using the Drilldown Manager.

Fit

Provides options to fit the data in the drilldown to:
- Contents – will fit the data to the width of the contents that are currently displayed in the drilldown.
- Headers – fits the contents to the width of the Headers for each column.
- Packed – shows the first few characters of the header and data for each column. Move the mouse over the column headers to see the name of the column. You can expand individual columns to see more data by double clicking on the right hand divider in the column header. Packed is a useful option when moving the column order within the drilldown.

Sort

Only By This (Ascending)

Sorts the Drilldown by this column only, in ascending order

Only By This (Descending)

Sorts the Drilldown by this column only, in descending order

Also By This (Ascending)

Adds this column to the sort order for the Drilldown, in ascending order

Also By This (Descending)

Adds this column to the sort order for the Drilldown, in descending order

Unsort

Removes the column from sorting entirely

Group

Only By This

Groups the Drilldown by this column only

Also By This

Adds this column to the grouping for the Drilldown

Ungroup

Removes this column from the grouping for the Drilldown

Insert Column

Adds a new column after the one being clicked on.

Custom

This displays the Expression Editor to allow you to create an expression using any combination of functions available.

Copy of

This option creates another column that is an exact replica of the selected one. This is useful when you need to create multiple calculated columns that have subtle differences, allowing modification of a particular expression as a starting point.

Aggregate

This option provides quick access to functions that aggregate the values in a column in a variety of ways, or change an already applied aggregation. These are normally applied to a Drilldown that has a grouping applied with the results displayed with one result row per grouping. If no groupings have been applied, the aggregation is performed on the entire drilldown resulting in a **single result row**.

The remaining menu options are functions by categories that accept a single input which will be the column you clicked on. Some functions do not have any inputs such as **Current Time** and **Current Date.** This list of functions will vary from release to release.

Transformation

This enables you to transform a column in a Drilldown, or to edit an already transformed column (one that was either subject to a previous transformation or was created via the Insert Column menu option).

Create/Edit

If the column has not been previously transformed then the option will be **Create** to apply a new transformation to the column. This option need only be used if you wish to apply a complex transformation using the expression editor, or to use a function that requires more than just a single click to achieve its result (for example, Substring). If the column already has a transformation applied, then the option will be **Edit** to allow you to change it in the Expression Editor

Save As

This option allows you to save a transformation for re-use in a number of ways:

Global Transformation Function

This saves the expression as a User-Defined Transformation Function which will appear as a new function block in the Expression Editor.

Global Validation Function

This saves the expression as a User-Defined Validation Function which will appear as a new function block in the Expression Editor and can only return true or false.

Table Expression

This only applies if the Drilldown is rows of a Table. This will save the expression as a Table Expression which is effectively the same as adding a user-defined Column to a Table. Table Expressions are particularly useful for the validation of rules on a Table or Column.

Remove

This removes the entire transformation from the original column.

The remaining menu options are functions by categories that accept a single input which will be the column you clicked on. This list of functions will vary from release to release and is different to the list in **Insert Column** as all of the functions accept an input in order to transform the column clicked on. These can be applied incrementally to progressively transform the data, resulting in a large end-to-end expression. This can be seen and edited by using the **Edit** option

Profile

This option provides the ability to dynamically profile the data in a column of a Drilldown without the need to go through the usual analytical load process. This provides frequency analysis of values, format patterns and phonetic patterns as well as analysis of missing values (nulls). These are provided both as entire frequency results, or top N and bottom N. This defaults to the top and bottom 10 values (i.e. the most frequent 10 and least frequent 10) but this number is a user preference that can be changed to a different value, for example the top and bottom 50.

More on Drilldowns

There is a vast array of Drilldowns available in Experian Pandora, and the easiest way to discover what is available from items in the Explorer or from selections on a Drilldown is to examine the right-click menus. Experian Pandora has been designed to be easy to navigate and to have the most appropriate Drilldowns available at your finger-tips. If you cannot find what you are looking for, contact your support representative who will be more than happy to assist.

Every release of Experian Pandora is likely to see the various available Drilldowns expand with additional columns, variants and even completely new Drilldowns.

Drilldowns and the Explorer

In general, if the Drilldown is the equivalent of a level in the **Explorer**, then you will find that the objects in the Drilldown have exactly the same menu options as the Explorer items, along with potentially some additional ones.

A common and very useful facility is to be able to find the item in question in the Explorer. We can look at the Drilldown of all Tables loaded in Experian Pandora, right-click on one of them and select **Locate in Explorer**. Experian Pandora will immediately navigate to where that item is in the Explorer.

Using the Drilldown Manager

The Drilldown Manager allows you edit various aspects of the Drilldown in combination and then apply them in a single execution, as follows:

- Make columns visible or invisible
- Delete one or more expression columns
- Arrange the order in which visible columns appear
- Change the way in which the data is sorted
- Specify how the rows of the drilldown are to be filtered
- Temporarily enable or disable individual filter conditions
- Group data rows together by one or more columns
- Change the method of grouping of data rows
- Limit the number of rows returned in the drilldown result
- Sample the returned results to return less information
- Apply a post-grouping filter to eliminate records after they have been grouped

The majority of the above operations can be completed one-at-a-time from the pop-up menu that appears when a column header on the drilldown is right-clicked.

The **Drilldown Manager** is implemented as a pop-up panel in all drilldown views that allow modification, which is the majority of them, and can be shown or hidden as necessary. Select the blue cog-wheel icon to show the panel.

The icon will change to orange. To close the Drilldown Manager click the orange cog-wheel.

![Experian logo]

Drilldown Manager Interface

Clicking on the drilldown manager button opens up the Drilldown Manager panel. The **mode** of the Drilldown Manager can be set to **Columns, Sorting, Grouping** and **Filtering** at the top-left of the Drilldown Manager by using the mode dropdown box or by clicking on one of the Mode Selector icons.

For all modes changes made in the Drilldown Manager are not applied until the **Run Query** button is pressed, at the top right of the panel. If this button is disabled, it means that you have not made any changes that can be applied. If this button's symbol appears red, it means that there is an incomplete Filter or Group Filter.

We will start by looking at the Columns mode:

Mode
This menu dropdown button changes the drilldown manager mode between Columns, Sorting, Grouping and Filtering.

Hidden Column List
Displays the full list of hidden columns that are available for the drilldown view. Columns may be dragged from here to the Visible Columns panel or use the arrow at the bottom of the panel.

Mode Selector
Click the relevant icon to move to the required mode for Columns, Sorting, Grouping or Filtering.

Column Search Field
Typing in the Search field will filter the columns in the Hidden or Visible list respectively to show only columns whose name contains the search text.

Visible Column List
Displays the full list of columns that are / will be visible for the drilldown view. Columns may be dragged from here to the Hidden Columns panel or use the arrow at the bottom of the panel.

Rows for Sales Order Detail (Table) — 6 of 8 Columns — 526,374 Rows

Columns | Reset | Run Query

Hidden Columns | Visible Columns

Search | Search

Row Id | Order Id
Source Row Id | Product Id
| Quantity
| Shipping Date
| Price
| Discount Code

Order Id	Product Id	Quantity	Shipping Date	Price	Discount Code
1	1150873	RTNR985	250	18-Jan-2010	0	M
2	1076603	NT646	295	10-Jul-2009	0.5	L
3	1076609	NT646	370	08-Jul-2009	0.5	H

Move to Visible Columns
Move the column from the Hidden columns to the Visible Columns.

Move to Hidden Columns
Remove the column from the Visible Columns to the Hidden Columns.

Up / Down
Repositions a column up or down the list. Columns can also be dragged.

Reset
Resets back to the initial state before any changes were made.

Run Query
Apply all changes and execute the query for a new drilldown result.

Columns

The Columns function allows you to specify column visibility and column order in the drilldown. All the Hidden columns which are available for the drilldown are on the left hand side in the Hidden Columns panel and listed in alphabetical order. All the columns which will be Visible in the drilldown are in the right hand side in the Visible Columns panel and can be positioned to the order you require them.

Moving a Column from Hidden to Visible

A hidden column can be moved to the visible columns in any of the following ways:

- Double click a column name in the Hidden Column panel
- Drag a column from the Hidden Columns panel to the Visible Columns panel. The column can be dropped into the desired position.
- Using the Ctrl key select more than one column name and either drag them or click the blue right arrow at the bottom right of the panel.

Hiding a Column

A visible column can be moved to the hidden columns using any of the following methods:

- Double click a column name in the Visible Columns panel
- Drag a column name from the Visible Columns panel to the Hidden Columns panel.
- Using the Ctrl key select one or more columns on the Visible Columns panel and either drag them or click the blue arrow left arrow at the bottom of the panel.

Change the Order in which the Columns appear

Changing the order of the columns in the Visible Column panel can be done a couple of ways:

- Select a column name and drag it to the required position
- Select a column name and use the Up and Down blue arrows to move it to the required position.
- To move a group of columns, use the Ctrl key to select them and either drag them to the required position or use the Up and Down blue arrows.

Column Search Field

Typing text in this field will instantly filter the Column List to show only those columns whose name contains the text that you entered. This is useful when you need to find one column in several hundred.

Once you have decided which Columns you want to be visible in the drilldown and the order you want them in, click the Run Query button at the top right of the Drilldown Manager and your changes will show in the drilldown window below.

Sort Panel

Experian Pandora allows the user to interactively sort data on multiple columns. Sorting of data is achieved using a parallel sorting engine that uses multiple processors on the Experian Pandora server. Reversing a sort, once completed, is instantaneous and repeating a sort at some time in the future will also be instantaneous thanks to Experian Pandora automatically indexing operations such as these if they take more than a second or so. These indices are created and managed automatically, and are applicable to all operations that change the number of rows, or the order of them, for example filtering and grouping data.

The unsorted columns are on the left hand side in the Unsorted Columns panel and the sorted columns are on the right hand side in the Sorted Columns panel.

To sort by a column:
- Double click a column name in the Unsorted Columns panel to move it to the Sorted Columns panel
- Drag a column name from the Unsorted Columns panel to the Sorted Columns panel.
- Using the Ctrl key select one or more columns on the Unsorted Columns panel and either drag them or click the blue arrow left arrow at the bottom of the panel.

The Sorted Columns panel shows the names of the sorted columns of the drilldown view. The order in which the columns will be sorted is dependent on the list order in the Sorted Columns panel itself. To change the sort from Ascending, blue up arrow, to Descending, green down arrow, click the arrow. When the query has been run the columns in the drilldown will show which order they have been sorted in with a number in the column header and the direction arrow to indicate whether ascending or descending:

Change the Sort Order of a Sorted Column

The sort order of a column can be changed from Ascending to Descending by clicking on the arrow in front of the column name. Blue is Ascending and Green is Descending.

Remove a Column from the Sort Criteria

A column can be removed from the sort criteria (i.e. unsorted) using the following methods:

- Double-click on the column in the Sorted Columns Panel
- Drag the column to the Unsorted Panel
- Select the column and use the left arrow at the base of the panel.

Change the Sort Criteria Ordering

The order in which the column sorting is applied (i.e. the columns' sort indices) can be changed using the following methods:

- Select one or more sorted columns on the Sorted Columns Panel then drag them to the desired location.
- Select one of more columns on the Sorted Columns Panel then use the **Up/Down** buttons to change their position.

Grouping and Aggregating Data

Experian Pandora allows the user to interactively group data by multiple columns and apply aggregations which appear as new columns. Groupings do not suffer from the normal restrictions place by SQL databases, which can only display in its output those columns that are either results of an aggregation function or one of the columns that are being grouped.

Experian Pandora displays every column, automatically showing you the **first** occurrence of the value that column in the grouping. You are free to hide affected columns or select specific aggregation functions, if the default **first** behaviour is not what is desired. Experian Pandora provides full grouping and aggregation capabilities, which are easily accessible via the drilldown column header pop-up menu or the Drilldown Manager grouping panel. Virtually every drilldown can be grouped with single or multi-column groupings allowed.

Grouping and aggregation, like sorting, creates an incremental index so a repeat of the operation is instant. The result of a grouping can be post-filtered, for example, to show rows where the count of a particular column is greater than 3. This will use a previously grouped index automatically, so that regrouping is not required to simply filter the grouped result, saving a lot of precious analysis time.

Group Panel

The Grouped Columns Panel shows the grouped columns of the drilldown and the order in which they are grouped. This panel allows you to add or remove columns from the grouping criteria and change the order in which the drilldown is grouped.

Add a Column to the Group Columns

A column can be grouped in any of the following ways using the following methods:

- Double-click on the column in the Ungrouped Columns Panel
- Drag a column from the Ungrouped Column panel to the required position in the Grouped Columns Panel
- Select the column and use the right arrow at the base of the panel.

Change the Group Criteria Ordering

The order in which columns are grouped (i.e. the columns' group indices) can be changed using the following methods:

- Select one or more group columns on the Group Panel then drag them to the desired location.
- Select one of more columns on the Group Panel then use the **Up**/**Down** buttons to change their position.

Remove a Column from the Group Criteria

A column can be removed from the group criteria in any of the following ways using the following methods:

- Double-click on the column in the Grouped Columns Panel
- Drag the column(s) from the Grouped Columns panel to the Ungrouped Columns panel
- Select the column(s) and use the left arrow at the base of the panel.

Aggregation Functions

The following aggregation functions are provided by Experian Pandora:

Function	Description
Average	The average of values in the column
Count	The count of rows
First Value	The first value in a column within the grouping
Group Values	The full list of values in the column, as a comma-separate list up to a maximum of 32,767 characters
Last Value	The last value in a column within the grouping
Maximum Value	The largest value in a column within the grouping
Minimum Value	The smallest value in a column within the grouping
Reverse Group Values	The full list of unique values in the column, sorted in reverse, as a comma-separated list up to a maximum of 32,767 characters
Sorted Group Values	The full list of values in the column, sorted, as a comma-separated list up to a maximum of 32,767 characters
Standard Deviation	The standard deviation of values in the column
Sum	The sum of the values in the column
Variance	The variance of values in the column
Distinct Average	The average of unique values in the column
Distinct Count	The count of rows where the column has a unique value
Distinct Group Values	The full list of unique values in the column, sorted, as a comma-separated list up to a maximum of 32,767 characters
Distinct Reverse Group Values	The full list of unique values in the column, sorted in reverse, as a comma-separated list up to a maximum of 32,767 characters
Distinct Standard Deviation	The standard deviation of unique values in the column
Distinct Sum	The sum of the unique values in the column
Distinct Variance	The variance of unique values in the column

You can apply aggregation functions to a grouped drilldown to find out new information related to the data. In this example, we are using a table about the Mountains in the United Kingdom. We initially group by **Country** which results in 3 result groups, one for **England**, **Scotland** and **Wales**. The first requirement is to find out the number of mountains in each of these countries. This is simply achieved by applying a **Count** aggregation on the **Name** column

Having obtained the count, we now want to know the highest mountain in the respective countries, which is achieved by applying the **Maximum Value** aggregation function on the **Height** column. The final step is to hide all of the remaining columns, so that the result just shows the grouped countries, the number of mountains and the highest mountain.

Once data is grouped like this, an additional drilldown appears on the right click menu which allows the user of Experian Pandora to drilldown to the content of one or more of those groups. Experian Pandora automatically builds the correct drilldown as a result of this drilldown request. For example, to now see all the mountains in Wales, simply right-click on the row for Wales and select the drilldown option for **View Grouping 'Wales'**. The result is a list of all the mountains where the country is Wales.

Filter Panel

The Filter Panel in the Drilldown Manager allows you to define how the rows in a drilldown are filtered, either to specifically include or remove information. For example, you may want to only show rows representing orders above a specific value. This filter is applied **prior** to sorting or grouping of data. A filter is built using a point-and-click interface in a guided English form, resulting in a hierarchy of multiple sentences. No technical skill is required to filter data with Experian Pandora.

When dealing with large columns of data the user may choose to set up the filters on a smaller sample of data to ensure they are getting the expected results and then save the created filter. This can then be applied to the whole table. To do this, before opening the rows drilldown right-click on **Rows** and select **Rows (Sample of 100)**
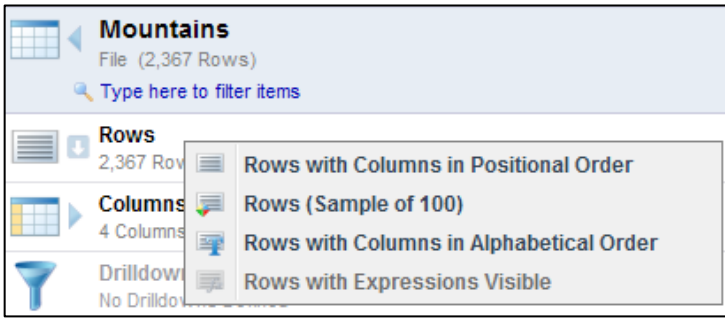
Group Filter Panel

The **Group Filter** allows you to define and apply a filter which uses the results of a grouping operation. For example once customer orders have been grouped by date, and a total order amount calculated for each date via a Sum aggregation, we can limit the results to only show dates where the total order value is greater than £100.

The difference essentially boils down to using the **Filter** to eliminate data from being grouped at all, and the **Group Filter** to eliminate results after grouping has completed. Obviously you can only filter on aggregated results such as counts with a Group Filter.

Drilldown Manager Filtering Panel

The following diagram illustrates the key features of the Filtering panel:

Filter Expression Panel
This is where you create the filter expression. Edit columns, operators and values by clicking on them here and editing them in the Value Selection Panel.

Filter Panel Toolbar
These are the common actions for manipulating the filter expression. More details below.

Value Selection Panel
This is where the various selections are made to complete the filter conditions.

Editing Filters

A filter consists of one or more **Conditions** which is a single statement that evaluates to either true or false. Each line on the Filter Panel represents a single condition. Multiple Conditions may be combined together using the **AND** or **OR** operators. If combined using **AND**, the combination is only true if **all** of the combined conditions are true. If combined using **OR**, the combination is true if **any** of the combined conditions are true.

Conditions may also be hierarchically nested or grouped so that **AND** or **OR** statements can be mixed in a single filter expression. Nesting is identical to enclosing individual AND & OR condition statements in brackets in a normal query language.

A Condition usually takes the following structure:

<LHS> <operator> <RHS>

LHS	The left-hand-side of the condition and must be a column in the Drilldown
Operator	The comparison operator, e.g. **equals**, **not equal to**, **is greater than** etc.
RHS	Can be another column, a constant value or a reference to a Business Constant.

Some operators, such as **is alphanumeric,** do not require a right-hand-side value. Other operators, such as **in edit distance,** require additional parameters.

Each of the components of a Condition can be clicked and edited in the context-sensitive panel on the right of the Filtering Panel.
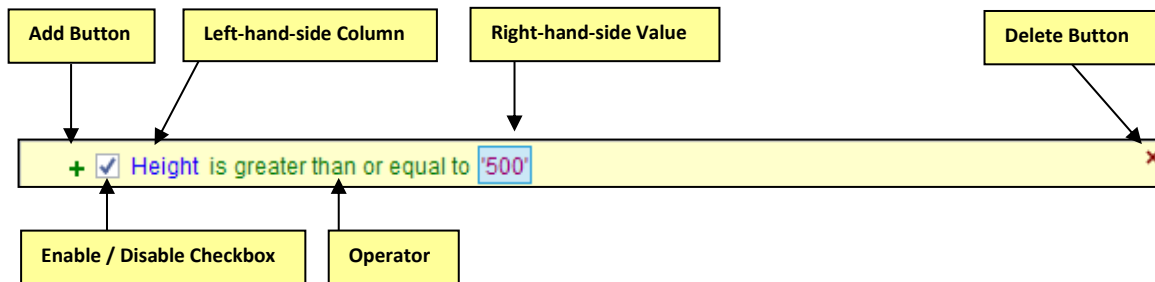
Filter Panel Toolbar

The following diagram illustrates the common actions available on the Filter Panel toolbar.

Add Button
Add a new condition. If an existing condition is selected, the new condition will be added at the same nesting level as the selected condition.

Delete Button
Deletes the selected condition(s)

Negate Button
This negates the whole condition.

Undo / Redo Buttons
Undo or redo the last add/delete/re-order action. NOTE: This undo/redo feature does not apply to changes of individual Condition value or operator changes.

Up / Down Buttons
These buttons move the selected condition(s) up or down the current filter expression.

Nesting Buttons
Moves the selected condition down a nesting level or up a nesting level (see subsequent examples for a better understanding of this function)

Filter Selection Panel

Clicking on a column, operator or value in the filter expression panel will show the appropriate edit panel in the selection panel on the right.

Depending on the currently selected operator for the condition, the right-hand-side value or parameters may be restricted to certain value types, e.g. constant-only, number-only, single value, single or multiple values etc.

An item in the Condition that may take multiple types will result in an edit panel with buttons along the top row.

Where a Filter has been applied to a drilldown the details of the filter can very often be longer than the space available to show in the information line. Clicking on the information line in the drilldown header will open an information panel detailing what has been filtered.

The following examples illustrate the different types of edit panel.

Single Column, Constant or Multiple Values

In this example the value **England** has been clicked, and the edit panel shows two type buttons on the top row to indicate that the value may be either a column or a constant value. The selection panel allows a single value or a list of values to be entered. If the **equals** operator is specified with a list of right-hand-side constant values, the condition will be true if the left-hand-side matches any one of the right-hand-side values. There are many operators like this that can take one or more right-hand-side values.

Here we extend the filter to **Multiple Values** of England and Wales:

Single numeric value or constant

In this example, the **in edit distance** operator takes additional parameters, one of which is the **Edit Distance** parameter, which can only be a number. Clicking this shows a simple edit panel with a numeric entry field.

Creating Filter Expressions

The following section will illustrate how to build filter expressions using examples based upon a simple table. The example table Mountains contains information about all mountains in the United Kingdom, and consists of 4 columns:

Country	The constituent UK country in which the mountain is located
Name	The name of the mountain
Height	The height of the mountain (in metres)
Range	The name of the mountain range in which the mountain is located

Creating a Simple Filter Expression

Objective: Show rows where the mountain height is less than 1000 metres

Starting with the drilldown of the Mountains table rows:-
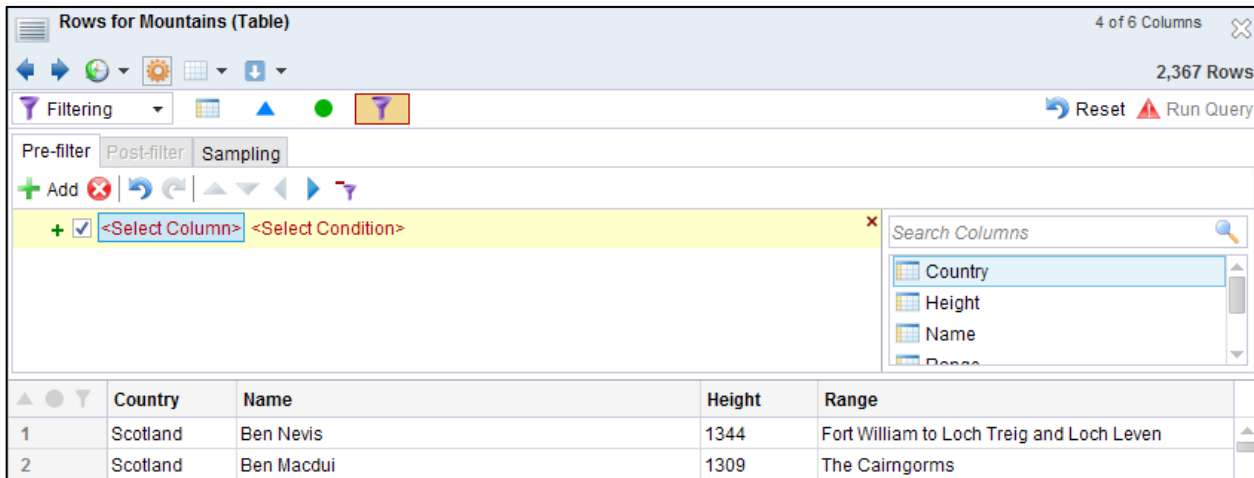
Step 1   **Click on the Drilldown Manager icon on the drilldown toolbar** and set the mode to **Filtering** on the mode dropdown or the Mode button in the Drilldown Manager panel. The drilldown window will appear as follows:

Step 2   Create the Condition
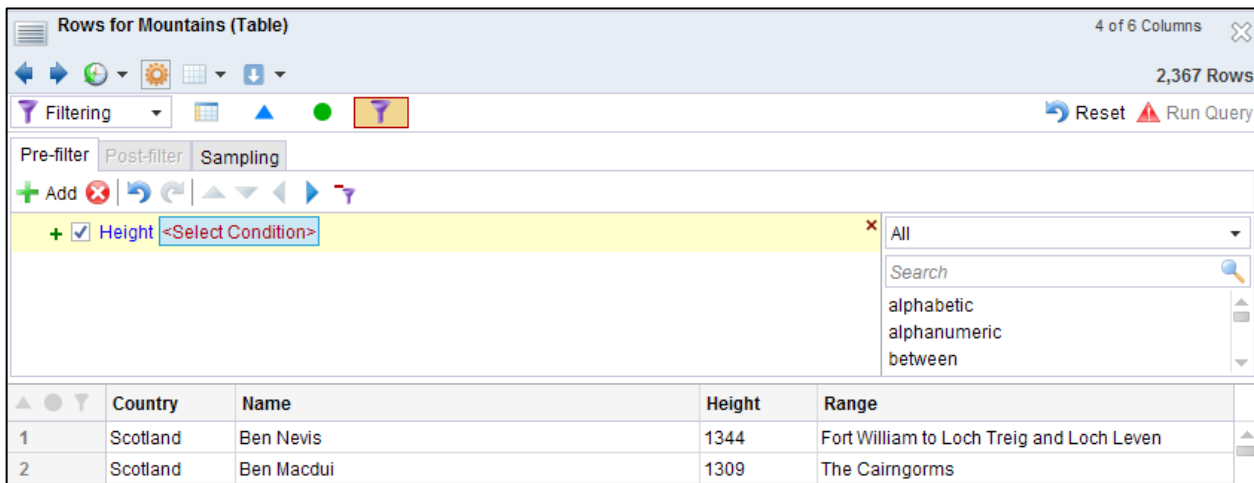
With a newly-created filter expression, clicking the **Add** button or clicking on the label that reads "<*Click here to create a new filter condition*>" will create a new condition in the filter expression panel. The drilldown window will appear as follows:

The left-hand-side column value in the expression is already selected and the Filter Selection panel reflects the fact that this value may only refer to a column not a constant value. Double-clicking on the Height column in the Filter Selection panel will populate the left-hand-side of the expression with that column. The filter expression will now appear as follows:

The filter expression now shows the operator selected and the Filter Selection panel shows a list of available operators.
We now need to find the **less than** operator. This can be achieved in one of two ways:

- Start typing in the operator search field at the top of the operator list. For example, typing **less** in this field will quickly filter the list to show only operators whose name contains **less**.

- Select an operator category to reduce the number of operators to select from.

Once you become familiar with the operators that are available, it is usually most efficient to use the search field to locate the required one.

Double-clicking the **less than** operator on the list will populate the operator value in the filter expression. The filter expression will now appear as follows:

The filter expression now shows the right-hand-side value selected and the filter selection panel has defaulted to showing the single value edit panel, which allows only one value to be specified.

At this point you should note the **Exact** checkbox to the left of the text field. If this is enabled, the right-hand-side value will be interpreted exactly as entered in the text field. If the checkbox is not enabled (the default), then the value will be standardized into the most appropriate data type and may have been modified by the time the Experian Pandora server evaluates the condition. This is normally the preferred method of operation.

For example, enabling the **Exact** checkbox and entering **0.0** as a constant value will result in the value being standardized to **0**. The same applies to date and money values. It normally has no value with alphanumeric values. In most cases, leaving the **Exact** checkbox disabled will not lead to unexpected results. However, if you wanted to compare against the alphanumeric value **10-12-2009**, the checkbox should be enabled to indicate that you did not want to match dates that resemble this value.

To continue the example, we now enter **1000** in the constant value text field and either press the RETURN key or press the **Apply** button to confirm the value. The filter expression will now appear as follows:

Note that the **Run Query** button is now enabled. Clicking this will invoke a new query with the filter expression applied:

Rows for Mountains (Table)
Filter: Height is less than '1000'

4 of 6 Columns

2,107 Rows (89.02%)

Filtering ▾   ▲ ● ▼

Reset Run Query

Pre-filter | Post-filter | Sampling

Add ⊗ | ↺ ↻ | ▲ ▼ ◀ ▶ ▾

+ ☑ Height is less than '1000'   ✕

	Country	Name	ΦHeight	Range
1	Scotland	Stob Ban	999	Fort William to Loch Treig and Loch Leven
2	Scotland	Sgurr Choinnich	999	Kyle of Lochalsh to Garve
3	Scotland	Sgurr nan Ceathreamhnan - Creag nan Clachan Geala	999	Loch Duich to Cannich
4	Wales	Glyder Fawr	999	Snowdonia
5	Scotland	Sgurr Breac	999	The Fannaichs
6	Scotland	Broad Cairn	998	Braemar to Montrose
7	Scotland	Sgurr nan Conbhairean - Creag a'Chaorainn	998	Glen Affric to Glen Moriston
8	Scotland	Stob Diamh	998	Glen Etive to Glen Lochy
9	Scotland	Ben Challum - South Top	998	Glen Lyon to Glen Dochart and Loch Tay
10	Scotland	Ben More Assynt	998	Scourie to Lairg
11	Scotland	Sgor Gaoith - Meall Dubhag	998	The Cairngorms
12	Scotland	Glas Bheinn Mhor	997	Glen Etive to Glen Lochy
13	Scotland	A'Chailleach	997	The Fannaichs
14	Scotland	Creise - Stob a'Ghlais Choire	996	Glen Etive to Glen Lochy
15	Scotland	Spidean Mialach	996	Glen Shiel to Loch Hourn and Loch Quoich
16	Scotland	Maoile Lunndaidh - Carn nam Fiaclan	996	Kyle of Lochalsh to Garve
17	Scotland	Mam Sodhail - Mullach Cadha Rainich	996	Loch Duich to Cannich
18	Scotland	An Caisteal	995	Loch Lomond to Strathyre
19	Scotland	Carn an Fhidhleir (Carn Ealar)	994	Glen Tromie to Glen Tilt
20	Scotland	Sgor na h-Ulaidh	994	Loch Linnhe to Loch Etive
21	Wales	Glyder Fach	994	Snowdonia
22	Scotland	Sgurr na Ruaidhe	993	Kyle of Lochalsh to Garve
23	Scotland	Creag Meagaidh - An Cearcallach	993	Loch Lochy to Loch Laggan
24	Scotland	Stob Poite Coire Ardair-Sron Coire a'Chriochairein	993	Loch Lochy to Loch Laggan

Mountains (Table)

Creating a Complex Filter Expression

Objective: Show rows for mountains in England or Scotland whose height is greater than 1000m or less than 500m

For this example, we need to create a filter expression that is the equivalent of:

(Country = 'England' OR Country = 'Scotland') AND (Height > '1000' OR Height < '500)

Starting with the Mountains table rows drilldown, click on the Drilldown Manager button and then the Filtering mode button to show an empty filter expression.

Step 1 – Create the Expression Structure

Click on the **Add** button to create two new conditions. Either the add indicated below, or the +
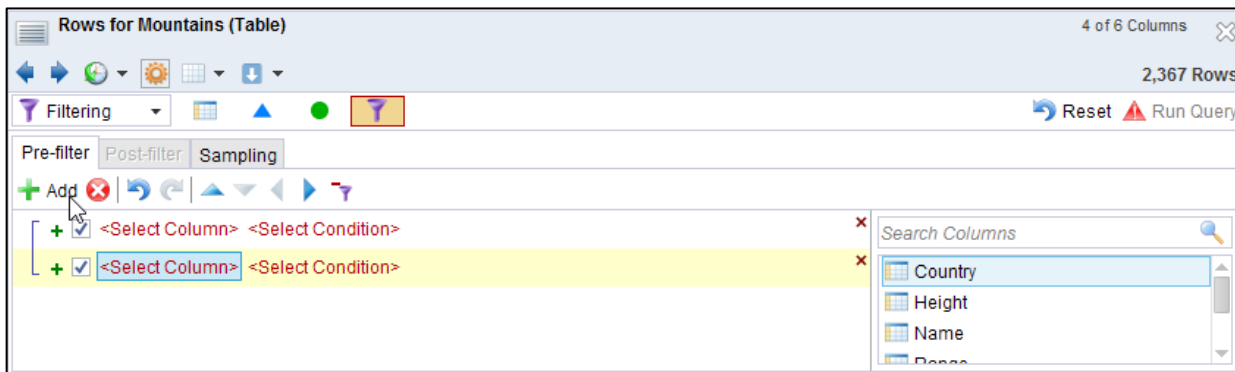
Before populating any values in the expression, we will first create the basic structure of the filter expression: Now we select the first condition by clicking anywhere on it. The condition line will highlight pale yellow. We then increase the nesting level of the first condition by clicking the indent button, as shown in the screenshot below:
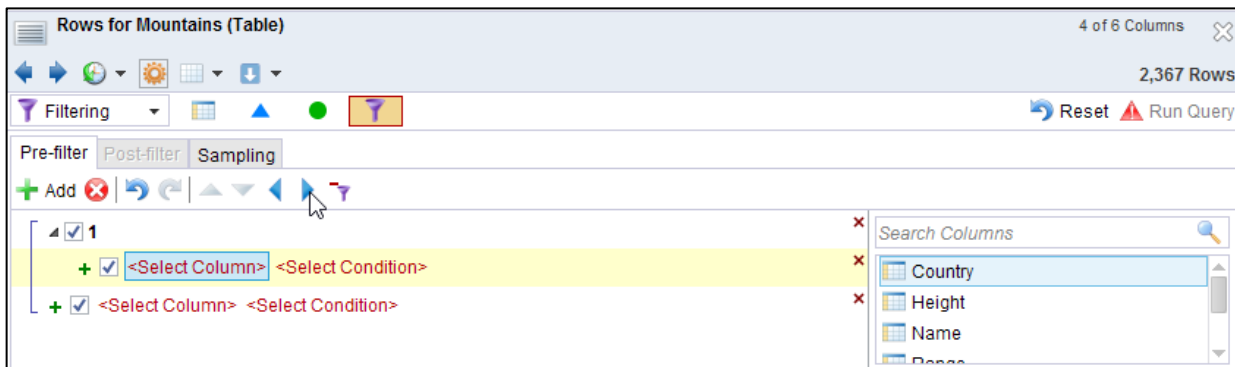
Note that the first condition is moved across and into the nesting group with the title "1". Now click on the Add button to create another empty condition at this level:
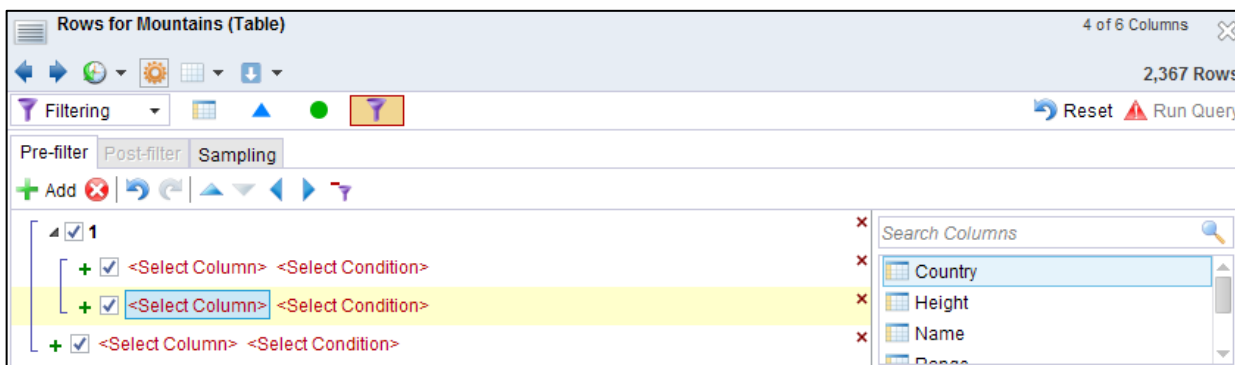
We repeat the same process on the second initial condition to obtain the following structure:
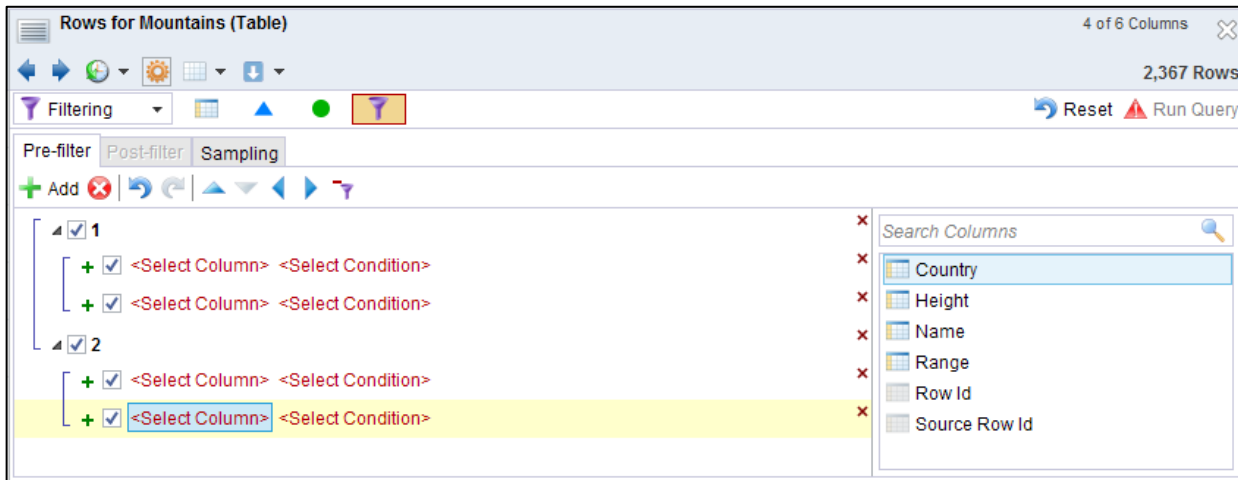
Step 2 – Fill in the Expression Values

Now that the structure of the filter expression has been defined, we now fill in the condition details as follows:
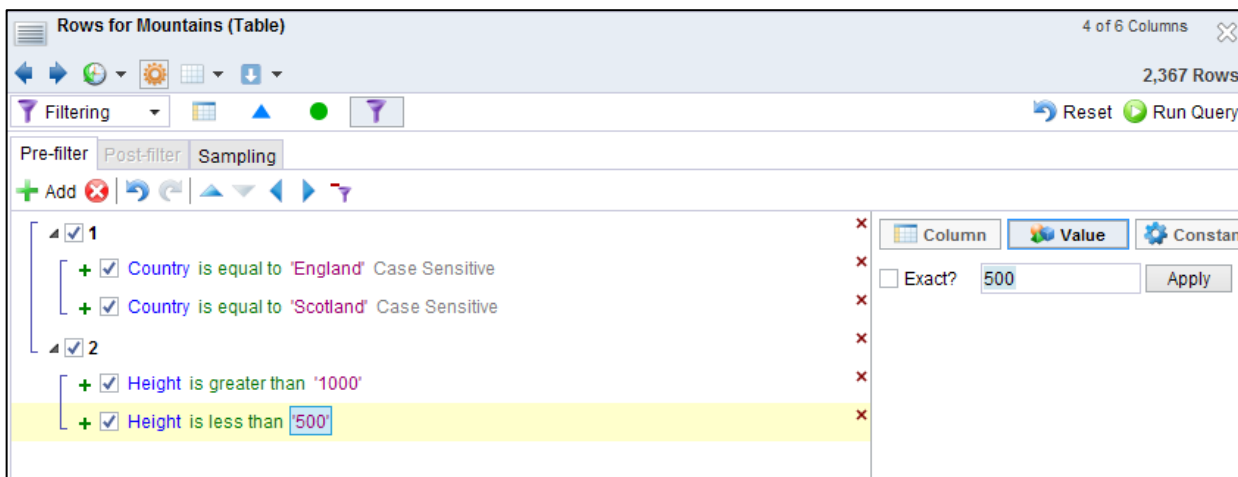
Step 3 – Define the Way the Logical Conditions are Combined

Recall that our target expression is:

(Country = 'England' OR Country = 'Scotland') AND (Height > '1000' OR Height < '500')

At the moment, the two nested Condition groups are combined by the default **OR** operation, not the **AND** as intended.

The bracketed blue lines indicated below, indicate that the conditions they surround are combined by an **OR** operator:

We need to change the top-level combination to be **AND**. This is accomplished by simply clicking on the leftmost angled line.

The filter expression now appears as follows:

Notice that the leftmost bracket line has changed to green and now contains an **&** symbol, to indicate that the enclosed conditions are combined using the **AND** operation.

When we **Run Query** we get the following result showing in the Drilldown Manager:

To view these results close the Drilldown Manager, by clicking on the icon and return to the drilldown.

Group Filter example

We will group by Mountain range, and then filter out any range which has less than 10 mountains.

First we need to Group by Range:

Σ Rows for Mountains (Table)

96 Groups

Grouping ▼

Ungrouped Columns

Search 🔍

- Country
- Height
- Name

Grouped Columns

Search 🔍

- ● Range

Group type: Normal ▼

▲ ● ▼	Country	Name	Height	Range
1	Wales	Plynlimon - Pen Pumlumon Fawr	752	Aberystwyth to Welshpool
2	Scotland	Ben Klibreck - Meall nan Con	962	Altnaharra to Dornoch
3	Wales	Yr Eifl	564	Anglesey and the Lleyn Peninsula
4	Scotland	Sgorr Ruadh	962	Applecross to Achnasheen
5	Scotland	Goat Fell	874	Arran and Holy Island
6	Scotland	Tinto	711	Ayr to the River Clyde
7	Wales	Aran Fawddwy	905	Bala to Welshpool
8	Wales	Arenig Fawr	854	Barmouth to Betws-y-Coed and Bala
9	Scotland	Heaval (Barra)	383	Barra to Barra Head
10	Scotland	Lochnagar - Cac Carn Beag	1155	Braemar to Montrose
11	Scotland	Askival	812	Canna, Rum and Eigg
12	Scotland	Farrmheall	521	Cape Wrath Peninsula
13	Wales	Hirfynydd	481	Carmarthen to the Vale of Neath
14	Scotland	Merrick	843	Carrick and Galloway
15	Scotland	Ben Cleuch	721	Central Scotland from Dumbarton to Montrose

Next we need to create a column with number of mountains per range.

Click right on **Name** then select **Insert Column > Aggregate** > **Count**.   As you see from the example below we now have a new column in the drilldown called **#Count (Name)** which also appears in the Ungrouped Columns list with a symbol to show it is an Aggregation:

Now we need to apply the **Group Filter** using the Count per range. Clearly this can only be done once the mountains have been grouped and the count calculated for each range. Complete the Filter details in the **Filtering** panel, selecting the **Post-Filter** tab and run the query:

Sampling

This capability allows a drilldown to be sampled at point of execution either regularly or randomly. You can also optionally limit the results returned to get a quick view of something.

The alternative is to open a drilldown of rows with a random sample of 100 rows to use this data to set up your filters, rules and expressions to then save them and apply them to the full set of data.

Background

If a drilldown has a query running that will take time to run you have the option to send the drilldown to the background to complete whilst you carry on with other work in Experian Pandora.   To set the options for this, go to **Home > Settings > My Preferences > Put Busy Drilldowns in Background.**   This will need to be set before running the query.
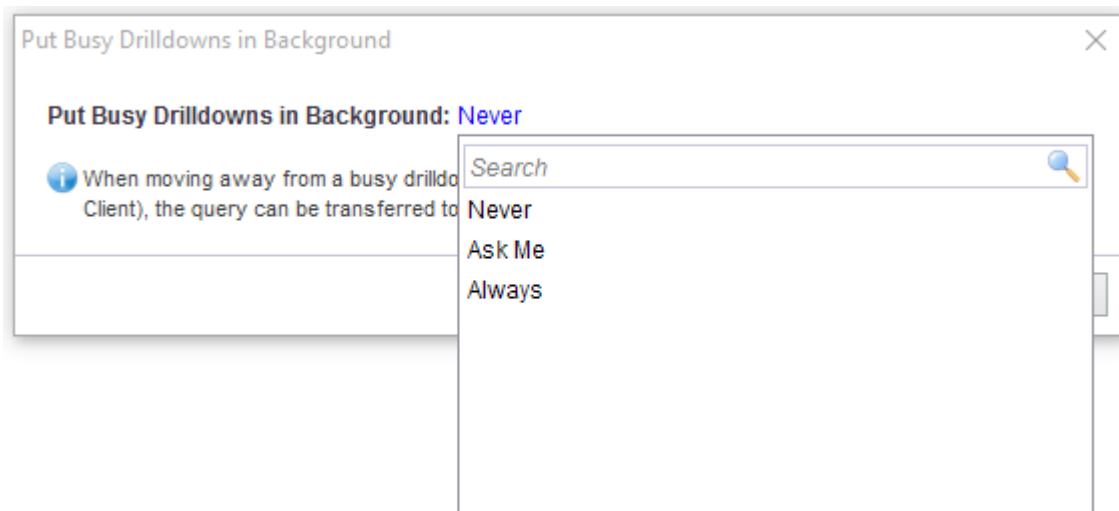
To use using the options whilst running a query, click the close **X** on the drilldown window and the query will run in the background. If you have the option set to never, clicking the **X** will close the drilldown completely.   The Background icon in the quick bar will now show a job running. To view all background queries that have been run or are currently running click on the icon and then double click the relevant line to view the results.

Quick Filtering

It is possible to quickly filter drilldowns, in some cases creating very complex filters, without even using the filter dialogs described in the previous examples.

Single Column Filter

In this example, we want to filter the Mountains table to only show mountains in **England**. In our drilldown scroll the data to show any row where the Country column is **England** and then right-click on the value, and select **Filter**.

Selecting **Keeping Selection** will create a filter for rows where the **Country** column is equal to **England**. Note that the filter details show in the drilldown title bar at the top.   The filter is also added to the Drilldown Manager in the Filter section:

If we had chosen **Removing Selection** then the opposite filter would have been applied and the result would show rows where the **Country** column is **NOT** equal to **England**.

Multi-Column Filter

In this example, we want to filter **Mountains** to only show mountains in **Scotland** in the range **Loch Maree to Lock Broom**.   In the drilldown for **Mountains**, we find a row that meets the above criterion and select the values **Scotland** and **Lock Maree to Lock Broom**:

Rows for Mountains (Table)

2,367 Rows

	Country	Name	Height	Range
313	Scotland	Stob Coire a'Chairn	981	Fort William to Loch Treig and Loch Leven
314	Scotland	Maol Chinn-dearg	981	Glen Shiel to Loch Hourn and Loch Quoich
315	Scotland	Mullach Coire Mhic Fhearchair - East Top	981	Loch Maree to Loch Broom
316	Scotland	Slioch	981	Loch Maree to Loch Broom
317	Scotland	Meall na Aighean	981	Loch Rannoch to Glen Lyon
318	Scotland	Lochnagar - Meikle Pap	980	Braemar to Montrose
319	Scotland	Beinn a'Chochuill	980	Glen Etive to Glen Lochy
320	Scotland	Stob Diamh - Stob Garbh	980	Glen Etive to Glen Lochy
321	Scotland	Slioch - Trig Point	980	Loch Maree to Loch Broom
322	Scotland	Beinn Eighe - Sail Mhor	980	Loch Torridon to Loch Maree
323	Scotland	Ciste Dhubh	979	Loch Duich to Cannich
324	Scotland	An Teallach - Glas Mheall Mor	979	Loch Maree to Loch Broom
325	Scotland	Stob Coire Sgriodain	979	Loch Treig to Loch Ericht
326	England	Scafell Pike	978	Central and Western Fells
327	Scotland	Beinn Dubhchraig	978	Inveraray to Crianlarich

Then we right-click on either of the selected values, select **Filter > Keeping Selection**. To view the details of the filter open the Drilldown Manager and select the Filter panel:

Rows for Mountains (Table)

Filter: Country is equal to *'Scotland'* and Range is equal to *'Loch Maree to Loch Broom'*

4 of 6 Columns

35 Rows (1.48%)

Filtering ▾   Reset ◷ Run Query

Pre-filter	Post-filter	Sampling

+ Add

⌐ + ☑ Country is equal to 'Scotland' Case Sensitive ✕
& + ☑ Range is equal to 'Loch Maree to Loch Broom' Case Sensitive ✕
└

	ΦCountry	Name	Height	ΦRange
1	Scotland	An Teallach - Bidein a'Ghlas Thuill	1062	Loch Maree to Loch Broom
2	Scotland	An Teallach - Sgurr Fiona	1060	Loch Maree to Loch Broom
3	Scotland	An Teallach - Lord Berkeley's Seat	1047	Loch Maree to Loch Broom
4	Scotland	An Teallach - Corrag Bhuidhe	1040	Loch Maree to Loch Broom
5	Scotland	Mullach Coire Mhic Fhearchair	1019	Loch Maree to Loch Broom

Optimised Filters

Experian Pandora automatically optimises quick filters based on selections by creating an expression using the minimum number of comparisons required to achieve the result.

Keeping and Removing Rows

To keep or remove entire rows then click the Row ID of the rows to keep or remove and use the Filter option:

Common Drilldown Capabilities

Drilldowns based on data, such as rows of a Table, or values of a Column, are provided with common drilldown operations for searching and cross-searching which can be used to navigate to other related drilldowns of information:

Menu	Option	Usage
Find Columns	Using this Value	Returns of Columns across all Tables in Experian Pandora that contain the selected value or values
	Using this Format	Returns of Columns across all Tables in Experian Pandora that contain the values matching the format of the selected value(s)
Find Values	With this Format	If the drilldown is a value-based drilldown such as values of a Column or Table, then this will filter all of the values for that object to find ones that match the format of the selected value(s)
	With this Format Globally	This will filter the list of all values in the entire Experian Pandora database for values that match the format of the selected value(s)
	With a Similar Format	If the drilldown is a value-based drilldown such as values of a Column or Table, then this will filter all of the values for that object to find ones that match the format of the selected value(s). A similar format is one that has a matching *simple format*
	With a Similar Format Globally	This will filter the list of all values in the entire Experian Pandora database for values that have a similar format of the selected value(s). A similar format is one that has a matching *simple format*
	Containing this Format	If the drilldown is a value-based drilldown such as values of a Column or Table, then this will filter all of the values for that object to find ones that have a format that embeds the format of the selected value(s)
	Containing this Format Globally	This will filter the list of all values in the entire Experian Pandora database for values that have a format that embeds the format of the selected value(s)
	Sounding Similar	If the drilldown is a value-based drilldown such as values of a Column or Table, then this will filter all of the values for that object for values that phonetically match the selected value(s)
	Sounding Similar Globally	This will filter the list of all values in the entire Experian Pandora database for values that phonetically match the selected value(s)

Naming Drilldowns

Individual drilldowns can be named and given a description. The name replaces the generated one for the drilldown in its title bar as well as updating the Drilldown History. The description can only be viewed in the history.

Renaming a drilldown does not implicitly put it into the history, this happens automatically depending on your preferences and the type of drilldown.

From the drilldown window, click on the Edit icon on the toolbar and select **Edit > Details** from the drop down. A simple dialog appears to allow you to change the name and description for the drilldown. This name will now appear in the Drilldown History making is easy for you to find the Drilldown again.

Drilldown History

The majority of drilldowns that are executed are automatically added to the user drilldown history.

To access the history drilldown, simply click on the history items entry on the user status panel at the bottom left of the Experian Pandora application window.

History is saved within 10 seconds of any change and is persistent. Upon drilling down, the history is presented in most-recent-first order. Drilling down to any history item presents the drilldown, and updates its timestamp to reflect its re-invocation. Like this, most often used items are always near the top. To stop the order of the items changing, simply sort the order by name.

Hovering over a history line entry gives you detailed information about the entry and right-clicking provides with a menu to view, edit and delete the history entry.

The history entries are colour-coded as follows:

Colour	Description
Red	The drilldown is no longer valid. It will also be entirely in italics.
Black	The drilldown has been grouped, aggregated or filtered (or any combination of)
Blue	The drilldown has been sorted or has expression columns added only (but not grouped, aggregated or filtered).
Grey	It is a plain drilldown and nothing has been done to it.

Example of recent History:

	Name	Drilldown Type	Timestamp	Filter	Sort Columns	Group Columns
1	Rows - Mountains (Table)	Rows	Today @ 17:11:00	Filter: Country is equal to 'Scotland' and		
2	Rows - Mountains (Table)	Rows	Today @ 17:08:52	Filter: Country is equal to 'Scotland' and		
3	Rows - Mountains (Table)	Rows	Today @ 17:05:07	Filter: Country is equal to 'Scotland' and		
4	Rows - Mountains (Table)	Rows	Today @ 17:04:11	Filter: Country is equal to 'Scotland' and		
5	Columns (All Versions)	Columns	Today @ 17:02:53	Filter: Columns with the value 'Loch Mar...		
6	Rows - Mountains (Table)	Rows	Today @ 17:02:27	Filter: Country is equal to 'Scotland'		
7	Rows - Mountains (Table)	Rows	Today @ 16:59:12	Filter: Country is equal to 'England'		
8	Rows - Mountains (Table)	Rows	Today @ 16:53:56	Group Filter: Count Name is greater tha...		Range
9	Columns (All Versions)	Columns	Today @ 16:40:20	Filter: Columns using the Format AAAA...		
10	Rows - Mountains (Table)	Rows	Today @ 16:39:15			
11	Rows - Mountains (Table)	Rows	Today @ 16:25:11			Range
12	Rows - Mountains (Table)	Rows	Today @ 16:23:34			Range
13	Rows - Mountains (Table)	Rows	Today @ 16:22:17			Range
14	Rows - Mountains (Table)	Rows	Today @ 16:05:37	Filter: (Country is equal to 'England' or .		
15	Rows - Mountains (Table)	Rows	Today @ 16:00:33	Filter: (Country is equal to 'England' or .		

Column Formatting

The display format of drilldown columns can be specified by editing the column heading. Right-click on the required Column, select **Edit,** which produces the edit box below. The default format of an actual column can also be edited using the edit option in the Explorer.

As well as giving the column an alias you can change the appearance. It is possible to format the content of the column with custom colours, rendering, display formatting, widths and so on. The result is persistent in the History, saved Views, note attachments etc.

The custom format can even add text into data values for display purposes. We are going to format the height with a comma separator for the thousands and add in Metres after the number.

Remember, this is for <u>formatting</u> only. The underlying data has not changed but if you export the data you will get what you see. Formatting has no effect on the generation of a **mapping** document however. If you need to ensure the format as part of a mapping, you should actually transform the data instead.

You can also set the sort options for the column such as whether sorting should be case sensitive and whether the Language setting should be used to collate items together when sorting.

Collaboration

The best results on projects are achieved as a result of collaboration and data-related projects are no exception. Experian Pandora provides a flexible, enterprise-wide environment which allows multiple people to work together regardless of physical separation to get the most out of their collective skills and knowledge.

It is increasingly common for project teams to be composed of people based in different offices or even working in alternate time zones. Experian Pandora provides a secure, multi-user environment allowing colleagues from anywhere to work together simultaneously on the same data and metadata with real-time access to everything.

All items regarding Collaboration can be found on the explorer at **Collaboration**

Meetings
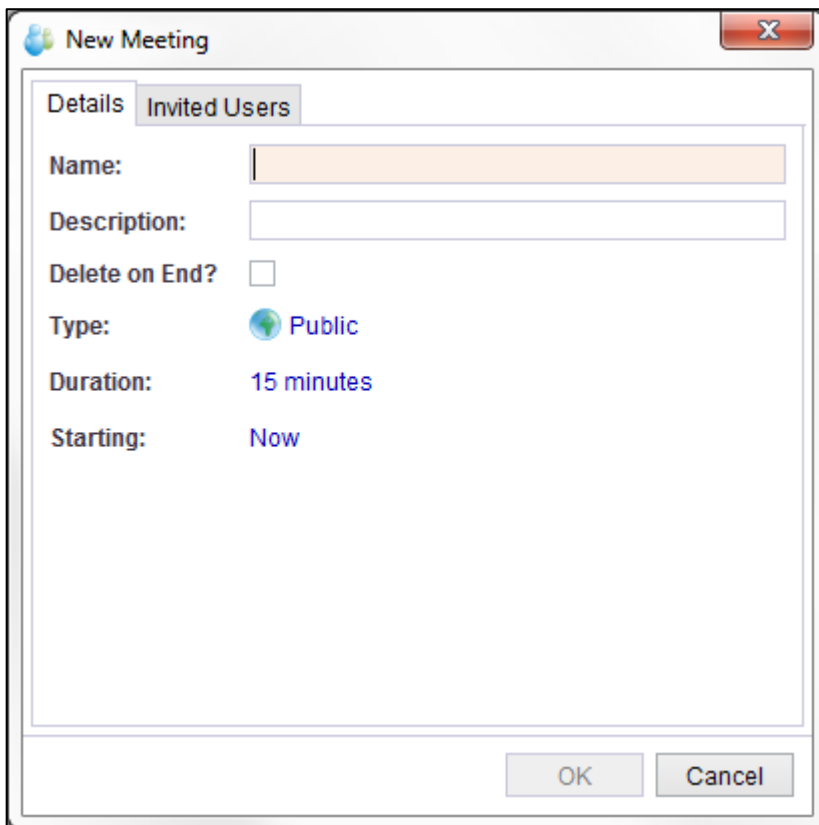
Experian Pandora allows users to collaborate in real-time in the following ways:

- Schedule online meetings and participate in those meetings
- Conduct a private chat with one or more other on-line users
- Share drilldowns and documents with meeting/chat participants

Creating a Meeting

Expand **Collaboration** in the explorer. Right-click on **Meetings**, and select **Create a Meeting** from the menu.

The **New Meeting** dialog appears:

Name

Name of the meeting e.g. **Weekly Team Meeting**. This field is mandatory.

Description

Enter an optional description for the meeting.

Delete on End

If enabled, the meeting (and all of its message history) will automatically be deleted as soon as the meeting ends.

Type

Meetings may be **Public** or **Invitation Only**.

Public meetings will be seen by all users, regardless of whether or not they have been invited.
Invitation Only meetings will only be seen by users specifically invited to the meeting.
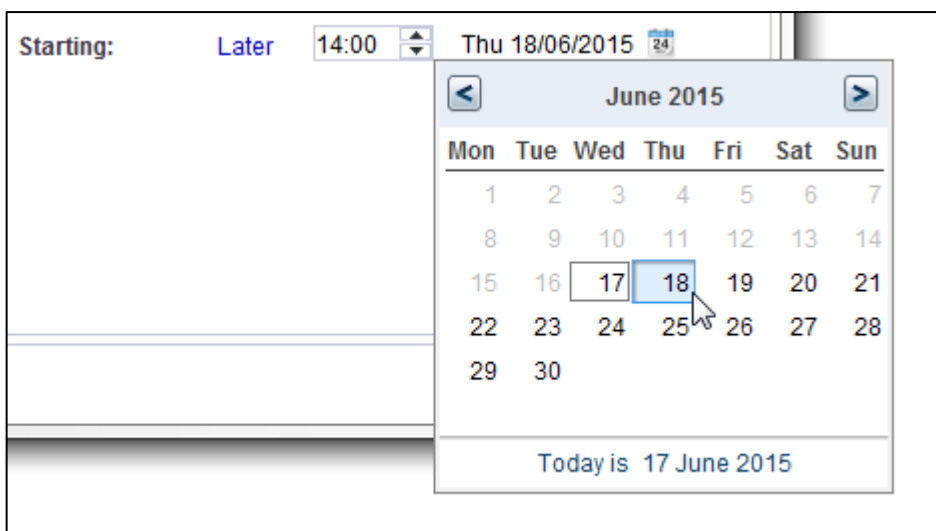
Duration

A number of default values are provided along with additional options that allow you to make the meeting perpetual. Selecting **Custom** will add a control that allows you to specify an arbitrary duration in hours and minutes.

Starting

The scheduled start time of the meeting which can be either **Now** or **Later**. Selecting **Later** will add controls that allow you to specify the start time and date of the meeting:

Invited Users tab

This tab allows you to choose which users are invited to the meeting.

Clicking on one or more users from the **Available Users** list and clicking the **Add** button will add them as meeting invitees.

Clicking on one or more users from the **Invited Users** list will remove them from the invited list.

NOTE: Any user may join a public meeting but they will not be alerted to the meeting's existence unless they are specifically invited.

Viewing Meetings

You can only see meetings under the following conditions:

- You are the meeting owner, i.e. you created the meeting
- The meeting is **Public**
- The meeting is **Invitation Only** and you are invited.

Available meetings are displayed both in the explorer and in a drilldown. Expanding **Collaboration > Meetings** will display them on the explorer. Clicking on the icon for the **Meetings** explorer item will display the drilldown:

Meeting Actions

Meetings displayed in the explorer or in the Meetings drilldown have identical right-click options. If you are the meeting owner you will see the following actions:
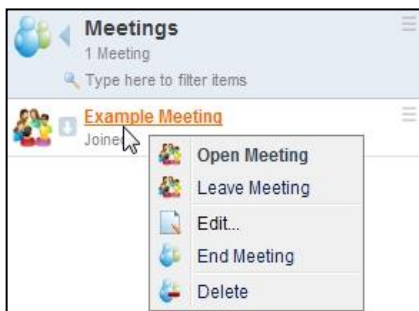
If you are not the meeting owner, but are invited to a meeting, you will see the following:
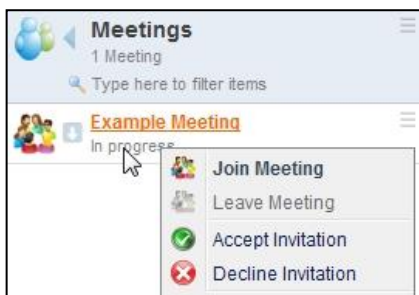
If a meeting is **Public** and you are not the owner and are not specifically invited to the meeting, you will see the following actions:
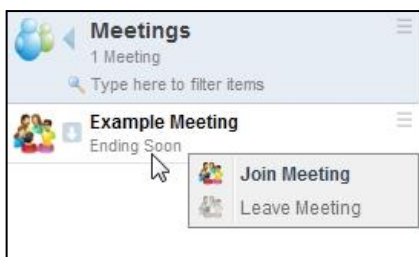
Participating in a Meeting

To join a meeting, right-click on the meeting and select **Join Meeting**
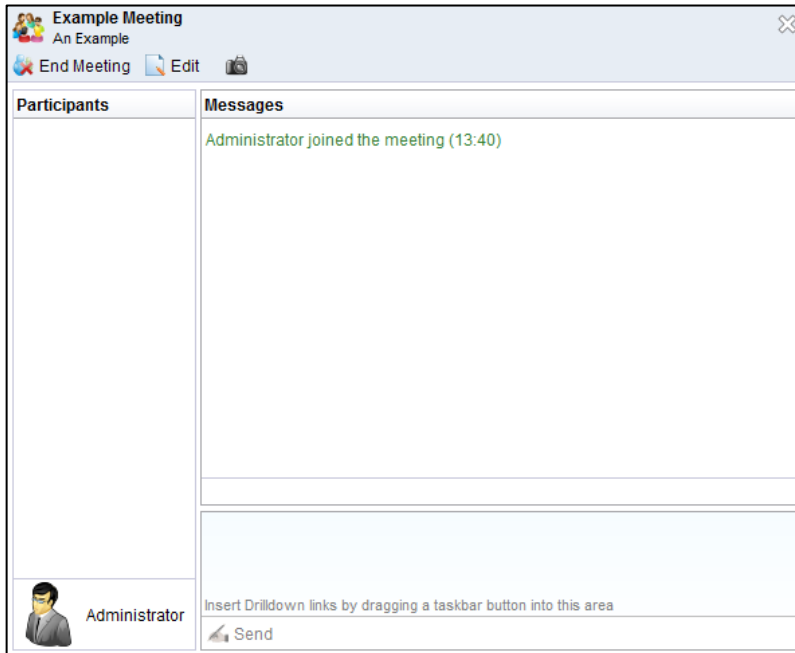
Participants

The list of users currently participating in the meeting

Messages

This shows messages that users have submitted and system-generated messages (e.g. users joining/leaving, meeting ending etc.)

Message entry area

This is the text area where you can type messages into the meeting. Pressing the RETURN key, or pressing the **Send** button will submit a message. Pressing Ctrl-RETURN will add a new line to the message that you are typing. You can also drag drilldowns from the Experian Pandora desktop task bar onto this area to provide other users with a link, which if clicked on, will present them with the very same drilldown.

Leaving a meeting

You can leave a meeting at any time by:

- Finding the meeting on the explorer or the Meetings drilldown and selecting the **Leave Meeting** menu action
- Closing the meeting window

Sharing Drilldowns and Documents

Links to drilldowns and documents can be included in your meeting messages by simply dragging any drilldown from the desktop taskbar onto the meeting message entry area.

The taskbar is located at the bottom of the desktop area. Here we can see a drilldown being dragged from the taskbar and about to be dropped onto the message area. The meeting window prompts you to insert the drilldown:
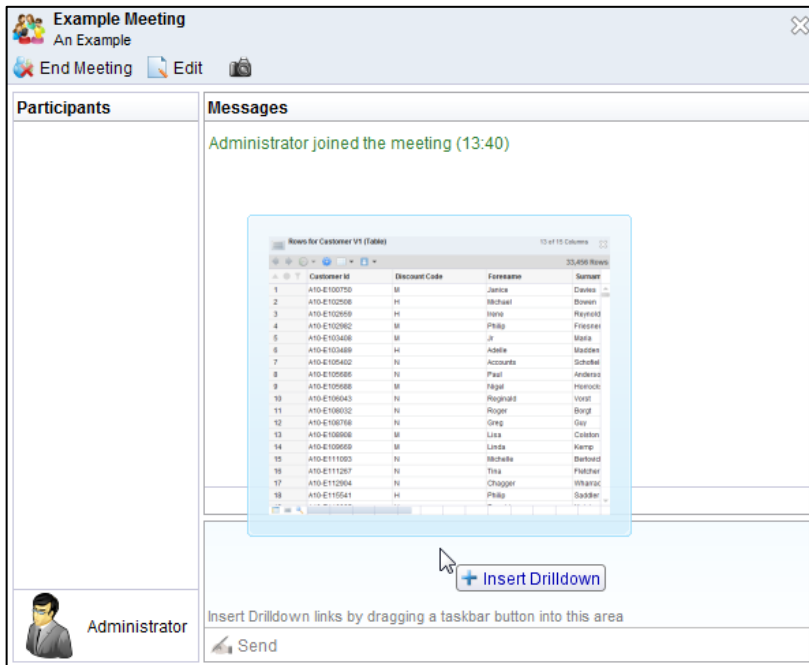
Submitting the above message will then share the drilldown link with all the other participants of the meeting. Another user clicking on this link will open the drilldown on their desktop.

In addition to standard drilldowns, documents (e.g. Mapping Document, Table Quality Report, Relationship Report etc.) and Venn Diagrams can also be shared in the same way.

Here we see what the meeting window looks like after the drilldown has been dropped onto it:
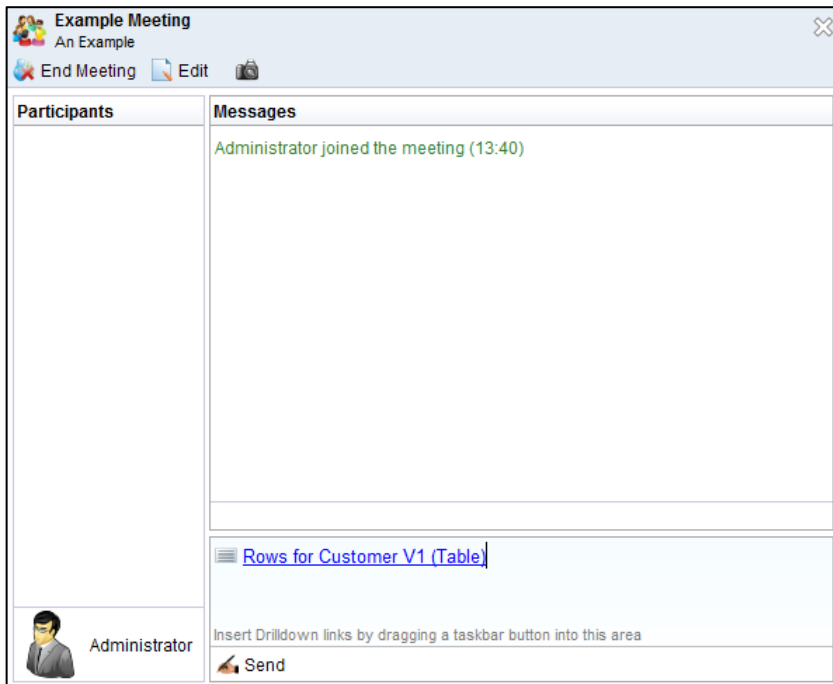
Multiple drilldowns can be attached at the same time:
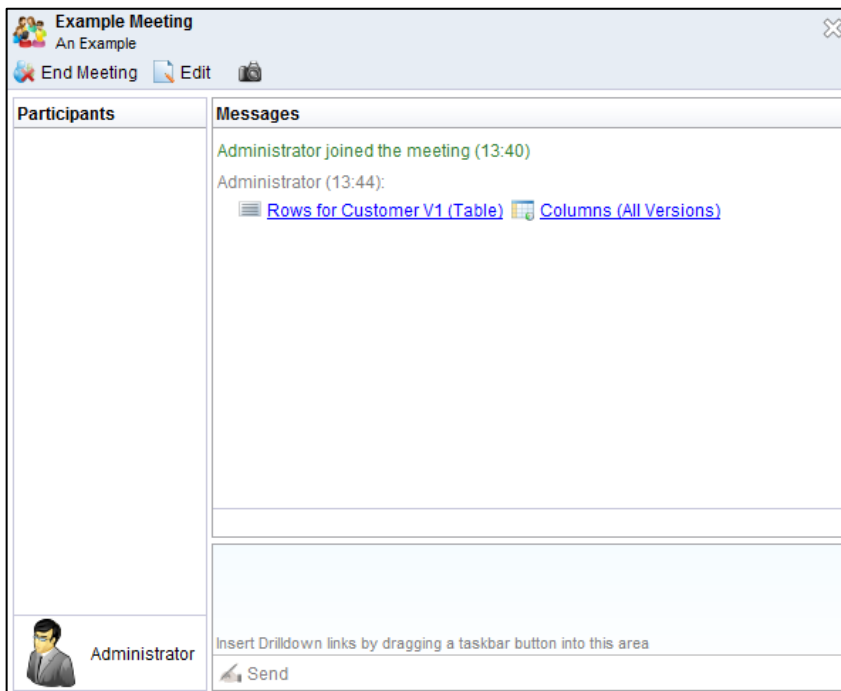
Ending a Meeting

Meetings will only end under the following circumstances:

- The meeting is due to end, according to its start time and duration, **and** there are no participants in the meeting. A meeting will never end automatically if users are still participating.

- The meeting owner, or an administrator, manually ends the meeting. This is achieved by right-clicking on the meeting in the explorer or the Meetings drilldown and selecting the **End Meeting** action from menu.

If the **Delete on End** option was enabled when the meeting was created, the meeting will be deleted immediately, including its message history.

If an ended meeting has not been deleted, the history of meeting messages can be reviewed by the owner of the meeting, any invitee, or indeed anyone at all if the meeting is **Public**. To review the meeting messages, right-click on the meeting and select **Review Meeting History** from the menu.

Meeting messages cannot be submitted once the meeting has ended.

Messaging

Private Chat

A private chat between two users is simply an invitation-only meeting (with two invitees) that starts immediately and ends when both users leave the meeting.

You can start a chat with a user wherever that user is displayed in Experian Pandora. This is easily achieved from **Collaboration > Messaging** on the explorer. There are other places a user will be displayed in the explorer or a drilldown, such as when the Steward of an object is displayed in the metadata for an object. In all cases, you can right-click on the user and select **Chat to** from the menu.
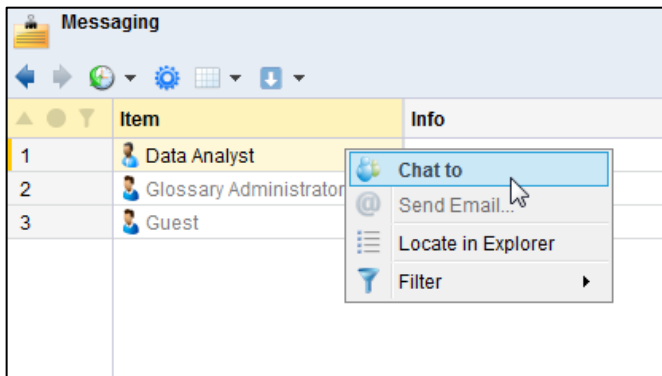
Private chats will appear in list of meetings for the two users but cannot be edited or explicitly ended. They will not appear in the list of meetings for any other user.

NB: If you select **Locate in Explorer** you will be taken to the details of the user, which if set up, will provide contact details and other information about the user.

Notifications

When you have been invited to a meeting, the **Meetings** button in the Quick Bar at the top of the screen will show in indication of this by displaying a small red icon:
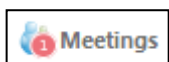
If a user has sent a message in a meeting you are currently participant in, the drilldown icon along the bottom will flash yellow. If you are invited to a meeting, you will similarly see a meeting alert appear. If a user requests a private chat, you will see a notification appear in the bottom-right of the desktop area. Simply click this notification to join the private chat immediately.
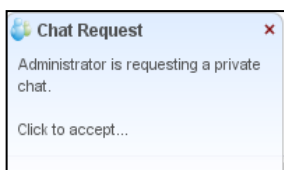
Sending data from a drilldown to a colleague

You can send any drilldown you are viewing to an on-line colleague, which results in an immediate chat session with the drilldown attached. From the drilldown toolbar, select the **Edit** button and select **Send to** and choose the available on-line user to send it to. If the user is off-line and you have an email address set up for that user, your default email service will create a new email to the users email address with a link to the drilldown which when clicked will open a web browser.
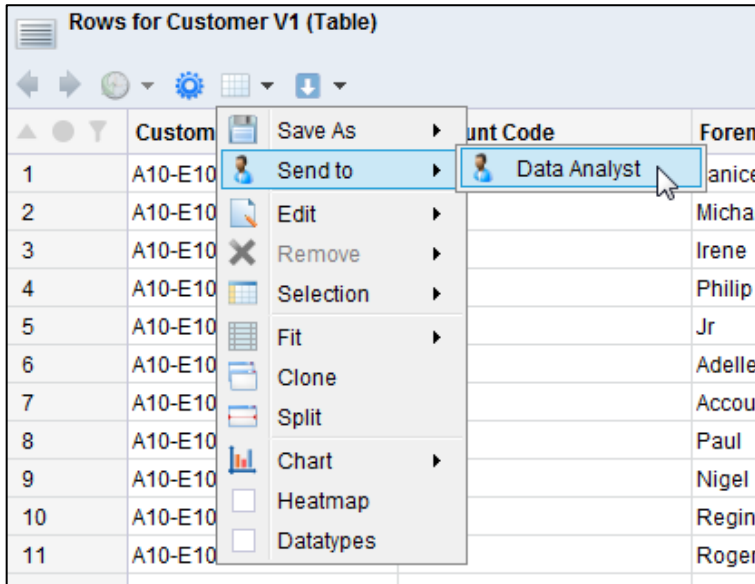
Emailing another user

Provided the user has a configured email address, you can email them directly from within Experian Pandora. For example, to email a colleague, right click on the **User's Name** and select **Send Email** from the right click menu. This will invoke your standard email application already filled-in with the email address of John Smith. This is useful for contacting users modelled in Experian Pandora but not actual users of the application.
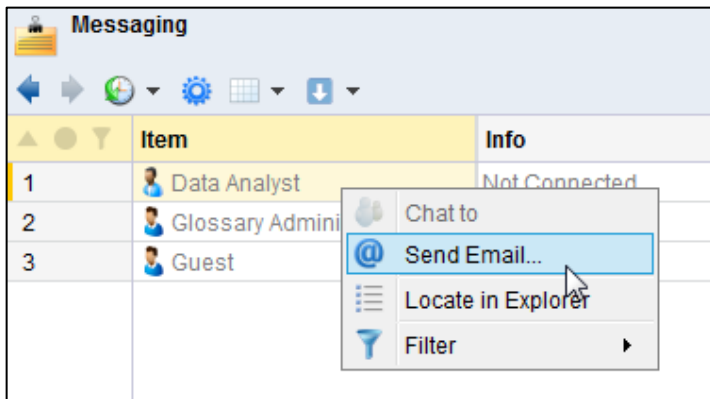
Notes

As data analysts review the results of the profiling and investigate issues, they need to be able to record context-specific notes and enter into discussions with other analysts. This essential part of the analytical data quality process is supported by Experian Pandora through the use of **Notes**. Any number of notes can be added to Schemas, Tables, Columns, Relationships, Dependencies, Keys and more, provided you have the rights to annotate things and the permission to annotate a given item.

Each note is given a **Subject**, **Category, Impacts, Priority and Status.**

There is no limit to the number of notes that can be created on any item. Once a note has been created, it can have additional detail items added to it, much like a bulletin board. Multiple users can interact with the note thread, adding detail or additional attachments as they go.

Notes can also be issued to users to effect workflow.

Creating Notes

A note can be created for a table (or any other object that supports it) by navigating to the corresponding item in the explorer. In the example of a Table, right-click on it and select **Notes > Add**.

You can attach any Drilldown by dragging it from the taskbar directly onto the **Attachment:** dropdown or by selecting one from the dropdown.

Retrieving Notes

It is possible to determine if an object has directly attached notes by examining it on the explorer. If it does, it will have a note icon attached to it. Simply click on the icon to see a drilldown of all attached notes.
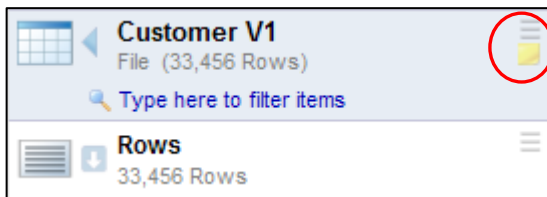
This will display the notes drilldown for the respective object selected.

	Subject	Category	Impact	Priority	Status	Progress
1	test	Audit	High	Medium	Under Investigation	50%
2	Customer ID clean	Compliance	Unimportant	Medium	Open	0%

Notes for Customer V1 (Table) — 18 of 30 Columns — 2 Rows

Note Detail

Right-click on one or more notes and Select **Entries**. A special type of drilldown will open, which is essentially a report of the information held in the note.

Viewing Attachments

You can also see an **attachment**, where we attached some example rows for evidence of the problem. To view the attachment in the note, right-click on it and select **View Attached Drilldown** or double-click the attachment. This is best done in split-view.

Changing the note heading

Right-click on the note heading and select **Edit Note**

This allows the author to change the Subject, Category, Impact, Priority, Status and who it is Issued To.

Changing the note entry

You can change note detail if you are the person who created it, by right-clicking on the detail header or text and selecting **Edit Entry.** You can now edit the Note Text then click **Save Changes** to make the change permanent. The note will be re-issued.

Adding more Entries

Following the bulletin-board theme, users can add Entries to a note provided they have the appropriate permissions. Right-click anywhere on the detail drilldown and select **Add Entry**, to open the **Add Entry to Note** dialog. Once the Entry is added, it can be seen at the end of the note detail report, in historical order. Edits from other users are shown in a slightly darker shade, to denote that they are not editable by you. All edits are time stamped so if you Edit a previous detail it will show the time and date of the last edit:

Entries
1 Note

Id:	**3**	
Subject:	**Customer ID clean**	
Table:	**Customer V1**	
Impact:	**Unimportant**	
Category:	**Compliance**	
Status:	**Open (0%)**	
Priority:	**Medium**	
Last Edited:	**Data Analyst- 15 October 2014 22:26:16 BST**	
Creator:	**Administrator**	
Issued to:	Administrator	15 October 2014 22:26:16 BST by Data Analyst
	Administrator	15 October 2014 22:24:41 BST by Administrator
	Administrator	15 October 2014 22:18:37 BST by Administrator

Last Edited: **Administrator - 15 October 2014 22:24:41 BST**   **(Edits=1)**
Attachment: **Rows for Customer V1 (Table)**

To match this column to the one in Purchase Order Header, the column will need to be edited using a Before function.
You have the option to change the colour of the font, and to use **BOLD** and/or *italics* to show any changes if you *want to :)*

Last Edited: **Data Analyst - 15 October 2014 22:26:16 BST**   **(Edits=0)**

And when someone else adds comments to the note the date and time of the comments will be added to the History of the note.

37 Rows

In the above example Note the Administrator edited their comments after the Analyst added theirs. To keep a dialog flow it is suggested that **Add Entry** is used when responding on a note rather than **Edit Entry.**

Note that once the change is made, reference to the number of edits made is shown on the Note Entry header. The original note detail entries are actually retained for historical purposes and can be viewed in Explorer under the Table Name. Go to Notes > Name of Note > Note Changes. These can be viewed in the Explorer or as a drilldown

Last Edited: **Administrator - 15 October 2014 22:24:41 BST**   **(Edits=1)**
Attachment: **Rows for Customer V1 (Table)**

Details of Note Entries and Note Changes

Full details of all Note Entries and Note Changes can be viewed in the Explorer or via drilldowns. Under the relevant table, select Notes and then the name of the note you wish to view. Full details of the note are shown as you click through the Explorer or down through the drilldowns, including the ability to open the relevant attachments where they are located.

Issuing Notes

Once a Note has been created you can issue it to a user by right-clicking on the note and selecting Issue To and choose the user. More than one note can be issued to a user at one time by highlighting the relevant notes and selecting the user to issue them to using the right-click menu option.

Viewing Issued Notes

To view notes that have been assigned to you, just click the **Issues** button in the Quick Bar.

Notes issued to a particular user can be listed using the option provided for a User by right-clicking on the user in the Explorer. Select **Settings > Users.** Right-click on one of the Users and you can select **Notes issued to User** or **Notes issued by User,** producing a drilldown of notes issued to or from that user.

The history of assignments is clearly shown when looking at the detail of a note:

Viewing multiple notes at once

Select two or more notes from a notes drilldown, right-click and select **Note Detail**. A composite note detail drilldown report will appear with each of the notes detailed sequentially.

Consolidating notes

As you can see, you cannot remove Entry items, but you can delete entire notes. As part of your management of notes, you can choose to allow free for all, bulletin board style interaction on notes over a long periods of time, and then finalise the information. Quite often, the final (or one of the final) note detail entries is the definitive result. Experian Pandora provides an additional option to right-click on a note, and select **New Note from This** .

This brings up the note editing window pre-populated with all the information from that note entry, and only the text and attachment from the note detail that the menu action was invoked from.

It is effectively a quick way to copy and paste all the details you need into your new note from an existing long note trail.

Exporting Notes

To export a note as a web page report (in HTML format), right-click on the required note(s) and select **Report**. This will open the note in a report format showing the information about the table, the Issue History and the Note Entries.   Click Save As on the Report toolbar to open the Export Document dialog. You will only be offered Web Page or PDF as the output format.
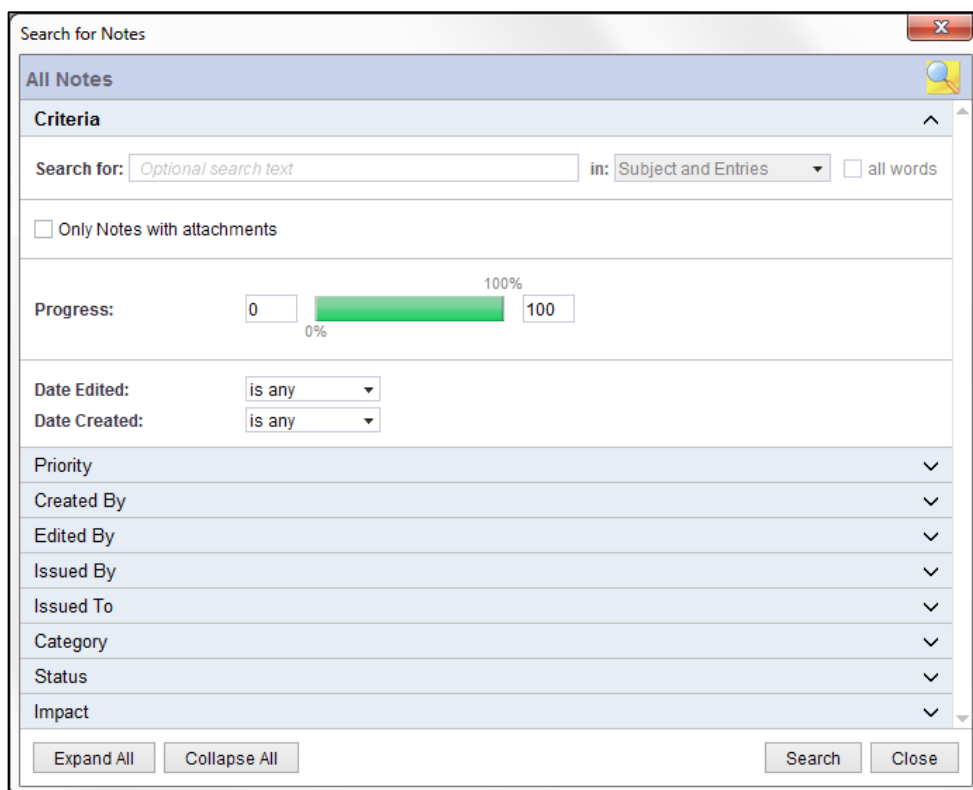
Searching Notes

The note search facility allows the analyst a context-specific, multi-criteria search which can return notes on dependent objects. The results of a note search can be exported, meaning that the final report of the analysis work often requested by project sponsors can be built up as the project progresses, ready to be exported later. The search criteria can be used to include only "conclusion" type notes, thus avoiding the inclusion of work-in-progress discussions in the final report.

Notes for a given object can be viewed from the object itself, as well as notes beneath the object – for example, clicking on a Table allows you to see notes just for the table or for all Columns beneath the table as well.

A general purpose search capability for Notes is also provided that allows detailed searching either across the whole repository, or localised to a given object (e.g. a Table)

The Note Search window can be accessed from any object supporting notes, or by right-clicking on **Home > Search > Search Notes** or Collaboration **> Notes** in the explorer, right-click on Notes and select **Search Notes**

Search For
This is the text or pattern to search for in the notes database, with each word being a separate search term.

Search In
You can choose to search only the Subject line, the note entries or both with this dropdown control.

Search for all words
This treats the search string as an exact string to search, i.e. looking for all the words together as opposed to treating them as individual search terms.

Only Notes with attachments
Optionally only consider notes that have attachments in their detail.

Progress
The progress can be altered to narrow your search. E.g. if you only want to look at those Notes which are below or above a set Status or even between a set of Status percentages.

Date Edited

Restrict the search to notes edited after or before a given date, or between dates.

Date Created
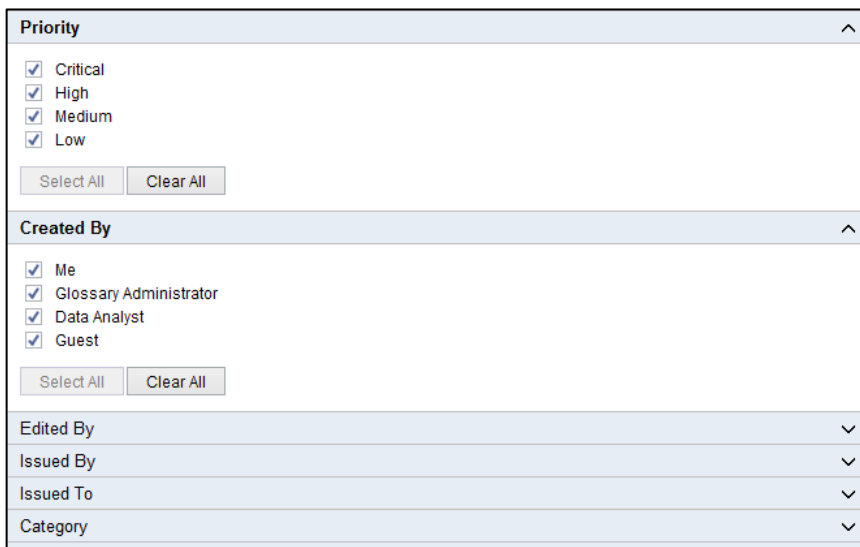
Restrict the search to notes created after or before a given date, or between dates.

Advanced search options

You can choose specific options and restrict the areas of the search:

- Priority
- Created By
- Edited By
- Issued By
- Issued To
- Category
- Status
- Impact

Clicking on any of the advanced search options, e.g. Priority and/or Created By will open the list for you to restrict your search selections:

Managing Note Categories

The categories for notes are definable repository objects in the Experian Pandora repository. From the explorer, expand **Settings > Note Categories** to create and delete categories and to view notes associated with a specific category.

Managing Note Impacts

The impacts for notes are definable repository objects in the Experian Pandora repository. From the explorer, expand **Settings > Note Impacts** to create and delete impacts and to view notes associated with a specific impact.

Managing Note Statuses

The statuses for Notes are definable repository objects in the Experian Pandora repository. From the explore, expand **Settings > Note Statuses** to create and delete status levels

Alerts

Experian Pandora can alert users when specified events occur, for example, when a Table has failed a rule. You can set configurations to:

- define the events you want to send Alerts for
- provide Alerts to selected users
- specify whether the notification is for Reference or Requires Acknowledgement.

Alerts requiring acknowledgment will be presented to the user with a pop-up on their screen when it is their time in the process and have to be acknowledged before the user can continue to use Experian Pandora. Alerts that require Acknowledgement have a time period for a response. If a user does not acknowledge the Alert within their time slot, the notification will disappear from their screen but they can still acknowledge the alert from their My Alerts drilldown.

Alerts can be sent and received internally in Experian Pandora or externally through email. External users who do not use the Experian Pandora interface can be alerted provided an email address has been provided and configured. The Administrator can configure an SMTP server under **Server Settings** > **Communication** > **SMTP properties**.

Receiving Alerts

Viewing Alerts

Each user has their own drilldown/inbox for the Alerts that they receive. This can be accessed from the quick bar or through the explorer under **Collaboration**:
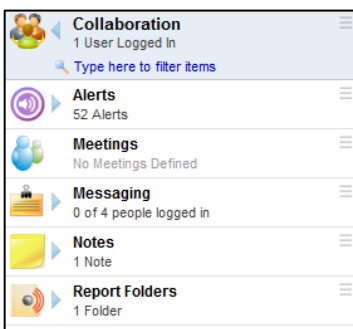
You can view your alerts in the Explorer by clicking on the Alerts icon or view the drilldown by right clicking on **Alerts** and selecting on the menu option. The quick bar will display the number of unread Alerts and open the My Alerts drilldown:
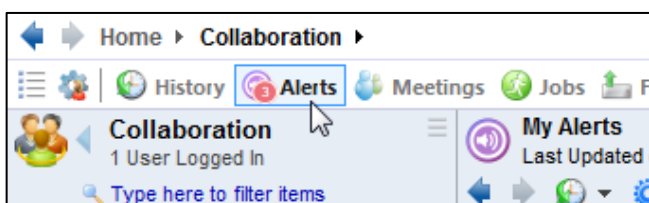
Right click Alerts in the Explorer to access the different Alerts drilldowns. My Alerts drilldown will list any Alerts you have received. **In Progress** are any Alerts that are currently progressing through the roster, where you are either scheduled later in the roster, or which you have not Acknowledged within the set timeframe. Administrators can see any alert that has ever been sent to any user, whether or not they are on the rosters.
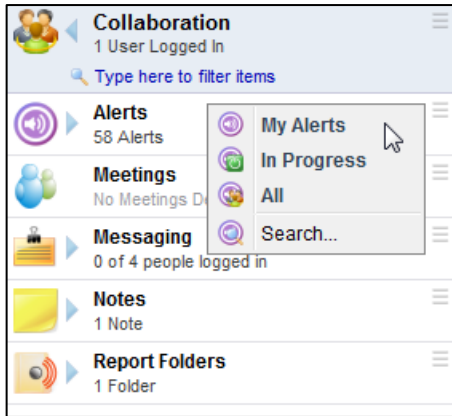
You can navigate between Alerts drilldowns using the Alerts menu button on the drilldown menu bar:
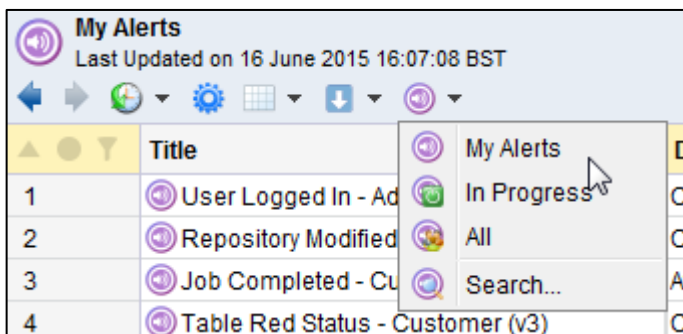
To view the detail of an Alert, right-click the relevant entry in the drilldown and select **View…**:
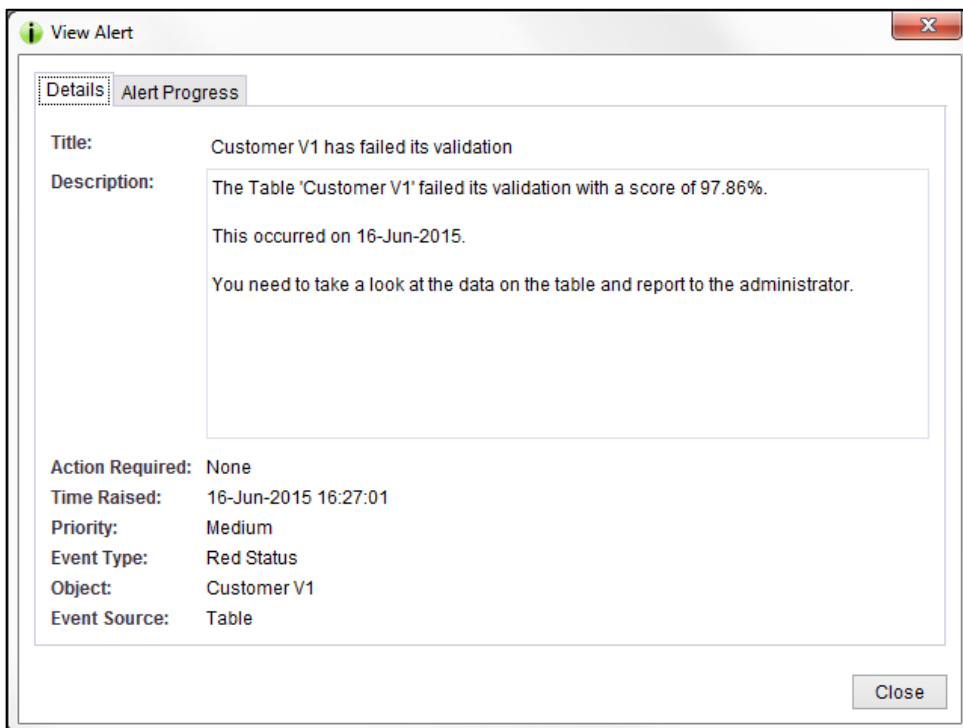
Acknowledging Alerts

When an Alert requires Acknowledgement and you are the current user in the Roster the Alert will pop-up on your screen. Before you can continue to use Experian Pandora you must acknowledge the Alert.   The Alert will remain on your screen for the duration that is shown in the Action Required line or until you acknowledge it.
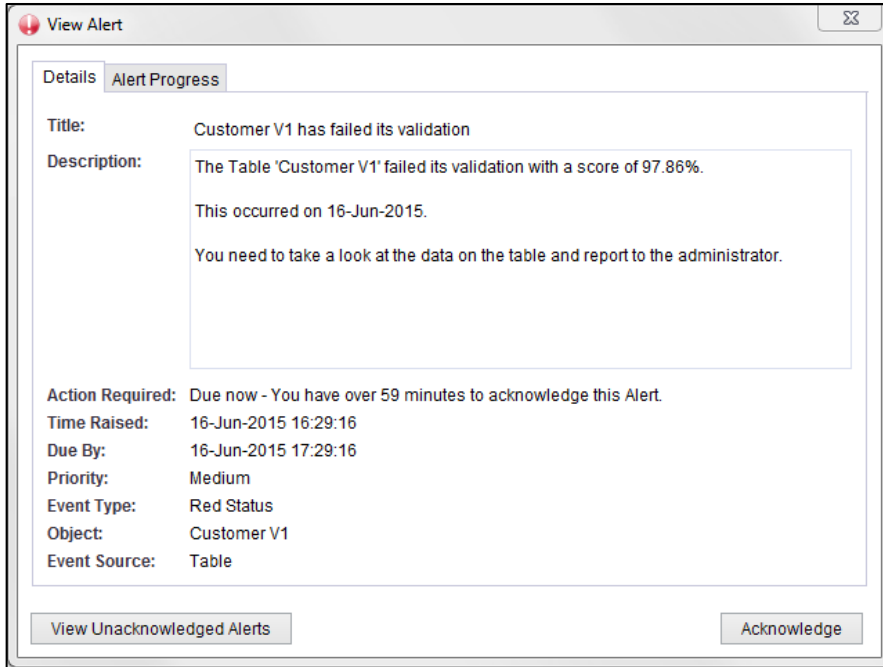
If you miss an Alert that required Acknowledgment and it has moved to the next person in the Roster, you can view it from the **My Alerts** drilldown and Acknowledge it from there.

The **Alert Progress** tab on the dialog shows the current state of the Alert and where it is on its journey.

External Response to Alerts

If a user receives an Alert via email or text message they can respond by clicking on the link provided that will access the Experian Pandora server via the REST API.

Configuring Alerts

To set up an alert, you need to configure a template, a list of users to receive the alert (a roster) and an event to trigger the alert.

Alert Templates
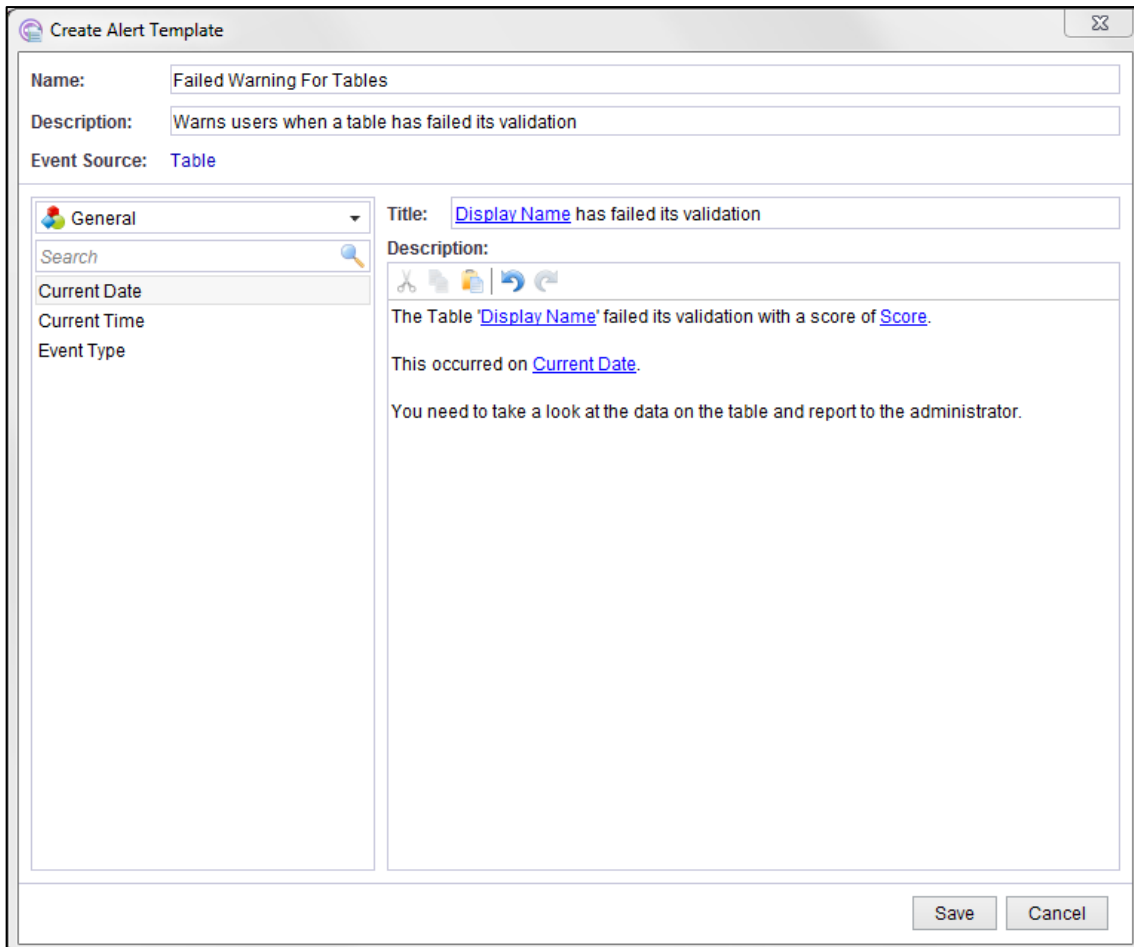
Alert Templates are used to set the text that will be displayed when a user receives the Alert. They are located in the Settings section of the Explorer. To create a template, go to Settings > Alerts, right-click Template and select New…
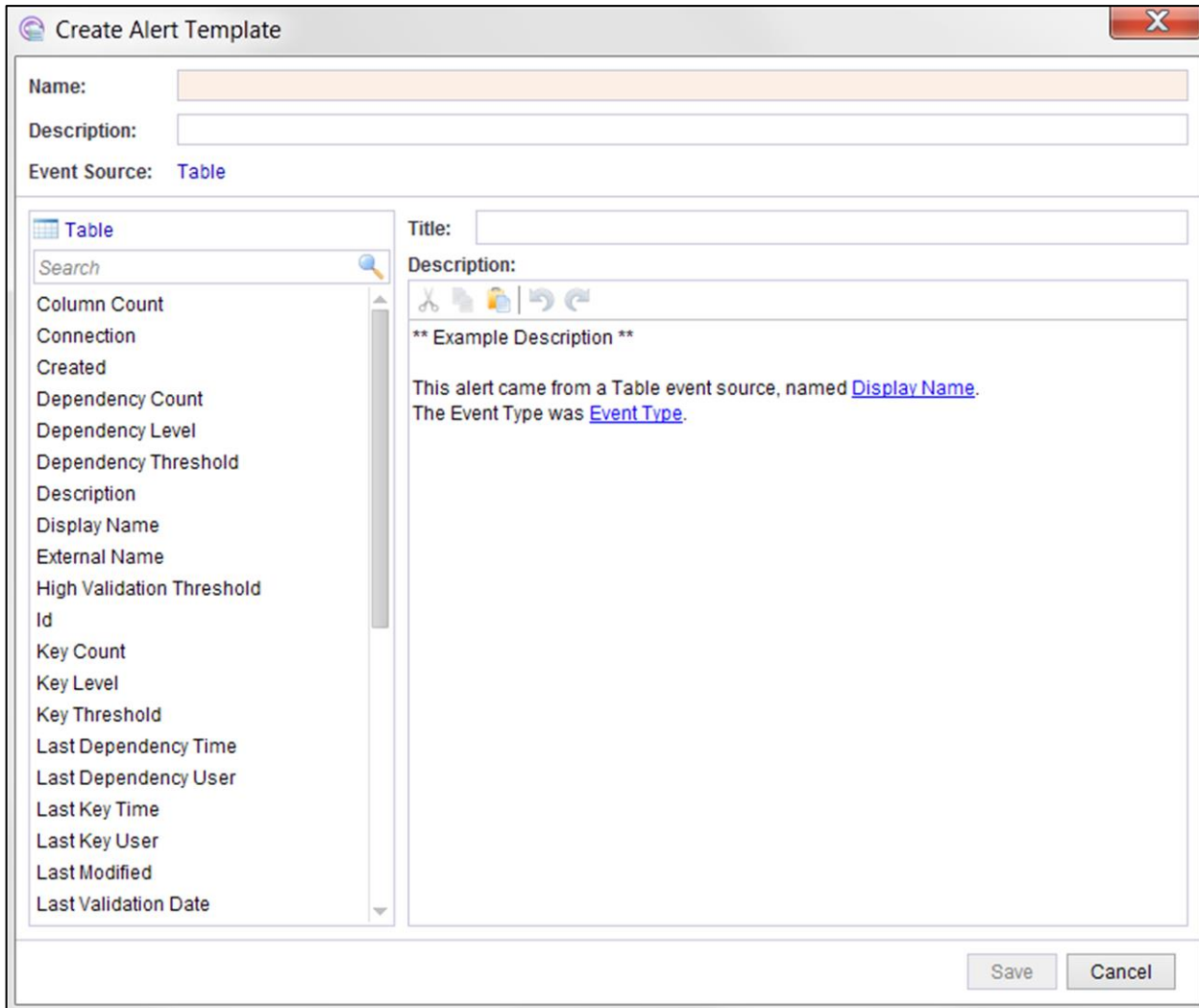
The Create Alert Template can also be opened when setting Event Handling, detailed below.

To create an Alert Template, set the type of object that the template will be used for, e.g. Table, Column Job etc., in Event Source, and define the text that will appear in the title and description of the Alert. There is an example provided in the dialog box, which will need to be removed when you create the new Alert description. Leaving the title and description blank will mean the example text will be used.

On the left is the list of attributes for each of the objects. When the Alert is sent it will be automatically populated with the relevant detail. Each attribute can be added by either double clicking it, or dragging the selection from the list.

An alert sent with this Template could appear with a title of:

'Customer V1 has failed its validation'

The description would look similar to:

'The Table 'Customer V1' failed its validation with a score of 87.64%.

The validation occurred on 25th May 2015.
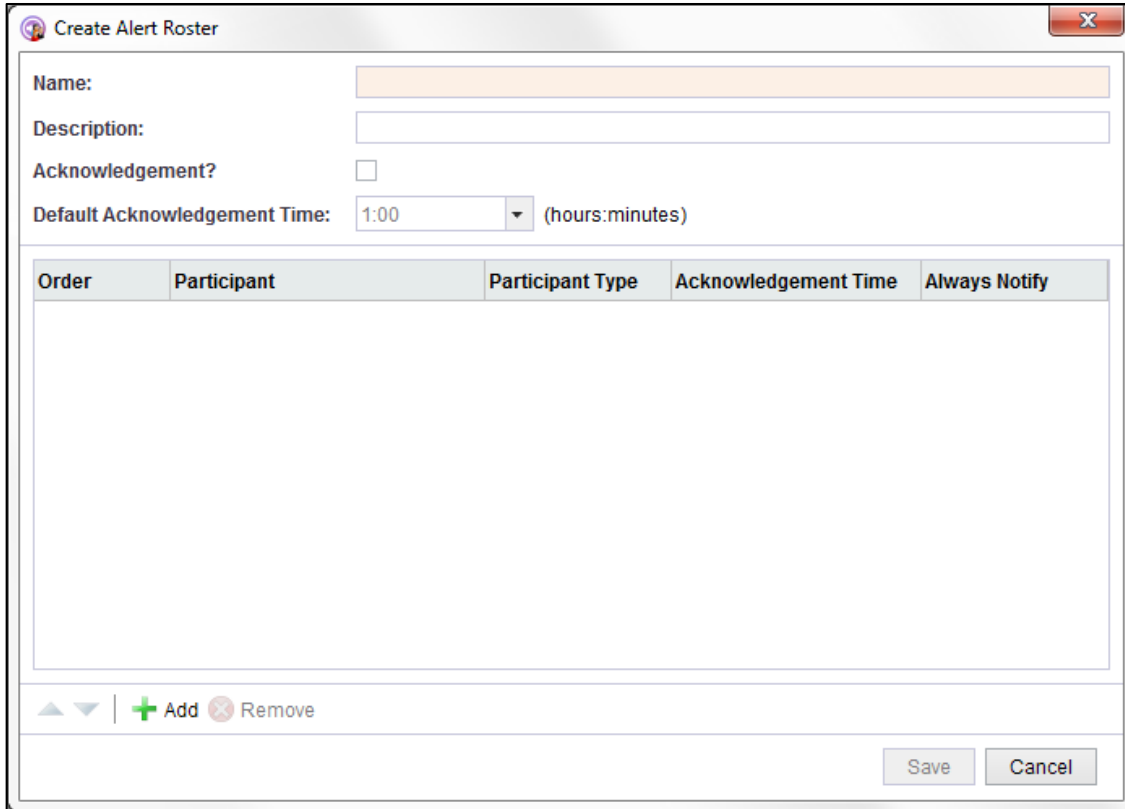
You need to take a look at the data on this table and report to the administrator.'

See Viewing Alerts section for how this Alert Template appears.

Alert Rosters

An Alert Roster determines which users will receive the Alert, the order they will receive it in, and the time period for each user to Acknowledge the Alert. Alert Rosters are located in the Settings section of the Explorer.

To create an Alert Roster select **Home > Settings > Alerts > Rosters**, right click on **Rosters** then select **New… :**

The Create Alert Roster dialog can also be opened when setting Event Handling, detailed below.

Alerts can be sent as notification only or as a notification which requires an Acknowledgement. If the **Acknowledgement?** checkbox is ticked to select, the Acknowledgment Time activates. This can be left as the default of 1 hour or changed to the length of time you wish to give users to respond. Acknowledgement time can be altered and changed for individual users if required. There is an option to return the Acknowledgment Time to default if so required.

Using the **Add** button you can add the next level in the notification process, this could be Users, Teams or Scripts. Using the dropdown menu select the Participant Type, and then the participant. Scripts can be used as a way of last resort if no user is available to respond to the Alert. Only the Administrator will be able to add scripts to an Alert Roster. To be able to select 'Script' as the Participant Type, you must have already prepared a script and saved it in Script Libraries.

A user will only be alerted once during the process, so if they appear in the list on their own and then also in a Team that is alerted later, they will be skipped over the second time. The order in which the users are alerted is determined by the order of the list, which can be changed by using the up and down arrows.

If Always Notify is ticked, a user will still receive the alert when another user responds to the alert before they are reached in the list or if they are incapable of responding to it at the time sent.

For example in the Alert Roster below:

Data Analyst will be alerted first and has 15 minutes to respond.

If they do not respond the Alert will pass to the Glossary Administrator who will have the default time of 1 hour.

After that, the team Analysts will each have 30 minutes to respond except for Data Analyst who has already been notified and will not receive a second notification.

Administrator will then have 1 hour to respond but will still receive an Alert even if the Alert is acknowledged before it reaches them.

If nobody acknowledges the Alert a script to revalidate Customer V1 will be executed. See Appendix at the end for details on Writing Alert Scripts.

Events

Alerts are triggered by specific Events. To view Events for Schemas, Tables and Columns, right click on the name of the object in the Explorer and select Events to open the drilldown for all the associated Event types and details of the set Rosters and Templates.   To edit an Event from the drilldown right-click on the relevant Event Type row and select Set Event Handling or use the right-click Edit… option in the Explorer on the Schema, Table or Column and select the Alerts tab.

Events are listed under their generator types, which are **Table**, **Column**, **Job**, **User**, **Note**, **Repository** or **System**.

By default objects will inherit the relevant events from the object above it. For example a Column will inherit all the Column and Repository events from its Table. A Table will inherit all of the Table, Column and Repository events from its Schema. Schema will inherit from the configuration of global events found in **Default Alerts** (explained later). These events can be overridden at the object's level by unticking the **Inherited** tick box which then allows you to change the configuration. Should you wish to revert back to the Schema's configuration you can tick the **Inherited** check box.
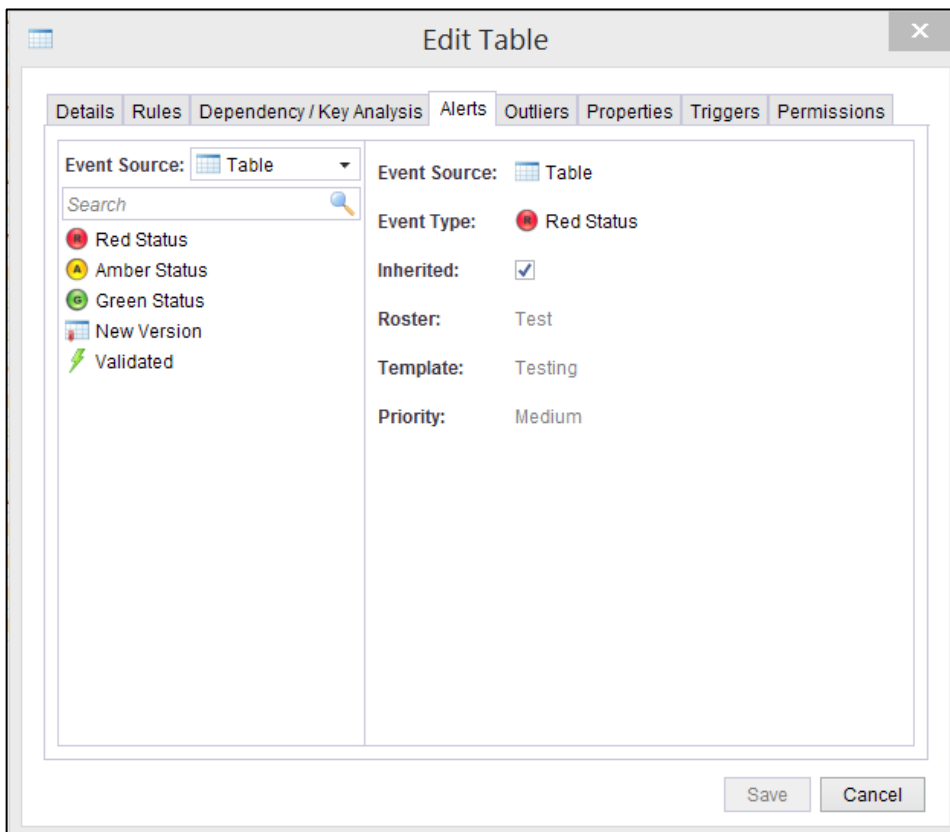
Schemas, Tables and Columns have an Alerts section in the edit dialog where you can see and adjust the list of events that can trigger Alerts. To open the dialog, go to the object in the Explorer or in a drilldown, right click on it, select **Edit…** and click on the **Alerts** tab.

The Alert Roster represents the order that users will be alerted and whether a response is required. Should you not want any alerts to be sent when the event occurs, you can select **<Disabled>** as an option for the Alert Roster. The Alert Template will be used to fill the text in the title and description of the Alert. If **<None>** is selected then a default title and description is given to the alert describing the event that has occurred and the details of the object involved. The priority is used to highlight the importance of the issue.
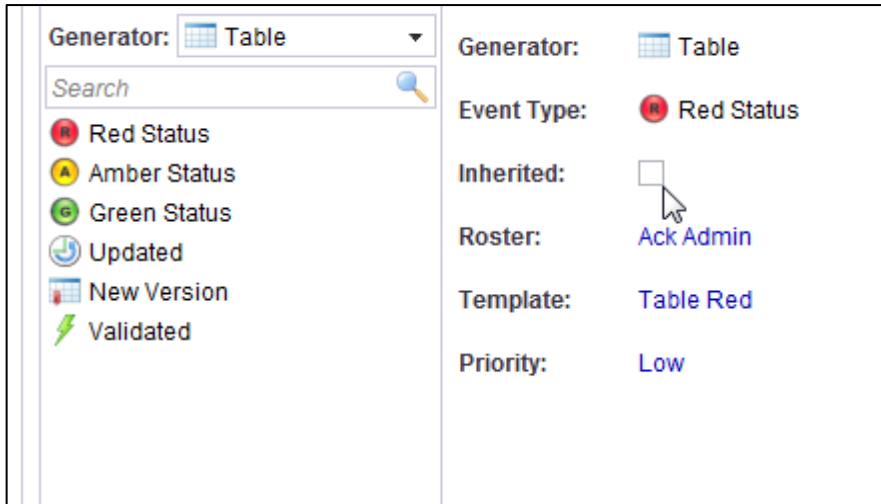
If a Table is set for events to be inherited and then that table moved into a new Schema it will inherit the Alert configuration of the new schema.

Searching Alerts

You can search for specific Alerts by right clicking on **Home** and select **Search > Search Alerts**.
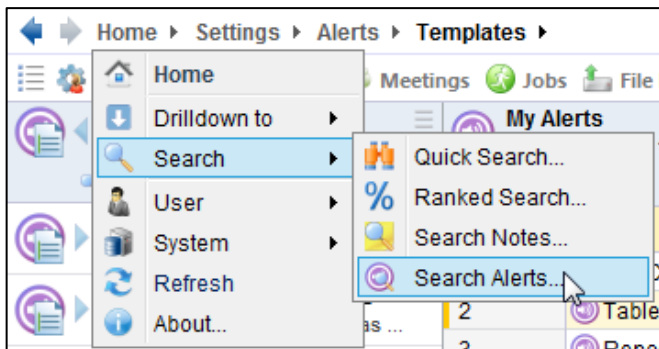
Alternatively it can be accessed by right clicking on the Alerts node in **Collaboration**:
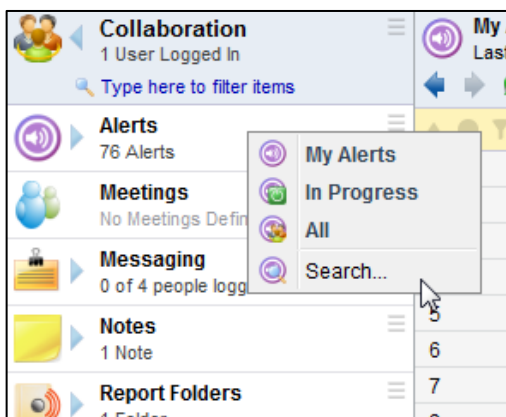
Clicking on this will bring up the following dialog:
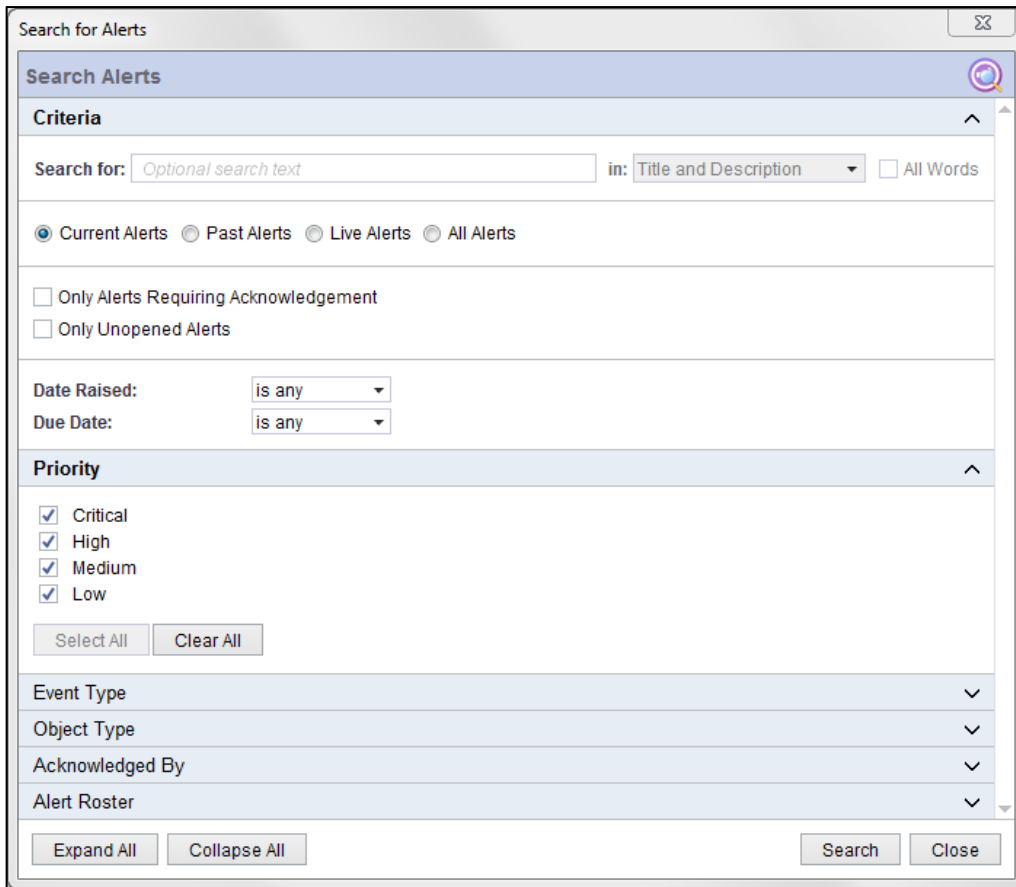
From here you can search for Alerts with particular words in the text, in the title or in the description or via other attributes attached to them.

Default Alerts

The administrator has the capability for selecting the default users to alert when events occur. This can be found in **Settings**.

The default Table, Column and Repository events are passed down to all of the Schemas in the database unless they are overridden at that level. They could then potentially be passed down to the Tables in the Schema and then the Columns of that Table.

Clicking on **Default Alerts** you can view the different categories of events that can occur. Right clicking on one of these categories allows you to change any of the events. Or you can quickly set all the events to one Alert Roster by using the **Set All Defaults** menu option.
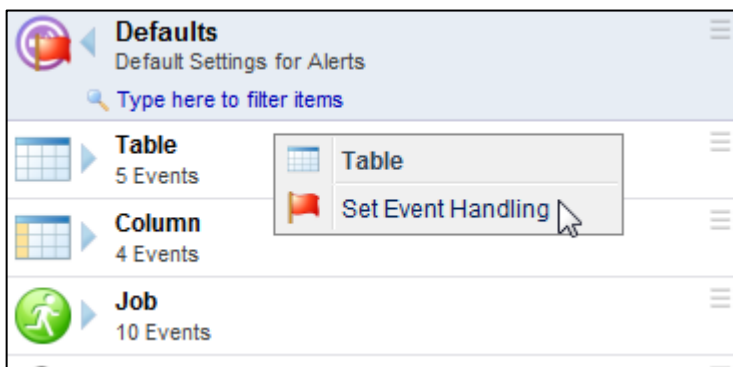
Selecting **Set Default** will bring up a dialog similar to the **Alert** section of editing a Table or Column.
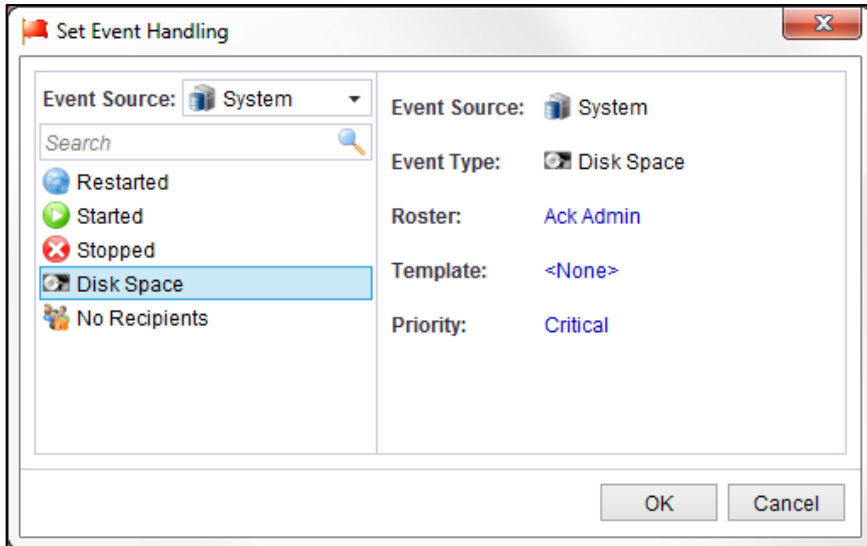
Here you can select the Alert Roster for who you want to send the Alert to, an Alert Template for how you want the Alert to appear and a priority for how important the Alert will be.

The No Recipients event will be triggered when an Alert is sent to a roster which all the users have unsubscribed from receiving Alerts.

Unsubscribing To Alerts

A user has the ability to ignore alerts that are generated from selected events. Right click **Home**, select **Users > Edit**. Then go onto the **Events** tab. Here you can untick all the events that you don't want to be alerted about. A user will only be able to see this tab if they have the Alert Management capability set in their role.
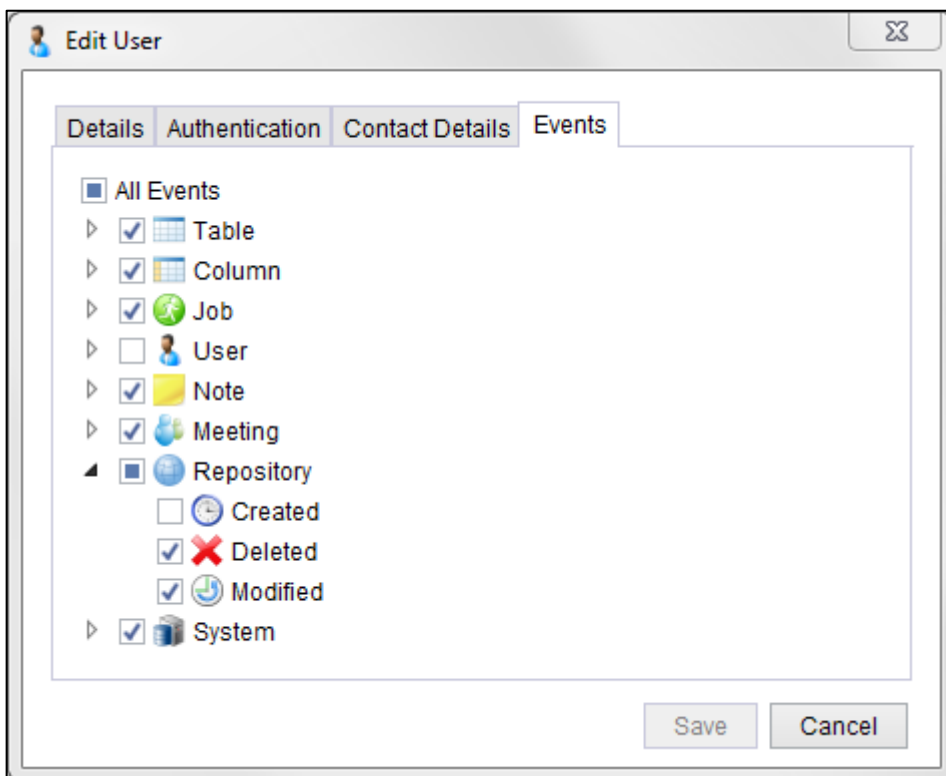
Reporting Folders

Overview

Reporting is the process of making data views available to external client applications via the DQP ODBC driver. These views, known as Reports, follow the ODBC naming convention of "Catalogue name","Schema name","Table name".

Reporting schemas are arbitrarily named folders under which you may group the Reports. A Report name must be unique within a Report Folder but can be shared among multiple Report Folders.

Before you create a Report you must first create one or more Report Folders.

Creating a Report Folder

In the explorer navigate to **Home > Collaboration > Report Folders** and right click on "**Report Folders**" and select "New" menu item.
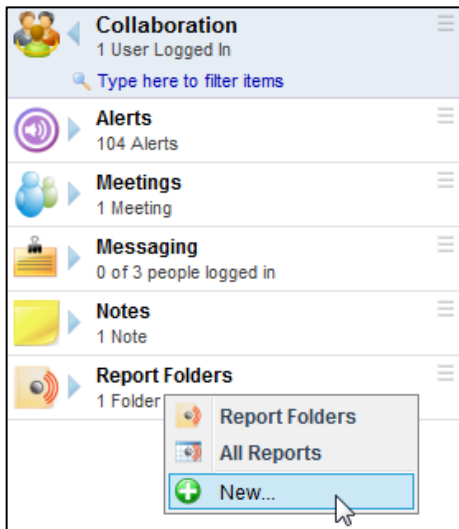
Create a name for the Report Folder and provide a description. You can also set the permissions for this Report Folder in the "Permissions" tab.
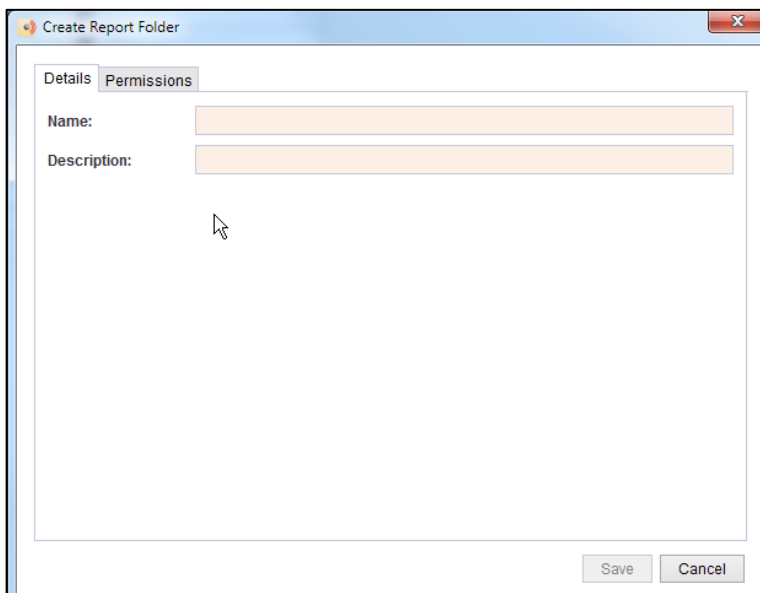
Note that since this Report Folder is visible to an ODBC enabled Client as an ODBC Schema then the name you give it must be an ODBC compliant Identifier name. In other words it should:

- Start with an Alphabetic character

- May only contain AlphaNumeric characters and Underscores
- Should not exceed 30 characters in length

Any characters that you type in the name field, other than letters, number and spaces (which will be converted to Underscores) will be ignored.
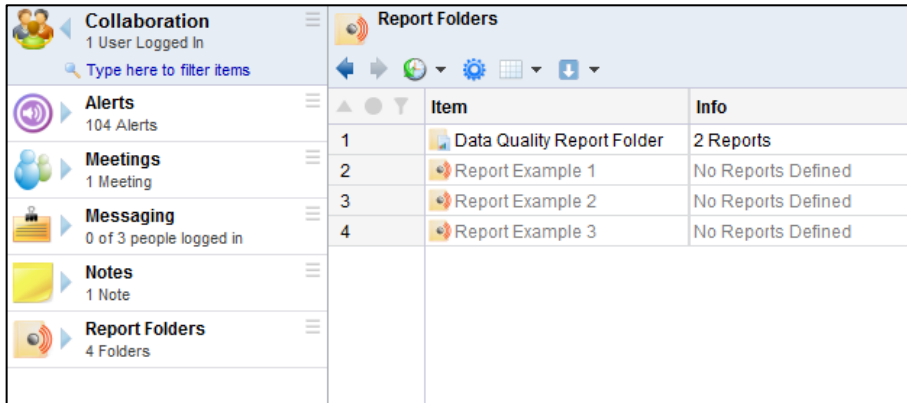
Publishing Reports

A **Report** can only be created from a saved Drilldown. Reports are published to one or more Report Folders. The visibility of the Report data is governed by the Owner of the Report Folder, so a Drilldown that has been created by the user "Administrator" and published to a Folder owned by "Administrator" may not have the same visible data as the same Drilldown published to a Folder owned by "Analyst". Using the permissions on a Report Folder you may publish the same Drilldown to different Folders but exclude confidential parts of the data depending on the owner of the Folder.

Right click on the Drilldown you wish to use to create a Report and select the folder you want to publish the report to.
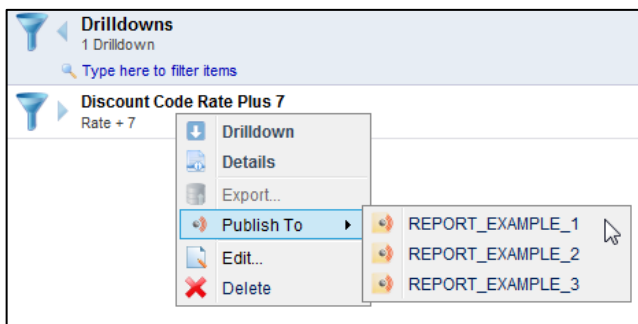
In this example we have 3 Report Folders to which we can publish the Drilldown. If this Drilldown has already been published to one or more Report Folders then the folder names will appear in the list but will be greyed out.
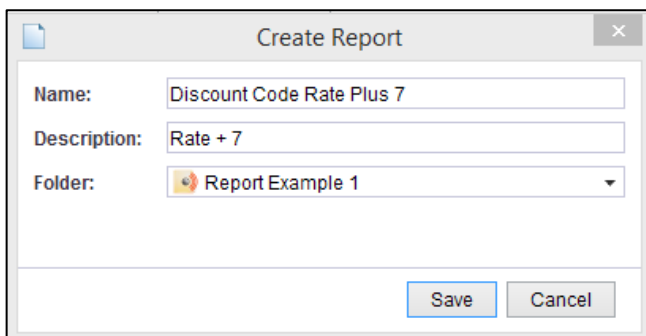
Just as the **Report Folder** ODBC Schema name is ODBC compliant, the **Report** ODBC Table name must also be ODBC compliant and is limited to 30 characters.

In this example the ODBC table name will appear as **Discount_Code_Rate_Plus_7**. When the Report has been Published it will be listed under its parent **Reporting Folder**.

A Drilldown that has been published will have an Icon indicating this.
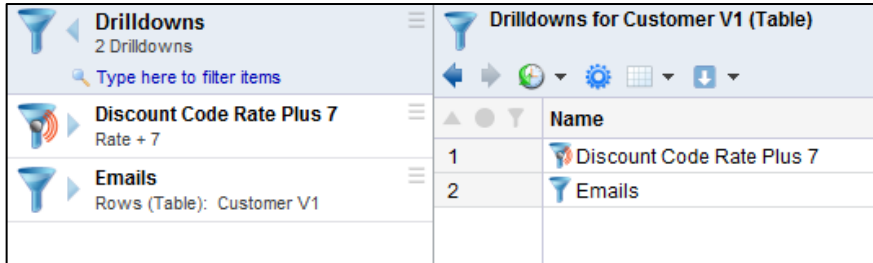
When an ODBC client connects to the ODBC DSN only the Report Folders that allow Read access to the User for that connection will be visible.

List of Reports

To view all the Reports for a Report Folder navigate to **Home > Collaboration > Report Folders > Folder_Name > Reports**
To view all the Reports for a Drilldown navigate to **Home > Data > Drilldowns**, right-click on the Drilldown name and select Details. Then double click on the row "Reports".

Deleting Publishing Reports

Reports may be unpublished either individually or collectively. Navigate to **Home > Collaboration > Report Folders > Folder_Name > Reports.** Then right click on the Report name, select Delete.
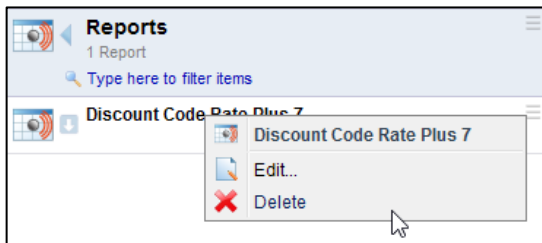
If you want to delete more than one report within the Report Folder, open the drilldown of the Report Folder, select the reports you want to delete, right click and select Delete Reports.

Reports may be collectively unpublished by Deleting their parent Drilldown. When a Drilldown is deleted all Reports that use that Drilldown will be deleted as well.

Note you only delete empty Reporting Folders.

Updating Reports

When a new version of the underlying Table is loaded and Novation has been enabled for that Table all Drilldowns associated with the table will be novated as well. Similarly any Reports associated with those Drilldowns will be novated automatically.

Profiling Columns

Introduction

Profiling of all data is performed automatically as part of the data load process in Experian Pandora along with analysis of relationships between newly-loaded columns and all existing columns in the database (unless this option has been turned off).

Experian Pandora provides 100% accuracy in data profiling since all of the data is used to derive over 180 different statistics per column. Experian Pandora calculates format patterns and uses a phonetic encoding scheme that can provide up to two phonetic patterns per value, allowing for pronunciation differences in different languages. For example, **José** can be pronounced **Ho-say**, or **Jo-si** depending on who is saying it, e.g. a Spaniard or an English-speaking person.

Tables and Columns

When data is loaded into Experian Pandora, a Table is created, with multiple Columns. Typically Tables are placed into Projects or Schemas. The entire set of Tables can be found in the explorer under **Data > Tables**. Equally the entire set of all Columns across all Tables can be found in the explorer under **Data > Columns.** Both items also have corresponding drilldowns available from their right-click menus.

Columns Drilldown

The Column Drilldown presents over 180 different statistics about columns. There are various scopes for this drilldown:

Scope	Description
Database	All Columns across all Tables in Experian Pandora
Schema	All Columns across all Tables in a given Schema
Project	All Columns across all Tables in a given Project
Table	All Columns in a given Table
Table Version Group	All Columns across all versions of a table
	All uniquely named Columns across all versions of a table
Business Term	All related Columns to the Business Term
Domain	All Columns using the Domain
User Function	All Columns using the Function
	All Columns used by the Function
Value Search	All Columns using a specific value
Format Search	All Columns using a specific format pattern

Experian Pandora calculates these statistics for all values in a Column. Additionally, it repeats these statistics (except for **Null Count**, **Actual Data type** and **Most Prevalent Data type**) for each of the datatypes discovered within a Column. For example, if a column contained mixed Alphanumeric and Integer values, you could view the column as a whole, from an Alphanumeric view point or from an Integer view point. Drilldowns are related to most of these statistics by double clicking on them. Wherever you see a Columns folder you can right click on it to view a Columns drilldown.

The Columns drilldown presents all metadata, whether documented, technical or inferred from the Data Profile, in a single table, allowing comparisons between all columns.

Menu Options

For single selection on the Columns drilldown, the menu options are almost identical to those on the explorer when right-clicking on a column, except for the additional feature to **Locate in Explorer**, which will find the selected column in the explorer for you, and the quick Filter feature found on most other drilldowns. If selecting multiple columns, an additional feature is provided to invoke the Relationship Finder to discover relationships between only the selected columns. Right-click on the multiple selection and select **Find Relationships.**

Documented Column Metadata

Every Column can have documented metadata which can either be imported from the metadata source for the original data, set at load time, editing post load or applied from another metadata source at any time. This information can be used in conjunction with built-in Validation Functions when building Validation Rules on a Column.

Metadata Item	Description
Length	The maximum length in characters of data in the Column
Datatype	The expected Datatype of values in the Column
Language	The language used for processing values in the Column. Ensures regional specifics are applied to comparison and sorting.
Minimum	The smallest expected value in the Column
Maximum	The largest expected value in the Column
Format	The expected format pattern for values in the Column
Scale	The expected numeric scale of decimal values in the Column
Precision	The expected numeric precision of decimal values in the Column
Key Threshold	The default threshold to set for detecting single column keys. Used in Key Analysis.
Nullable	True if NULLs are allowed in the Column
Value when null	The value to display instead of nothing when a NULL appears in the Column
Sort Case Sensitive	Whether case sensitivity should be ignored when sorting values in the Column
Sort with Collation	Whether collation should be used, i.e. the rules of the defined language, when sorting values in the Column

Editing Documented Column Metadata
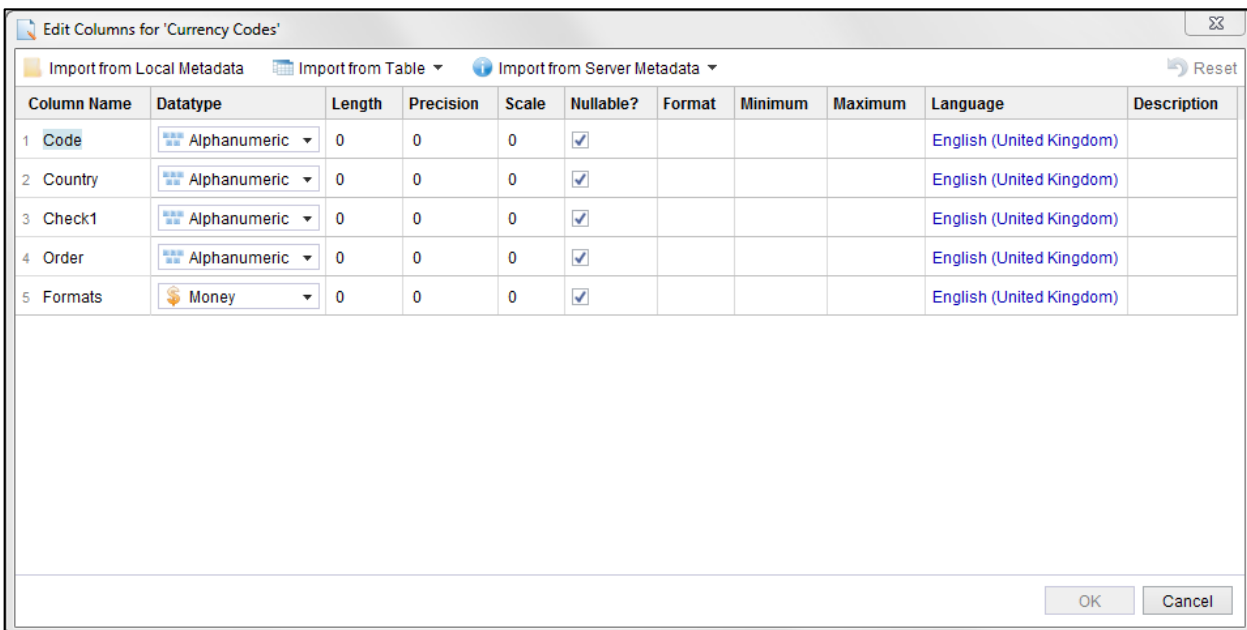
From the explorer, right-click on a Column and select **Edit**.

This displays the **Edit** dialog for the Column. Select the **Metadata** tab to edit the documented metadata items.

Click **Save** after making changes.

Editing Multiple Columns

You can also edit the metadata for all the columns of a table in a single editor. In the Explorer menu for a Table, right-click on **Columns** and select **Edit**.

The column editor shows all the metadata for all columns in a single dialog, and you can tab between all the fields. Clicking **OK** changes all the columns in a single step.

Applying Metadata from Other Sources

The columns editor also provides the ability to apply metadata from multiple sources. This means that data can be loaded without headers, the wrong headers or with missing documented metadata, and then the appropriate metadata can simply be taken from a pre-defined source, either on your local client machine, on the server or within the Experian Pandora database.

The sources available are:

Local Metadata

This is any local file of **ANSI DDL** statements taken from the creation of a relational database system, or local XML configuration files of the type produced by Experian Pandora.

Table

This is any existing Table in the database

Server Metadata

This is any file of **ANSI DDL** statements taken from the creation of a relational database system, or local XML configuration files of the type produced by Experian Pandora. These reside in the server import directories that are accessible to you which are the same ones that are shown as sources for loading data.

Column Metadata

Expand the **Info** folder for a Column to see all static and documented aspects of the Column.

Column Profile

- Right-clicking on the profile folder presents the option for the **Profile** drilldown.
- Expanding the Profile folder for a Column presents the Profile statistics for the Column accumulated during data loading.
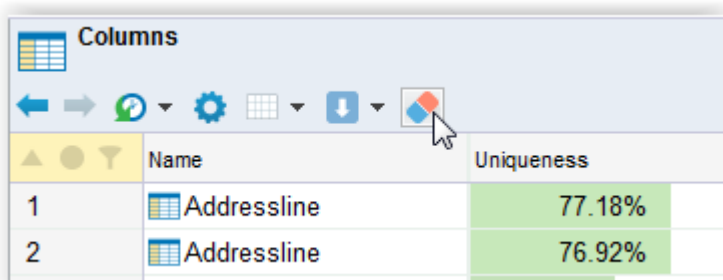
Persisting Drilldown Configuration

Note: This feature is only available in Experian Pandora v5.9.2 and later.

Users can persist changes made to reordering, hiding/showing and transforming columns when looking at the Column Profile drilldown and these changes will be kept the next time the user goes back to it. To turn this feature on, the user can go into their User Preferences options (found in Settings or by clicking the User Preferences button on the quick bar) and turn on the 'Persist Drilldown Configuration' option.

Note that going onto a saved drilldown of the Column Profile drilldown, and any changes you make to it, will not overwrite the current persisted configuration. Also note that this feature does not include the drilldown shown when right clicking on a data value and selecting 'Find Columns > Using this Value/Format'.

You can clear your configuration using the Clear Drilldown Configuration button found near the top of the drilldown.

Or if the user is unable to reach the drilldown, it can be reset by right clicking on 'Home' and selecting 'User > Clear Drilldown Configuration'.
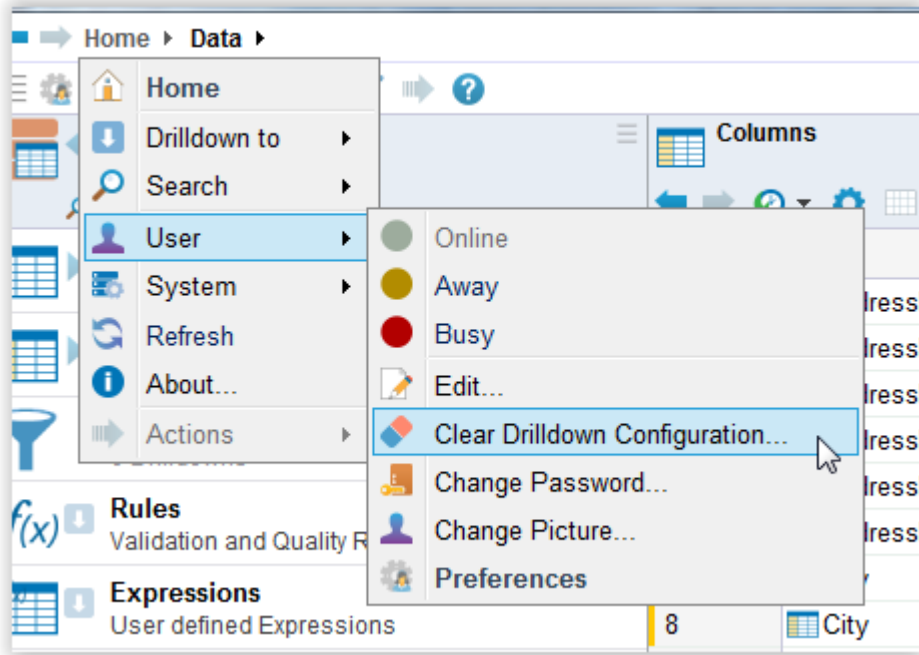
Column Profile Statistics

The statistics presented are self-documenting and self-explanatory. We will examine some of the key information.

Datatype

The overall datatype required to contain the values in the column. This is the same as the datatype required if the column was moved verbatim to a target system. The corresponding drilldown displays all rows in the Table the Column belongs to where the column has this datatype.

Dominant Datatype

The most common datatype that occurs in values in the column. The corresponding drilldown displays all rows in the Table the Column belongs to where the column has this datatype.

Nulls

The number of nulls in the column and its distribution. The corresponding drilldown will display all rows in the Table this column belongs to where the column is NULL.

Unique Values

The number of unique values in the column and its distribution. The corresponding drilldown will display the frequency analysis of all values in the column (the Values Drilldown).

Integers

The number of unique Integer values in this Column out of the total number of rows in the Table that are Integer and the percentage of rows in the Table that contain Integer values.

Alphanumerics

The number of unique Alphanumeric values in this Column out of the total number of rows in the Table that are Alphanumeric and the percentage of rows in the Table that contain Alphanumeric values

Minimum Value

The smallest value in the column and the number of times it occurs. The corresponding drilldown will display all rows in the Table the column belongs to where the value in the column is equal to the inferred Minimum Value.

Maximum Value

The largest value in the column and the number of times it occurs. The corresponding drilldown will display all rows in the Table the column belongs to where the value in the column is equal to the inferred Maximum Value.

Least Frequent Value

The least frequently occurring value in the column and the number of times it occurs. This is the first instance of a value occurring that number of times and there may be other values with the same level of occurrence. The corresponding drilldown will display all values in the column which occur the same number of times as the least frequent value.

Most Frequent Value

The most frequently occurring value in the column and the number of times it occurs. This is the first instance of a value occurring that number of times and there may be other values with the same level of occurrence. The corresponding drilldown will display all values in the column which occur the same number of times as the most frequent value.

Unique Formats

The number of distinct format patterns for values in the column. This corresponding drilldown displays the frequency analysis of the format patterns in the column.

Least Frequent Format

The least frequently occurring format pattern in the column and the number of times it occurs. This is the first instance of a format pattern occurring that number of times and there may be other values with the same level of occurrence. The corresponding drilldown displays the frequency analysis of the format patterns in the column filtered by formats that occur the same number of times as the least frequent format.

Most Frequent Format

The most frequently occurring format pattern in the column and the number of times it occurs. This is the first instance of a format pattern occurring that number of times and there may be other values with the same level of occurrence. The corresponding drilldown displays the frequency analysis of the format patterns in the column filtered by formats that occur the same number of times as the most frequent format.

Precision

The maximum precision of all Integer, Decimal or Money values in this Column. Precision is the total number of digits to the left and right of the decimal point. For example, 2.35 has a precision of 3

Shortest Length

The shortest length of any value observed in the column. The corresponding drilldown displays the frequency analysis of values in the column filtered by the shortest length.

Longest Length

The longest length of any value observed in the column. The corresponding drilldown displays the frequency analysis of values in the column filtered by the longest length.

Average Length

The average length of all values in the column. There is no corresponding drilldown.

Deviation of Length

The deviation of the length of values in the Column.

Spaces

The number of blank values in the column. This is only accounted for if the option to not strip entirely blank values is set on loading, otherwise by default varying length space values are treated as NULLS. The corresponding drilldown displays all rows in the Table the column belongs to where the value in the column is variable spaces.

Negative Values

The count of negative values in the column. The corresponding drilldown displays the frequency analysis of values in the column filtered where the value is a negative number.

Zeroes

The count of values at zero in the column. The corresponding drilldown displays all rows in the Table the Column belongs to where the column has a value of zero.

Total Sum

The sum of all Integer, Decimal and Money values in the Column

Average

The average of all Integer, Decimal and Money values in the Column

Value Deviation

The Standard Deviation of all Integer, Decimal and Money values in the Column

Checksum

Checksum of values calculated from the data.

Encoding Errors

Values that failed character set translation resulting in errors. This points to an invalid character set or corrupt data.

Frequency Distribution of Values for a Column

Right-click on a column, and select **Values** from the menu:

This displays the Frequency Distribution (or Profile of Values) for the column:

Values for Mountains.Country　　　　　　　　　　　　　　　　　　　　　　3 Rows

	Value	Rows	Distribution %	Value Length	Global Occurrences	Column Usage Count	Datatype	Format
1	England	379	16.01%	7	386	2	Alphanumeric	AAAAAAA
2	Scotland	1,703	71.95%	8	1,703	1	Alphanumeric	AAAAAAAA
3	Wales	285	12.04%	5	285	1	Alphanumeric	AAAAA

Note that values will appear in **bold** if they are different to the Original Value (i.e. different to what they looked like in the source.) This does not include leading/trailing space differences.

Heat map Capabilities

The Frequency Distribution of values for a column is heat map enabled. It will colour the **Value** column based on the number of times the value occurs in that column. This provides an instant visual indicator of how frequently a value appears in that column. Additionally, the drilldown separately highlights the heat of the **Format** column based on how often the format occurs in the column.

Rows for Customer V1 (Table)　　　　　　　　　13 of 15 Columns　　　　33,456 Rows

	Customer Id	Discount Code	Forename	Surname	Email
1	A10-E100750	M	Janice	Davies	Janice_Davies@Su
2	A10-E102508	H	Michael	Bowen	M.Bowen@Oregon_
3	A10-E102659	H	Irene	Reynolds	Irene.Reynolds@Ms
4	A10-E102982	M	Philip	Friesner	Philip.Friesner@Pa
5	A10-E103408	M	Jr	Maria	Jr.Maria@Vls_secu
6	A10-E103489	H	Adelle	Madden	Adelle.Madden@Ph
7	A10-E105402	N	Accounts	Schofield	Accounts_Schofield
8	A10-E105686	N	Paul	Anderson	Paul.Anderson@No
9	A10-E105688	M	Nigel	Horrocks	Nigel.Horrocks@Mu
10	A10-E106043	N	Reginald	Vorst	R.Vorst@Infousa.cc

Frequency Distribution of Formats for a Column

Right-click on a column, and select **Formats** from the menu:
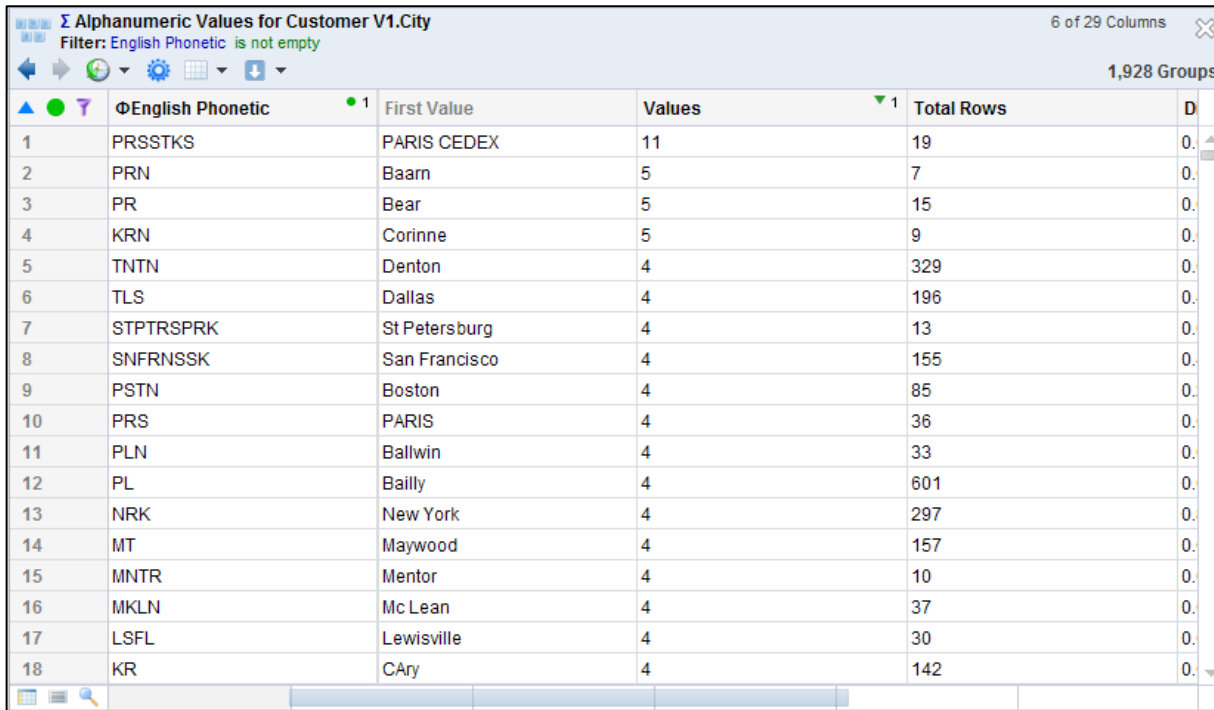
This displays the Frequency Distribution (or Profile of Formats) for the column:

82 Groups

	Format	First Value	Simple Format	Values	Total Rows	Distribution %	Format Length
1	( ) S999-999-9999	( ) 312-441-9080	( ) S9-9-9	1	1	0.01%	15
2	(9) 9.99.99.99.99	(0)1.39.30.06.93	(9) 9.9.9.9.9	2	2	0.01%	16
3	(9) 99-9999999	(0)10-4060922	(9) 9-9	4	4	0.01%	13
4	(9) 9999S9999-9	(0)7141 7856-0	(9) 9S9-9	1	1	0.01%	14
5	(9) 9999S9999S-S9	(0)2234 1821 - 0	(9) 9S9S-S9	1	1	0.01%	16
6	(9) 999S999999	(0)182 544544	(9) 9S9	3	4	0.01%	13
7	(9) 999S9999S99	(0)341 4652 21	(9) 9S9S9	1	1	0.01%	14
8	(9) 999S999S999	(0)317 455 114	(9) 9S9S9	3	3	0.01%	14
9	(9) 999S99S9999	(0)346 59 8700	(9) 9S9S9	4	4	0.01%	14
10	(9) 999S99S99S99	(0)113 34 76 93	(9) 9S9S9S9	50	51	0.15%	15
11	(9) 99S9999999	(0)26 3663641	(9) 9S9	5	5	0.01%	13
12	(9) 99S9999S999	(0)33 4220 220	(9) 9S9S9	2	2	0.01%	14
13	(9) 99S999S9999	(0)20 553 9111	(9) 9S9S9	19	20	0.06%	14
14	(9) 99S999S99S99	(0)10 211 35 72	(9) 9S9S9S9	293	306	0.92%	15
15	(9) 99S99S999S99	(0)71 71 017 00	(9) 9S9S9S9	1	1	0.01%	15
16	(9) 99S99S99S999	(0)40 27 38 648	(9) 9S9S9S9	1	1	0.01%	15
17	(9) 9S9999999	(0)1 41417777	(9) 9S9	2	2	0.01%	13
18	(9) 9S99S99S99S99	(0)1 34 88 72 98	(9) 9S9S9S9S9	56	56	0.17%	16

Frequency Distribution of Phonetic Patterns for a Column

Experian Pandora uses a phonetic encoding scheme that can provide up to two phonetic patterns per value, allowing for pronunciation differences in different languages (for example José can be pronounced **Ho-say**, or **Jo-si** depending on who is saying it (i.e. a Spaniard or an English person). You can drilldown to the first type (English Phonetic Patterns) or the second type (Non-English Phonetic Patterns). Note that these drilldowns are only available if there are phonetic patterns in the column, which is only likely if there are alphanumeric values.

Right-click on a column and select **Values,** then right-click on the row(s) you want to view the **English Phonetic** or **Non-English Phonetic** frequency for and select the relevant menu item.

	ΦEnglish Phonetic	First Value	Values	Total Rows	D
1	PRSSTKS	PARIS CEDEX	11	19	0.
2	PRN	Baarn	5	7	0.
3	PR	Bear	5	15	0.
4	KRN	Corinne	5	9	0.
5	TNTN	Denton	4	329	0.
6	TLS	Dallas	4	196	0.
7	STPTRSPRK	St Petersburg	4	13	0.
8	SNFRNSSK	San Francisco	4	155	0.
9	PSTN	Boston	4	85	0.
10	PRS	PARIS	4	36	0.
11	PLN	Ballwin	4	33	0.
12	PL	Bailly	4	601	0.
13	NRK	New York	4	297	0.
14	MT	Maywood	4	157	0.
15	MNTR	Mentor	4	10	0.
16	MKLN	Mc Lean	4	37	0.
17	LSFL	Lewisville	4	30	0.
18	KR	CAry	4	142	0.

Σ Alphanumeric Values for Customer V1.City — 6 of 29 Columns — Filter: English Phonetic is not empty — 1,928 Groups

In the above example, we have sorted the output in reverse order for the number of values matching the format pattern, and in reverse order for the length of the phonetic pattern. What we are most interested in are phonetic patterns of reasonable length (more than a few characters) that have more than one value associated.

In the example above, we can see two values of immediate interest **STPTRSPRK** and **SNFRNSSK**. The value examples of these are St Petersburg and San Francisco.

Double-clicking on the **SNFRNSSK** and we can see the values that match the phonetic pattern, here shown in a split-view:

Σ Alphanumeric Values for Customer V1.City
Filter: English Phonetic is not empty
6 of 29 Columns
1,928 Groups

	ΦEnglish Phonetic	First Value	Values	Total Rows	D
5	TNTN	Denton	4	329	0.
6	TLS	Dallas	4	196	0.
7	STPTRSPRK	St Petersburg	4	13	0.
8	SNFRNSSK	San Francisco	4	155	0.
9	PSTN	Boston	4	85	0.
10	PRS	PARIS	4	36	0.
11	PLN	Ballwin	4	33	0.

Alphanumeric Values for Customer V1.City
Filter: English Phonetic is not empty and English Phonetic is equal to 'SNFRNSSK'
26 of 27 Columns
4 Rows (0.18%)

	ΦEnglish Phonetic	First Value	Distribution %	Phonetic Length	Row
1	SNFRNSSK	San Francisco	0.454	8	152
2	SNFRNSSK	San Francisoc	0.003	8	1
3	SNFRNSSK	San Franscisco	0.003	8	1
4	SNFRNSSK	San Fransiso	0.003	8	1

The example shows 4 variants which may, in fact, be varied spellings of **San Francisco**. The longer the phonetic pattern, the higher the chance of duplicated being found. We can see much the same information looking at **STPTRSPRK**. In this case the spelling is correct, but variances in case and punctuation have resulted in differences in the values.

Σ Alphanumeric Values for Customer V1.City
Filter: English Phonetic is not empty
6 of 29 Columns
1,928 Groups

	ΦEnglish Phonetic	First Value	Values	Total Rows	D
5	TNTN	Denton	4	329	0.
6	TLS	Dallas	4	196	0.
7	STPTRSPRK	St Petersburg	4	13	0.
8	SNFRNSSK	San Francisco	4	155	0.
9	PSTN	Boston	4	85	0.
10	PRS	PARIS	4	36	0.
11	PLN	Ballwin	4	33	0.

Alphanumeric Values for Customer V1.City
Filter: English Phonetic is not empty and English Phonetic is equal to 'STPTRSPRK'
26 of 27 Columns
4 Rows (0.18%)

	ΦEnglish Phonetic	First Value	Distribution %	Phonetic Length	Row
1	STPTRSPRK	St Petersburg	0.003	9	1
2	STPTRSPRK	St. Petersburg	0.03	9	10
3	STPTRSPRK	st Petersburg	0.003	9	1
4	STPTRSPRK	st petersburg	0.003	9	1

Outliers

An outlier is a value that appears to deviate markedly from other members of the sample in which it occurs. For data, it is a value which has an unusual attribute or characteristic.

E.g. an unusually large salary amount in the HR system, a very frequent postcode in the "goods returned" system, a strangely-formatted product code in the sales system, and so on.

The Outliers function in Experian Pandora highlights values which are very different from the average (mean or expected) value in some way. The report can be configured to highlight values which are more or less "unusual".

For characteristics based on the relative sizes of a set of values, Experian Pandora uses a statistical calculation based on standard deviations and the average.

The default configuration for the Outliers report will highlight values which are at the extremes of a normal distribution curve, in other words those occurring no more than twice per thousand values.
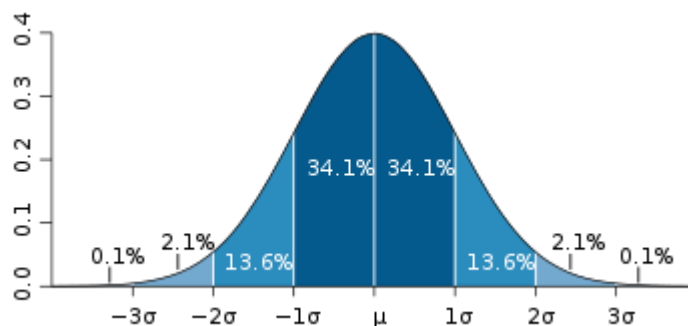
Dark blue is less than one standard deviation from the mean. For the normal distribution, this accounts for 68.27 percent of the set; while two standard deviations from the mean (medium and dark blue) account for 95.45 percent; three standard deviations (light, medium, and dark blue) account for 99.73 percent; and four standard deviations account for 99.994 percent (based on figure by Jeremy Kemp, on 2005-02-09)

For characteristics based on a single statistic, Experian Pandora uses a threshold value to determine whether something unusual is present.

Configuration

There are two main parameters that control the Outliers Report, and these are set for each table using the edit dialogue. The Outlier parameters can be set for a table.

There are also check boxes to turn on/off different types of checks at the schema, table and column, accessible by editing the object. If Experian Pandora highlights an Outlier and you decide you don't want to be informed about it subsequently, you can switch off the warning here.
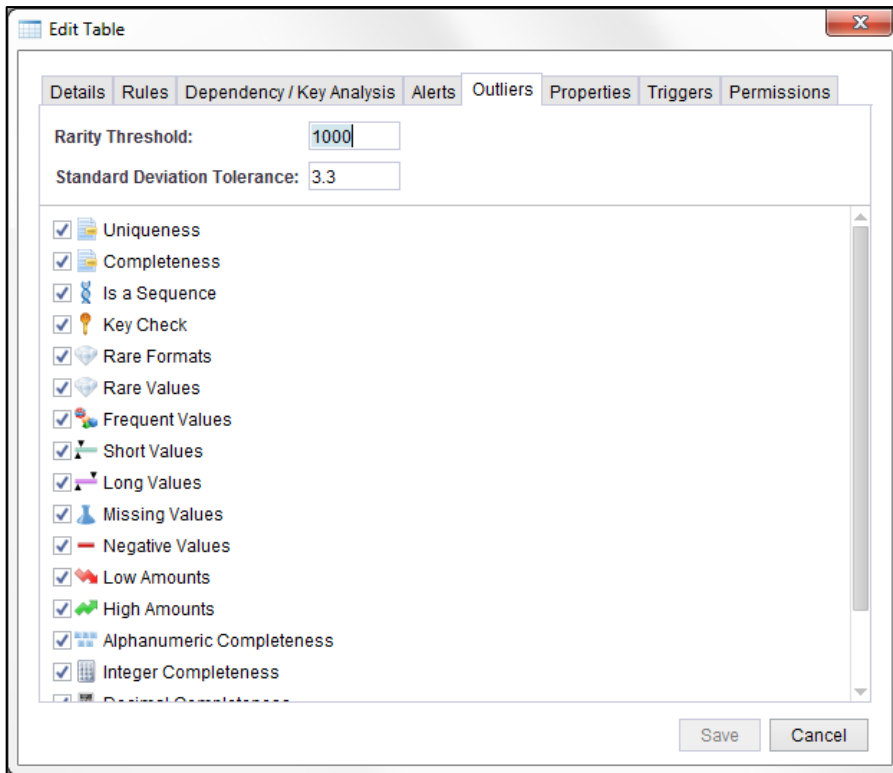
Rarity Threshold

Highlights a characteristic that should be highlighted if it occurs less frequently than one time in <threshold>. E.g. the screen shot above specifies that Experian Pandora should highlight a key as being "broken" if the number of duplicate values is less than 1 in 1000. The Rarity Threshold is used for the following Warnings:

- Is a Key
- Rare <data type>
- Unusually missing values

Standard Deviation Tolerance

Specifies the number of Standard Deviations from the norm that is tolerated. The example above will highlight records whose values are more than 3.3 Standard Deviations above or below the average value.

The Standard Deviation Tolerance is used for the following Warnings:

- High/Low amounts (numeric only)
- Frequent/Rare values
- Rare formats
- Long/short values

Not Configurable

The following warning is not configurable:
- Is a Sequence

Relevant Statistics

Uniqueness

Test for uniqueness of values in a column. The higher the percentage, the more it is unique (i.e. has less duplicates)

Completeness (by datatype and overall)

Test for how complete a column is in terms of how much of it is made up of values as opposed to Nulls. The higher the percentage, the more complete it is.

Negative Values

Returns the number of negative values in a column because negative numbers are unusual in most business applications.

Derived Outlier Warnings

Is a Sequence

Highlights any column containing sequential integer values (e.g. a generated key). If a column contains sequential integers but less than 1% of these are missing or have been duplicated, the message "broken sequence" is displayed.

Is a Key

Uses the uniqueness to detect if the column is a key. If it is 100% unique, then it is a key. If it is above the Rarity threshold then it is a broken key.

Long/Short values

Highlights values which are unusually long/short based on the Standard Deviation Tolerance.

Frequent/Rare Values (by datatype and overall)

Highlights values which occur unusually often/rarely, based on the Standard Deviation Tolerance.

Frequent /Rare Formats (by datatype and overall)

Highlights formats which occur unusually often/rarely, based on the Standard Deviation Tolerance.

High/Low Amounts

Highlights numeric values which are unusually high/low, based on the Standard Deviation Tolerance.

Table, Project, Table Group and Schema Outlier Test Flags

These control which columns are shown in the Outlier's drilldown and work in conjunction with the column flags. Each of these have their own columns drilldowns (e.g. columns in a table, columns in a schema etc.) so you can have different Outlier views at each of these levels.

Using Outliers

To view the Outliers drilldown for a table, right-click on the **Outliers** section of the Explorer, either for a Table, a Project, Schema or under the top level Data and select **Outliers**.
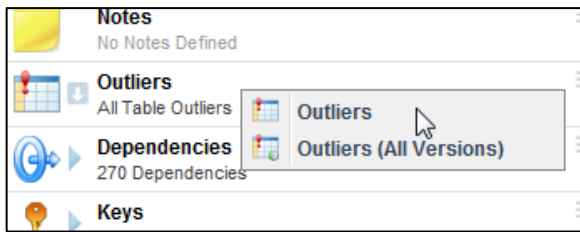
For each column, this report shows statistics such as uniqueness and completeness, then displays warnings based on the Outlier tests.

Forename	Customer	10.12%	100%				Rare		Long Values	Frequent
Good Sequence	Sales Order Detail Seq	100%	100%	Sequence	Key	Rare		Short Values	Long Values	
Grams	Annual Ground Coffee Sales	0.42%	100%							
Individual Units	Annual Ground Coffee Sales	0.08%	100%							
Order Date	Sales Order Header	0.28%	100%							
Order Id	Sales Order Detail	14.11%	100%							Frequent
Order Id	Sales Order Header	100%	100%	Broken Sequence	Key					
Order Id	Sales Order Detail Seq	14.11%	100%							Frequent
Organic	Annual Ground Coffee Sales	0.08%	100%						Rare	
Pack	Annual Ground Coffee Sales	0.27%	100%					Short Values	Long Values	Rare
Post Code	Customer	82.23%	99.43%					Short Values	Long Values	Frequent
Price	Sales Order Detail	0.02%	100%			Rare			Long Values	Rare	Frequent
Price	Sales Order Detail Seq	0.02%	100%			Rare			Long Values	Rare	Frequent
Product Description	Annual Ground Coffee Sales	2.64%	100%							

To drill down to the records, values or formats which have provoked the Warning, right click on the row of the report to view the details relating to the value.

For this example we drill down to 4 rare values (occur once only) out of the 8 possible values in this column.

Values for Customer V1.Discount Code (Column)

	Value	Rows	Distribution %	Value Length
1	H	6,647	19.87%	1
2	L	4,584	13.7%	1
3	M	5,270	15.75%	1
4	N	16,951	50.67%	1
5	A	1	0.01%	1
6	P	1	0.01%	1
7	X	1	0.01%	1
8	Z	1	0.01%	1

Profiling Tables

Viewing Rows for a Table

The entire row set for any Table can be viewed as a drilldown. Click on the table name in the Explorer, right-click on the table icon and select **Rows**.

Each row displayed in the drilldown is numbered in much the same way as a spreadsheet. This column is static and always visible on the left hand side of the drilldown.

The rows drilldown always includes two additional columns. The first is the **Row Id** column which is the row number generated internally in the Experian Pandora database. The second is the **Source Row Id** column which is the original row number from the source when the table was loaded. This may be different if a header was removed from the source, or if the source was sampled at load time. These columns are not visible by default but can be added to the drilldown using the Drilldown Manager and moving them from Hidden to Visible.

Here we can see the Row ID column is visible and we removed a couple of rows of data.   The row number has adjusted but the Row ID shows that data for rows 22-25 has been removed. The data that has been removed is shown in the Filter detail under the Drilldown detail at the top of the drilldown window.   The filter can be clicked on to open a panel which will show all the details of the filter.

Rows for Currency Codes (Table) — 6 of 7 Columns
Filter: Country is not any of 'Belarus, Rubles','Belize, Dollars','Bhutan, Ngultrum','Botswana, Pulas'
167 Rows (97.66%)

	Row Id	Code	ΦCountry
20	20	BRL	Brazil, Brazil Real
21	21	BSD	Bahamas, Dollars
22	26	CAD	Canada, Dollars
23	27	CDF	Congo/Kinshasa, Congolese Francs
24	28	CHF	Switzerland, Francs
25	29	CLP	Chile, Pesos

The rows can be manipulated, filtered, grouped and sorted in any way you desire to gain further analytical insight. In many cases, the default profiles are not sufficient as there may be redundant or noise data which is skewing the results. This can be removed through either selection filtering or creating custom filter rules prior to dynamically profiling the data.

Alternate row drilldown options

If you expand a table in the explorer and right-click on the **Rows** item, there are four row drilldown options. The first presents rows with the columns in **Positional** Order (the normal default), the second option is to view a sample from the table of 100 rows, the third to present rows with the columns in **Alphabetical** Order.   The forth, as states, will also display additional columns in the output, one for each Table Expression attached to the Table. If this is greyed out there are no Table Expressions attached to the table.

Heat map Capabilities

The Table Rows drilldown is heat map enabled. Values in columns are highlighted according to how often they occur in **that column**. This provides easy visual identification of values that occur frequently.

Table Metadata

Expand the **Info** folder of a Table in the Explorer to see all static and documented aspects of the Table.

Tables Drilldowns

You can view drilldowns of Table metadata in different contexts within the Experian Pandora database. For example **Home > Data > Tables** will show all the latest versions of the tables. To see all the versions of all the tables right-click **Tables** and select **Tables (All Versions)**

Menu Options

For single selection on the Tables drilldown, the menu options are almost identical to those on the explorer when right-clicking on a Table, except for the additional feature to **Locate in Explorer**, which will find the selected Table in the explorer for you. If selecting multiple Tables, several additional features are provided on the Tables drilldown.

Dynamically Profiling Columns from Data Rows

A column profile can be produced on-demand for any drilldown of rows being viewed, provided that drilldown has not made use of grouping or aggregation. If it has, it can still be profiled by simply saving the drilldown as a View first, and then profiling the rows of the View.

You can profile any column for Values, Formats, English/Non-English phonetics and Nulls.

Then select the type of profile you wish to view. Note for each kind of profile (except Nulls) there are 3 variants:

Variant	Description
Full Profile	A profile of all items in the column of the drilldown
Top *n*	A profile of the top n most frequent items in the drilldown
Bottom *n*	A profile of the bottom n least frequent items in the drilldown

For **Top**/**Bottom *n*** the default is *n*= 10. This can be set by changing your **User Settings** under **Server Settings**.

Right-click on any column header and from the menu and choose **Profile** to see the available dynamic profiling drilldowns.

Table Quality Report

Experian Pandora provides a quality report for every table loaded. This is available from any drilldown related to a table from the Drilldown and Reports Menu, and from right-clicking on a table name and selecting **Quality Report.**

Clicking the **Save as** button on the document viewer toolbar will show a standard file save dialog allowing you to specify the location of the export file. Table Quality reports can currently be exported as web pages (HTML) or as portable documents (PDF).

If viewing the Quality Report for a group of tables, they will appear as tabs in this drilldown. Selecting **Save All As** will export the group of Quality Reports to the desired location.

The Profiles and Summary Table can be disabled from the output. Additionally, if there are any Validation rules applied to the Table or Quality rules applied to any Columns in the table, their quality dashboards can optionally be seen in the report, both as a summary and as a dashboard per attached rule.

Charts

Chart Types

Experian Pandora has 10 different Chart Types. Each chart type has a limit to the number of rows which it will display. If a user defines a Chart with a set of data with more than the maximum available for display, the results shown will be the first set of rows up to the row limit. If you wish to use a sample of the data from the drilldown you can set the Sampling in the Drilldown Manager.

The chart types and the row limits are:

Category Charts:	Limit	Pie Charts:	Limit	XY Charts:	Limit	Other:	Limit
Bar Chart (2D)	100	Pie Chart (2D)	15	Line Chart	500	Area Chart	100
Bar Chart (3D)	100	Pie Chart (3D)	15	Scatter Chart	500	Spider Chart	100
		Ring Chart	15	Spline Chart	500		

Creating a Chart

In the Drilldown Action Menu select Chart > Add to open the dialog.   At Chart Type select the type of chart you would like to use to view your data.

Data Tab

Data Label

When selecting Bar, Pie, Area and Spider charts you have the option to specify a Data Label. Click <none> to open a drop down list and select the relevant label. For Bar and Area charts this will be the X Axis for the chart and for Pie and Spider charts this will provide the section label.

Data Series Columns

Now select the column(s) you want to view the data for from the list of Available Columns, by either dragging across or double clicking on the name. The data should be numeric - any alphanumeric data will not plot onto the Chart. There can be more than one column used in charts which support multiple results.

Chart Preview

In the top right-hand corner of the Edit Dialog, a preview of the chart will be displayed and it will alter in real time with the options you select. The chart preview can be expanded for more detail by clicking on it, and then click anywhere on it to return to the setting dialog.
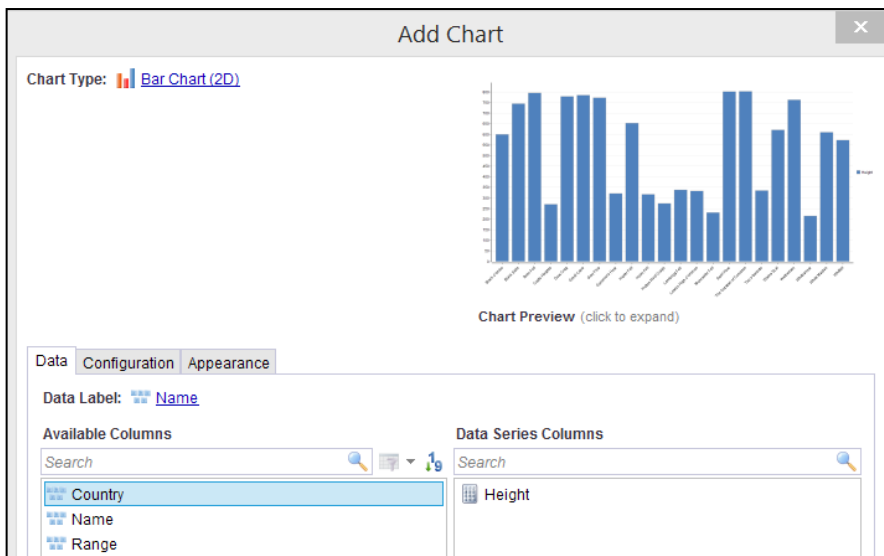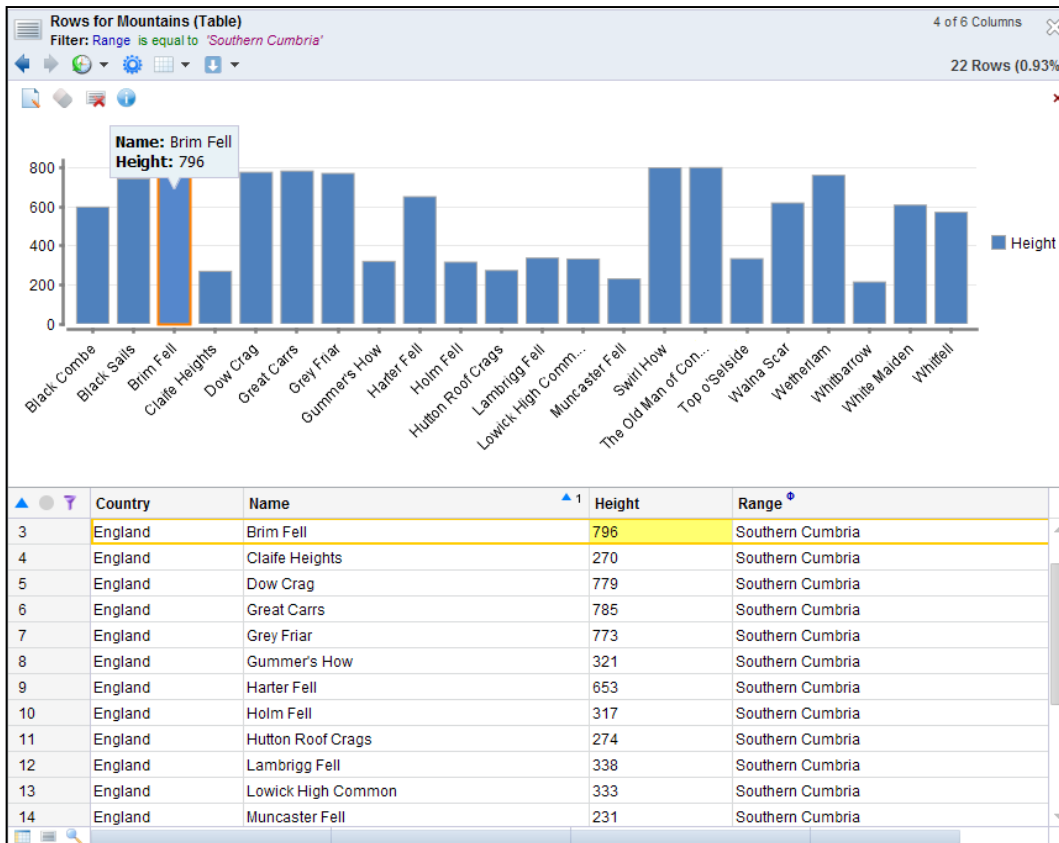
Configuration Tab

Here you can define how the Chart will be drawn and the information that will be displayed. The options available will depend upon which chart type you have chosen.

Title

Allows you to set the Title of the Chart. Text is automatically generated but can be overwritten with your specific choice of title.

X-Axis Label

Allows you to set the X-Axis of the Chart. Pie Chart and Spider Chart types don't offer this option. Text is automatically generated but can be overwritten with your specific choice of label.

Y-Axis Label

Allows you to set the Y-Axis of the Chart. Some Chart types don't offer this option. Text is automatically generated but can be with your specific choice of label.

Show Section Labels

Used only by Pie Chart types or Spider Charts, this option controls whether or not the section labels are drawn around the Chart.

Font Size

This will affect the font size of the Titles, Labels and Legend for the Chart.

Trendlines

On Bar and XY Charts selecting this option will draw a trendline for every series visible.

Logarithmic Y-Axis

When selected it will apply a logarithmic scale on the Y-Axis. This can only be applied to Bar Charts or XY Chart Types.

Logarithmic X-Axis

When selected it will apply a logarithmic scale on the X-Axis. This can be applied on XY Chart Types.

Show Horizontal Gridlines

When selected gridlines will be drawn on the Y-axis. These can be applied to XY Chart Types and Bar Charts

Show Vertical Gridlines

When selected gridlines will be drawn on the X-axis. These can be applied to XY Chart Types.

Show Shadow

When selected adds a shadow to the chart display.

Stacked

When using Bar and Area Charts you can select multiple columns in Data series' Y-Axis columns, selecting this option will stack the Y-Axis values on the top of each other.

Show Legend

This option gives you the choices as to whether or not you want the Legend displayed

Split Series

When you have multiple series' (Y-Axis Columns) selecting this option will display each series in its own Chart. To view the different results click the down arrow on the series selector, which is on the name of the displayed result below the title and select the required option. Alternatively you can use the up and down arrows on your keyboard to switch between the individual results.

NB This is not available for Pie or Ring Charts.

Web Filled

This is an option for Spider Charts, and if selected a light shading appears in the background inside the points drawn on the Chart.

Appearance Tab

Charts are initially created with a default appearance, but you can customise the way that charts are displayed get your message across more clearly, or to follow company style guides.
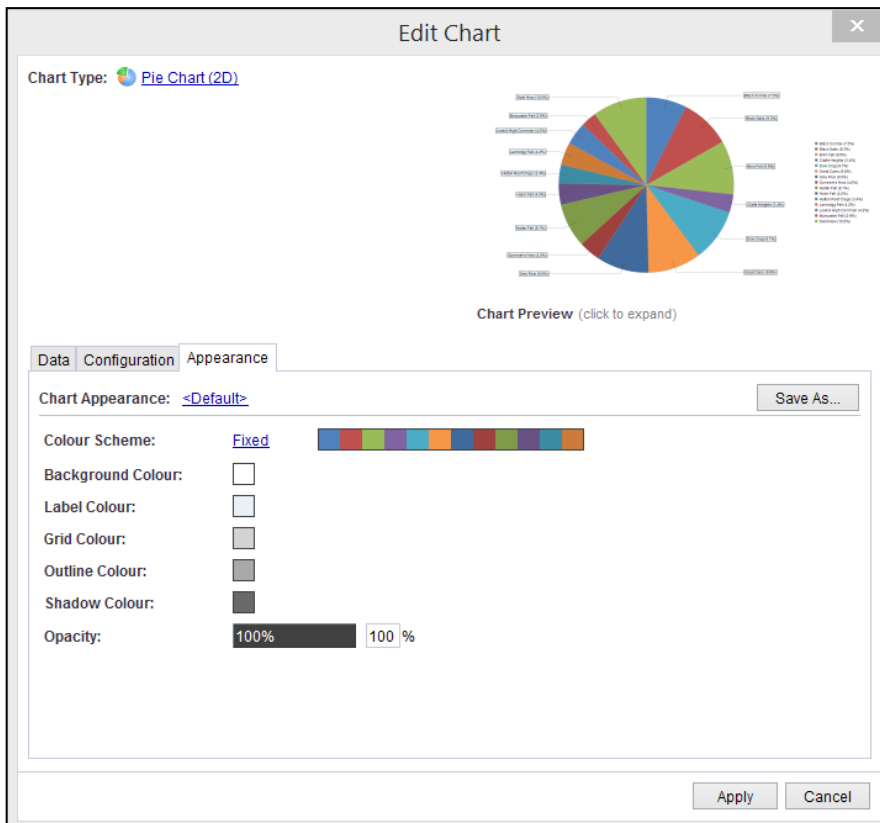
Chart Appearance

Charts are created with the chart < Default > appearance selected. As soon as you start to configure your own colour scheme, this switches to < Custom > to indicate that you have customised the appearance. If you wish to make use of the customised appearance in other charts you can save the scheme you have created.   As soon as you make a selection, all of the fields below update and the Chart Preview shows you how it will appear.

Save As…

Once you have made all the changes select Save As…   In the dialog box you must specify a name for the Chart Appearance, and can optionally give it a description. The name must be unique amongst the collection of saved Chart Appearances (see System > Repository Objects), and if you give a duplicate name Experian Pandora will verify whether you want to overwrite or not.

Colour Scheme

There are four types of colour scheme to choose from: a fixed default palette, shades of one colour getting lighter or darker, a gradient between two colours, or a custom palette of colours.

Shades

To switch between shades getting lighter or getting darker, click on the dark/light button to the right of the colour scheme preview strip.

Gradient

When you select a gradient colour scheme the colours at either end of the gradient can be changed by clicking on their swatches, and the preview strip shows how the gradient will appear. In actual use on your chart the exact appearance will depend on the number of steps required in the gradient.

Palette

When you select a Palette colour scheme, you have control over each colour individually and can adjust the size of the palette by pressing the - and + buttons.

Colour Chooser

To change the colours for a Colour Scheme, click on the colour swatch. The colour picker window pops up to allow you to change the colour. You can either click on the colour wheel and the brightness slider, or type in a specific colour. Inputs include:

RGB Colour Values

To use RGB values, type: rgb(<red>, <green>, <blue>) where <red>, <green> and <blue> are decimal integer values between 0 and 255. For example rgb(255, 0, 0) is bright red.

Hexadecimal Colours

If you want to use Hexadecimal colour codes, you can enter them as #<red><green><blue>, where <red>, <green> and <blue> are either 1 or 2 hexadecimal numerals between 0 and F or 00 and FF. For example, both #F00 and #FF0000 are bright red.

Colour Names

Finally, if the colour you want to use has a well-known name, you can type that in. Capitalisation and punctuation are ignored, so light salmon and Light-Salmon both select a pale pink colour.

The most descriptive name will always be displayed, so however you select the colour – from the colour wheel, or typing the RGB, Hexadecimal, you will see the colour name.

When you have chosen the colour that you want, the colour picker window can be dismissed by hitting enter or escape, or clicking anywhere outside of the popup window.

Background Colour

This specifies the background colour for the chart.

Label Colour

This specifies the colour that appears as the background to labels on Pie Charts, or the font colour for Bar Charts

Grid Colour

This specifies the colour used for gridlines.

Shadow Colour

This specifies the colour used for drop shadows on the charts that include them.

Outline Colour

This specifies the colour used for the outline of charts.

Opacity

This controls how opaque the chart is painted. At 100% it is totally opaque and none of the background can be seen through. At 50% the chart is semi-transparent.

Ring Width

When you have the chart type set to Ring Chart the Ring Width option is visible. This allows you to specify what proportion of the ring's radius is solid, versus the hole.

Saved Chart Appearance

Viewing Chart Appearances

You can view a Drilldown of saved Chart Appearances by navigating through the explorer to Home > Settings > Chart Appearances

Deleting Chart Appearances

To delete a Chart Appearance that is no longer required Home > Settings > Chart Appearances. Right click on the scheme you no longer require and select Delete. If the scheme is in use it will be retained in the chart but the Chart Appearance selector will now revert to <Custom>

Editing Chart Appearances

To edit a Chart Appearance, open the Edit Chart dialog and navigate to the Appearance tab. Then select one, make improvements, and save it, specifying the same name to overwrite it.

Interacting with Charts

You can interact with the Chart in a number of ways.

Working with the drilldown

You can work with the data in the drilldown, for example changing the sort order and the result is displayed immediately in the Chart. You can apply filters and change the sampling etc. in the drilldown manager.   This will open over the Chart and then when the query is run and the drilldown will change and when the drilldown manager is closed the changes are seen in the Chart.

Group on a column in the drilldown and the Chart will display the result of the first line of each group. Click on the result in the chart and the drilldown will display all the rows relating to the group.   If you wanted to view a Chart of the results for the group, close the graph, then start a new one with the results in the resultant drilldown.

Rollover Highlighting

When you roll the mouse over any result in the Chart a text label showing the information for the selection will appear and in the drilldown information below it will highlight the row and column in the data.

	Country	Name	Height	Range
23	England	Great Dun Fell	848	The Northern Pennines
24	England	Great Shunner Fell	716	The Northern Pennines
25	England	Great Stony Hill	708	The Northern Pennines
26	England	Grey Nag	656	The Northern Pennines
27	England	Grey Nag South Top	637	The Northern Pennines
28	England	Harwood Common (Redgleam)	718	The Northern Pennines
29	England	High Seat	709	The Northern Pennines
30	England	Hoove	554	The Northern Pennines
31	England	Hugh Seat	689	The Northern Pennines
32	England	James's Hill (Westernhope Moor)	675	The Northern Pennines
33	England	Killhope Law	673	The Northern Pennines
34	England	Kisdon	499	The Northern Pennines

Left-click Selection

When you click on one of the Bars, Points or Sections of a Chart, a selection occurs. A selection will act on the data below, drilling-down to the value selected. If the data came from a grouped column, it will display the grouping. Just re-click to deselect.

Right-click Menu

You can right-click on the chart to access a quick action menu. This provides options: **Edit, Clear Selection, Hide Rows** and **Remove.**

Chart Toolbar and Right-click Menu actions

The toolbar specific to Charts sits at the top left of the Chart.

Edit

Opens the edit dialog to give the User control over the current chart

Hide Data

Select this button to hide the data below the graph. Click again to return the data view.

Information

Clicking on the information button will tell you how many rows the chart is limited to.

Close

The small cross in the right-hand corner of the Chart toolbar will clear charting information from the drilldown. To get your chart back you can press the History back button on the drilldown window or you must start again.

Keyboard Arrow Keys

Up and Down

When viewing a Split Series the up and down keys are used to change the Chart Series that is currently displayed. When Split Series is disabled, the down key will clear any Selection currently on the chart.

Left and Right

The left and right arrow keys will change the current selection in the Chart, or if there is none, a left or right arrow will make a Selection.

Saving a Chart

You can export your Chart and Data as a PDF File, save it as a Drilldown, or email a link or view it as a Web Link by using the **Save As** option in the Drilldown Action Menu.

Frequency Charts on Columns

Frequency Charts are available for all columns. From either the Explorer or within a drilldown of Columns right click on the column and select from the options available.

Example: The column Discount Code can be displayed as both Frequency and Frequency Distribution, whereas City can only be a Frequency Distribution due to the number of unique values:

Values for Customer V1.Discount Code (Column)
Results Sampled & Limited to 100

8 Rows

"Discount Code" Value Frequency

	Value	Rows	Distribution %	Value Length	Global Occurrences	Colum
1	H	6,647	19.87%	1	111,318	4
2	L	4,584	13.7%	1	76,535	3
3	M	5,270	15.75%	1	87,383	4
4	N	16,951	50.67%	1	284,548	4
5	A	1	0.01%	1	14	3
6	P	1	0.01%	1	16	3
7	X	1	0.01%	1	22	3
8	Z	1	0.01%	1	10	2
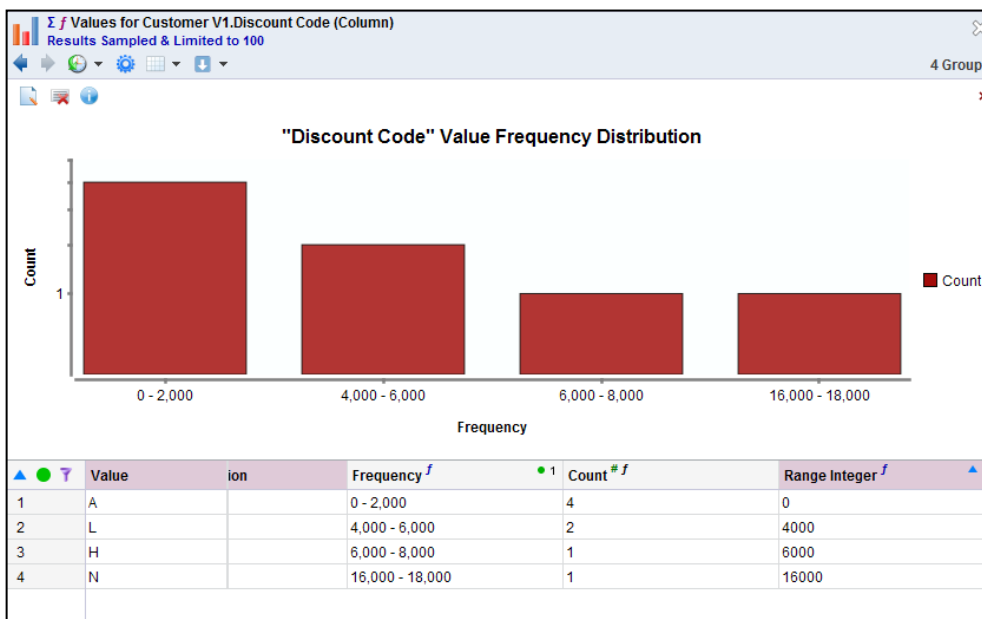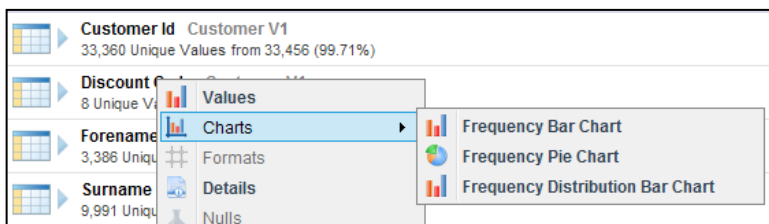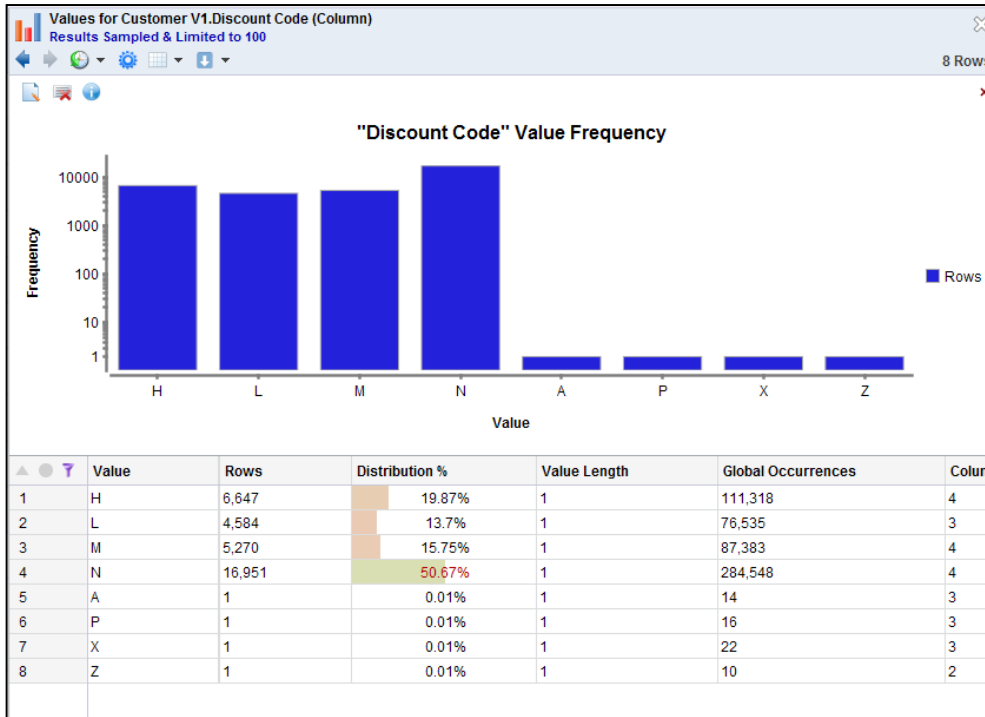
Addressline Customer V1
25,822 Unique Values from 33,456 (77.18%)

City Custome
2,269 Unique Va

Country Cus
7 Unique Values

Post Code C
27,511 Unique V

- Values
- Charts ▶
  - Frequency Bar Chart
  - Frequency Pie Chart
  - **Frequency Distribution Bar Chart**
- Formats
- Details
- Nulls
- Dashboard

Σ ƒ **Values for Customer V1.City (Column)**
Results Sampled & Limited to 100

10 Groups

"City" Value Frequency Distribution

	Value	n	Frequency ƒ	● 1	Count # ƒ	Range Integer ƒ	▲ 1
1	Black Mountain		0 - 80		2176	0	
2	Atlanta		80 - 160		27	80	
3	Birmingham		160 - 240		28	160	
4	Bath		240 - 320		15	240	
5	Carlisle		320 - 400		10	320	
6	Gloucester		400 - 480		2	400	
7	Ipswich		480 - 560		5	480	
8	Exeter		560 - 640		3	560	
9	Swansea		640 - 720		1	640	
10	Aberdeen		720 - 800		2	720	

Dependency & Key Analysis

Dependencies

Dependencies and Keys can be viewed at Global, Schema and Project levels. Experian Pandora can discover them at time of Load by changing the configuration of a table prior to loading or manually discover them as and when required.

What is a Dependency ?

A Dependency is a relationship from one or more Columns (the left-hand-side) to a single Column (the right-hand-side) within the same Table, such that the left-hand-side uniquely determines the right-hand-side. The inverse, however, may not be true.

For example, a table of Employee information may have a **NAME** Column for the employee name, and a **SSN** Column for the employee Social Security Number. There should be a Dependency between **SSN** and **NAME** as a given Social Security Number should always give back the same Name

e.g. 1234578 could be the SSN for a particular Mr Smith, but each different Mr Smith will have his own SSN.

Normalisation

Historically, Dependencies have been used in database design to decide how to break up Tables into smaller ones to avoid repeated information, a process known as Normalisation.

For example, if we had an **EMPLOYEE_ACTIVITY** table which had the clocking in and out times of **EMPLOYEES**, it would be wasteful to repeat the **SSN**, **NAME**, **ADDRESS** etc. of the employee for each activity they did. It would also make updating details of a single employee very difficult to do as it would require updating every single activity entry! We would normally design the database to have an **EMPLOYEE** table and reference the employee from the **EMPLOYEE_ACTIVITY** table to the **EMPLOYEE** table using **SSN** as a common key between them. This is known as a foreign key relationship.

Unfortunately, data in the real world is rarely of suitable quality to find 100% Dependencies. Experian Pandora discovers Dependencies with any number of Columns in the left-hand-side automatically, by simply looking at the data, with no additional user guidance. It also is able to do this very quickly, and with 100% accuracy as it uses **all** of the data to achieve this. The process is very computationally intensive, but Experian Pandora uses a unique technique leveraging the intelligent storage mechanisms in the database to achieve this much faster than any other available technology.

Experian Pandora can find Dependencies of any quality, although the minimum threshold for automatic discovery is limited to 90% since results below that are often coincidences. Experian Pandora automatically analyses and grades Dependencies in the derived Dependency, and ignores cases where both the left-hand-side and right-hand-side are NULL. For a Dependency to have an error, it will either have a problem on the left-hand-side (NULL in any of the left-hand-side Columns) or a problem in the right-hand-side (NULL or missing value, or different values on the right-hand-side) for a given left-hand-side value.

For example, in an **ADDRESS** Table you may discover a Dependency between **POSTAL_CODE** and **CITY**. For a given **POSTAL_CODE** you may expect to be given the same **CITY**. This may not hold true for a number of reasons.

Reasons for Dependency Errors

Missing Values

The **CITY** value may be missing for a given **POSTAL_CODE**. If the same **POSTAL_CODE** appears elsewhere in the **ADDRESS** Table with a value in the **CITY** Column then you may be able to use that Value to infer how to correct the missing Value.

Different Values

The **CITY** value may be different for a given **POSTAL_CODE**. This could be because the wrong information is in the wrong Column (e.g. someone put the value for a **COUNTY/STATE** in the **CITY** field).

It could also be because of a different interpretation of what should go in the **CITY** field. For example, the **POSTAL_CODE SW7** could have **Knightsbridge** in the **CITY** field, whereas other instances of that **POSTAL_CODE** have **London**.

Experian Pandora will assume that the most frequent right-hand-side value is correct (unless it is NULL), and everything else is a discrepancy error or null error. Of course only a subject matter expert can determine whether the assumption is reasonable or not.

Misspelled Values

Variances in spelling and case can break a Dependency. For example, **LONDON, London** and even **Lunden** may be intended to be the same thing. Dependency analysis will highlight these exceptions when the analyst may not have been expecting any.

Error Estimation

Experian Pandora never samples data when discovering Dependencies, and always uses 100% of the data. However, part of the speed of the Dependency analysis algorithm comes from taking an almost human viewpoint of the quality of a Dependency. In effect, it takes a quick look at *shape* of the data and judges whether a Dependency is definitely 100% or definitely has errors but is within the quality threshold given to it by the analyst. It does not need to actually count at this point all of the errors as that is wasteful from a time point of view, especially as the analysis process is likely to throw up lots of uninteresting and coincidental Dependencies.

In some cases, Experian Pandora can get a 100% answer from the shape of the data. In cases where it cannot, the results presented to the user with regard to the number of errors found will occasionally be marked as **estimated**. Drilling down to the discrepancies completes the analysis and refines the error count to make it 100% accurate. This only happens to Dependencies at or near the specified quality threshold, and Experian Pandora will never report a Dependency as being 100% accurate if, in fact, it is not.

If a Dependency has been fully validated (100% reliable) during analysis it is marked as being **Tested**. If it has not been fully validated it is marked as **Discovered** and becomes **Tested** simply by drilling down to the discrepancies.
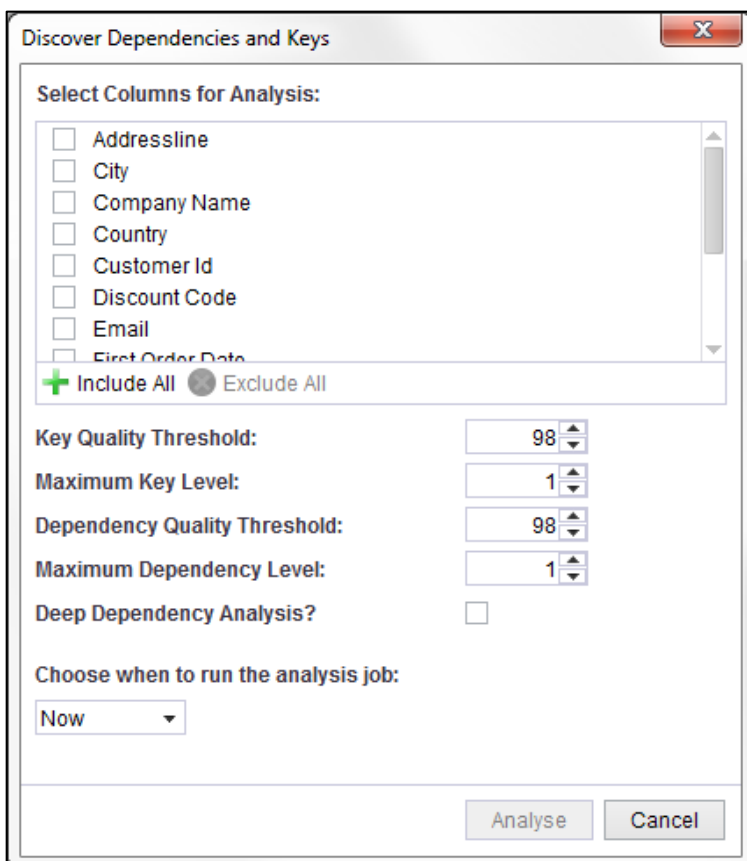
Uses for Dependencies

- Discovering unknown relationships automatically.
- Discovering expected relationships do not hold - Apart from the obvious cases where a Dependency can break down due to data errors, a Dependency may not hold for unsuspected reasons. An example is a **SONGS** Table. It may be expected that the **ARTIST** Column would always identify the same **RECORDING_COMPANY**. However this may not be true as you discover that the **Artist** may have moved between **Recording Companies** over time.

- Finding hidden data - an example of this is a **CUSTOMER** Table that had some undocumented spare Columns, intended for some long forgotten "future use". One of them may contain email addresses, for example, and the Dependency Analysis may discover that the **NAME** column identifies a spare column which actually contains email addresses that they did not know were being recorded! This suddenly gives the analyst access to **EMAIL** information they did not necessarily realise that they had available to them.

- Ensuring expected relationships hold - obvious examples are checking that an expected relationship is valid. For example, in an **ORDERS** Table which contains product information you would expect the **PRODUCT_NAME** to always give the same **PRODUCT_DESCRIPTION**, which it may not.

- Normalising Tables - Discovering left-hand-sides with 100% accuracy that point to multiple right-hand-side Columns means that a Table has the potential to be *normalised* which means it can be split into multiple tables with the left-hand-side being a common key. This allows for smaller tables and the removal of duplicated data.

Columns can be excluded from the analysis if necessary, and the analyst has control over the quality threshold and *depth* of Dependencies to be derived.

Automatically discovering Dependencies

Note that the analysis for Dependencies will also discover single or multi-column keys at the same time as it is a by-product of the same analysis process. Several of the options here are related to Key Analysis which is discussed in the next section. To launch Dependency Analysis, go to the Table in the explorer right-click **Dependencies > Discover…** The following window appears:

The window allows you to select which Columns to include in the analysis, individually ticked.

Dependency Quality Threshold

This denotes the minimum level of quality that will be accepted when discovering Dependencies.

Maximum Dependency Level

Denotes the maximum number of columns that should be considered in combination on the left-hand-side of the Dependency. The higher this is set the more analysis is required. The analysis effort increases exponentially with this setting.

Deep Dependency Analysis

This option will ensure that analysis continues to a further level even if a Dependency is discovered at a lower level. This is typically not required if the preference for Dependencies discovered is that there be less Columns in the left-hand side.

Key Quality Threshold

Denotes the lowest level of quality of keys to allow when discovering multi-column keys.

Maximum Key Level

Denotes the maximum number of Columns to consider when discovering multi-column Keys.

Scheduling Analysis

You can choose to either run the analysis now or later by clicking the **Now** dropdown and selecting a different time.

Starting the Analysis

Click on **Analyse** to start the analysis. This is queued as a job in the scheduler, the interface will automatically update when processing is complete.

In the Quick Bar at the top of your screen, the Jobs button will show a symbol indicating that you have one job running

Viewing Dependencies

To view the Dependencies for a Table, in the table in the Explorer select **Dependencies** to open the drilldown or click the icon to open the results in the Explorer.

The **Status** of the Dependency varies. *Tested* implies the Dependency has been validated, either by the user drilling down onto it to see the detail for the Dependency, or because it is a 100% Dependency and can therefore be deemed automatically tested.

Drilling down on a Dependency displays the conflicts for the Dependency, i.e. the values that break the Dependency rule.

	Forename 1	Gender	Rows
1	Bradley	female	1
2	Bradley	male	201
3	Charlie	female	176
4	Charlie	male	172
5	Daniel	Both	1
6	Daniel	male	164
7	Jake	Both	1
8	Jake	male	188
9	Jordan	female	1
10	Jordan	male	163

Here we can see for the Dependency **Forename 1** *identifies* **Gender** that there are a number of cases where the forename has different genders. This may have been due to a data quality error or, as is the case here, forenames that can legitimately be applied to both genders. An analyst may have assumed that they could identify the gender from the forename, and this would prove that clearly, that is not the case.

We have highlighted both cases of **Bradley** in the drilldown above, and can right-click and chose **View Conflicts** to see the rows

	Application Reference	Title	ΦForename 1	Forename 2	Forename 3	Surname	ΦGender
1	700000000041243	Dr	Bradley	Jay	Reece	Knowles	female
2	700000000001064	Mr	Bradley			Patterson	male
3	700000000001481	Dr	Bradley			Hawkins	male
4	700000000001528	Mr	Bradley	Thomas		Warren	male
5	700000000001609	Dr	Bradley	Christopher		Moss	male
6	700000000001694	Mr	Bradley	Hayden		Pritchard	male
7	700000000002115	Dr	Bradley	Toby		Little	male
8	700000000002285	Dr	Bradley	Ewan		Barnes	male
9	700000000002866	Dr	Bradley	Alex		Williams	male

Actually, on further inspection it transpires that Dr Bradley Knowles is infact a man and we have spotted a gender data issue. Note that when we first looked at the Dependencies drilldown, **Forename 1** to **Gender** was marked *Discovered*. If we return to the Dependencies drilldown again, you will see that it is now marked as *Tested* as simply drilling down on the Dependencies of interest is enough to complete the statistical results.

	Name	Identity Columns	Identified Column	Quality	Tested At	Status	Rows To Cleanse
1	Application Reference ▸ Title	Application Reference	Title	100%	98%	Tested	0
2	Application Reference ▸ Forename 1	Application Reference	Forename 1	100%	98%	Tested	0
3	Application Reference ▸ Surname	Application Reference	Surname	100%	98%	Tested	0
4	Application Reference ▸ Gender	Application Reference	Gender	100%	98%	Tested	0
5	Application Reference ▸ Date Of Birth	Application Reference	Date Of Birth	100%	98%	Tested	0
6	Forename 1 ▸ Gender	Forename 1	Gender	99%	98%	Tested	344

As you can see the original quality estimate by Experian Pandora was actually completely accurate and remains at 99%.

Discovering data errors with Dependencies

Here we see another example with Purchase Order Detail. There are some discrepencies related to the price of certain products, visible by the Product Id to Price rule being only 99% in quality.

	Name	Identity Columns	Identified Column	Quality	Tested At	Status	Rows To Cleanse
1	Order Id ▸ Discount Code	Order Id	Discount Code	100%	98%	Tested	0
2	Product Id ▸ Price	Product Id	Price	99%	98%	Discovered	1

Looking at product id **KT412** we can see that 536 rows are priced at £5.00 and only one row is priced at £6.00.

	Product Id	Price	Rows
1	KT412	GBP 5	536
2	KT412	GBP 6	1

Clearly the one purchase was either incorrectly priced or had been subject to a change in price. Double click the conflict with the £6.00 price and note the date that it was shipped.

Drill down to all the rows by highlighting both conflict results, right click and select **View Conflicts**, we can carry on our investigation to discover that the product had been shipped several times on the same date with the lower price so this is looking more likely that it was incorrectly priced at £6.00

	Order Id	ΦProduct Id	ΦPrice	Quantity	Discount Code	Shipping Date
11	1079411	KT412	GBP 5	55	N	06-Jul-2009
12	1076987	KT412	GBP 5	280	L	07-Jul-2009
13	1077025	KT412	GBP 5	66	N	07-Jul-2009
14	1077524	KT412	GBP 5	8	N	07-Jul-2009
15	1078202	KT412	GBP 5	77	H	07-Jul-2009
16	1079383	KT412	GBP 5	110	M	07-Jul-2009
17	1080647	KT412	GBP 6	95	N	07-Jul-2009
18	1084055	KT412	GBP 5	3600	M	07-Jul-2009
19	1084377	KT412	GBP 5	390	N	07-Jul-2009
20	1078649	KT412	GBP 5	3200	H	08-Jul-2009
21	1079563	KT412	GBP 5	300	N	08-Jul-2009

This is obviously not conclusive and the next step is to actually see when the item was **ordered**. This can be done by now joining Purchase Order Detail to Purchase Order Header, and filtering out for that Shipping Date. (See Relationships section for details of how to join tables together.)

With a few mouse clicks, I have drilled down to the common rows between Purchase Order Header and Purchase Order Detail, filtered out for shipping dates of the 5th July 2009 and for the product number KT412.

I can see straight away that the item was ordered on the 4th of July, and several rows up there is another instance of an order on the same date of the same product, at the more common £5.00 price. I now have very strong evidence that the price of £6.00 is wrong and a data quality error.

Relationship for Purchase Order Detail to Purchase Order Header on Order Id
Filter: Product Id is equal to *'KT412'* and Shipping Date is equal to *'07-Jul-2009'* - Common Rows

9 of 11 Columns

8 Rows (0.01%)

	Order Id	ΦProduct Id	Price	Quantity	Discount Code	ΦShipping Date	Order Id 1	Customer Id	Order Date
1	1076987	KT412	GBP 5	280	L	07-Jul-2009	1076987	E117104	06-Jul-2009
2	1077025	KT412	GBP 5	66	N	07-Jul-2009	1077025	E149630	06-Jul-2009
3	1077524	KT412	GBP 5	8	N	07-Jul-2009	1077524	E121845	04-Jul-2009
4	1078202	KT412	GBP 5	77	H	07-Jul-2009	1078202	E137990	12-Jul-2009
5	1079383	KT412	GBP 5	110	M	07-Jul-2009	1079383	E128311	05-Jul-2009
6	1080647	KT412	GBP 6	95	N	07-Jul-2009	1080647	E115638	04-Jul-2009
7	1084055	KT412	GBP 5	3600	M	07-Jul-2009	1084055	E108183	05-Jul-2009
8	1084377	KT412	GBP 5	390	N	07-Jul-2009	1084377	E121479	15-Jul-2009

Manually testing expected Dependencies

It is not necessary to run the analysis process to test an expected Dependency. Here we show creating a Dependency to test without the need for running the analyser. Expand a Table in the explorer and right click **Dependencies** and select **New**

The **Create Dependency** dialog appears. We are going to test our belief that a **City** should uniquely identify a **Post Code**

Edit Dependency

Name: Postal Code To City

Description:

Select the Identity Columns (1 column selected)

- [ ] Customer Id
- [ ] Discount Code
- [ ] Forename
- [ ] Surname
- [ ] Email
- [ ] Telephone
- [ ] Company Name
- [ ] First Order Date
- [ ] Sales Last Year
- [ ] Addressline
- [ ] City
- [ ] Country
- [x] Post Code

Include
Exclude
Include All
Exclude All

Identified Column: City

Save   Cancel

Firstly, we set the **Identified Column** to be **City**. Next, we choose our left-hand-sides. In this case, we select only **Post Code**. Finally, we must give our Dependency a name. Click **Save** to create the Dependency.

This **User Defined** Dependency is added to the list of Dependencies. The quality and statistics are totally unknown at this point because it has not been analysed or tested in anyway. Simply double-click on the Dependency (or right-click and select **View > Dependency Conflicts** to see the results of the Dependency. I will first split the screen.

Simply by drilling down, and by virtue of the split pane, I can see the Dependency has updated in real-time and in-fact is at a 98% quality. There are a number of conflicts visible, for example the first two show a conflict due to case differences in **Amsterdam**.

Additive Analysis

Dependency Analysis is additive. Given the performance of analysis, it is often better when you are discovering Dependencies you are unaware of to start with a high quality level. If you then drop the quality on a subsequent analysis, Experian Pandora will not overwrite any Dependencies that already exist from previous analysis runs. This allows you to incrementally analyse without having to wade through lots of noise that can occur with coincidental Dependencies due to a low quality threshold.

Deleting Dependencies

To remove or delete Dependencies, right-click on the one you wish to remove, select **Delete.**

Key Analysis

At load time, each column is analysed to discover how unique its values are. A column whose values are totally unique is obviously a key. However, it is often the case that the records in a table are identified by a key **composed of values from more than one column**. An obvious example would be the line item records associated with a customer order. The line item key could be composed of an order number and a line number. It could even be that there is more than one line item with this same key, and that there's a third column to be taken into account, say a date or historical sequence.

While carrying our Dependency analysis, Experian Pandora also **derives the potential multi-column Keys** for each table based on the same columns as have been selected for Dependencies. The analyst must specify the maximum number of columns to be considered and the quality threshold below which candidate Keys can be ignored. Via drilldown the analyst can see the values in conflict with the derived key, and the reason for the conflict.
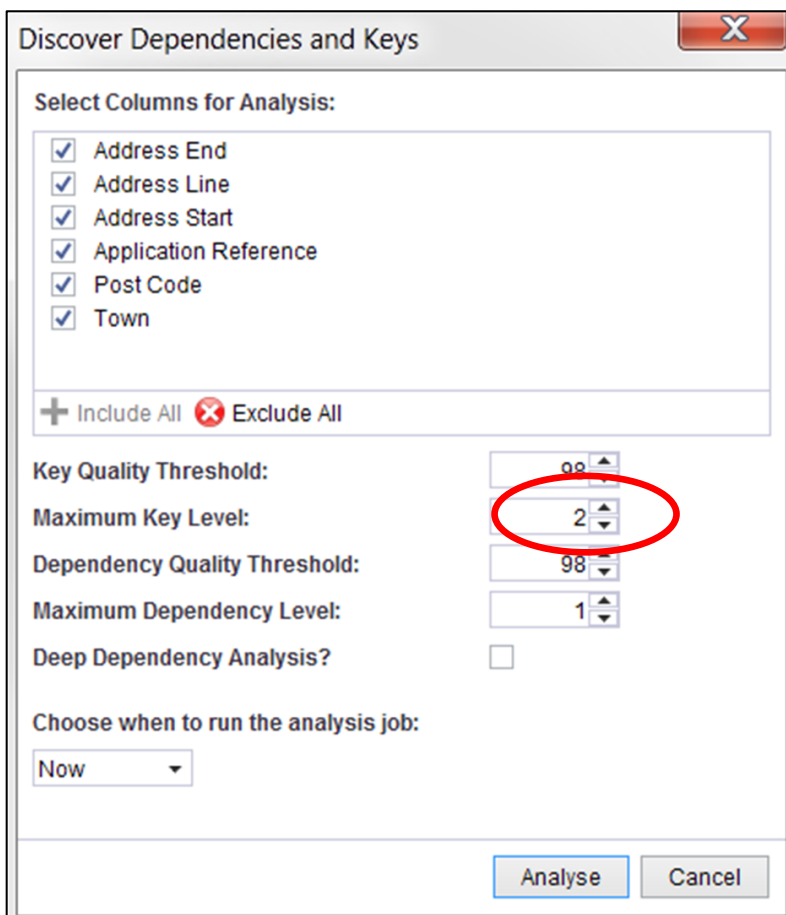
Further drilldown can be performed to show the actual rows of data in conflict.

It is also possible to specify a composite key and Experian Pandora will simply validate it.

Keys and Dependencies can be defined (and validated) and also *deleted* from the list of derived Keys if they are of no interest.

Automatically discovering multi-column Keys

This is the same process as discovering Dependencies, covered in **Automatically discovering** . Simply set the number of columns required in the key analysis (in this case to 2) and the required key quality threshold.

Here we see the results of Key analysis for the **Applicationsaddresses** Table, using 2 columns on the left-hand-side. Go to the Table in the explorer and select **Keys** to see the Keys for a Table.

Keys for Applicationaddresses (Table) 9 of 19 Columns

7 Rows

	Business Name	Primary Key	Key Columns	Quality	Tested At	Status
1	Application Reference, Address Start	✖ No	Application Reference, Address Start	99%	98%	Tested
2	Application Reference, Address Line	✖ No	Application Reference, Address Line	99%	98%	Tested
3	Application Reference, Town	✖ No	Application Reference, Town	99%	98%	Tested
4	Application Reference, Post Code	✖ No	Application Reference, Post Code	99%	98%	Tested
5	Address Start, Address Line	✖ No	Address Start, Address Line	99%	98%	Tested
6	Address Start, Town	✖ No	Address Start, Town	99%	98%	Tested
7	Address Start, Post Code	✖ No	Address Start, Post Code	99%	98%	Tested

Drilling down to the key composed of **Application Reference** and **Address Line**, we can see the conflicts for the key. In this example we see that we have several key values that result in duplicates.

Key Conflicts for Application Reference, Address Line (Key) 4 of 5 Columns
Key 683

4 Rows

	Application Reference	Address Line	Rows	Conflict Description
1			1	Null Value
2	700000000004752	177 Park Avenue	2	Duplicates
3	700000000023114	193 St Andrews Lane	2	Duplicates
4	700000000031165	11 Scotsburn Rd	2	Duplicates

We can further drilldown to see any or all of the rows where the Keys are in conflict:

Rows for Applicationaddresses (Table) 6 of 8 Columns
Filtered by 3 Conflicts

6 Rows (0.01%)

	ΦApplication Reference	Address Start	Address End	ΦAddress Line	Town
1	700000000004752	18-Jul-1995	04-Oct-2003	177 Park Avenue	LAUDER
2	700000000004752	05-Oct-2003	05-May-2006	177 Park Avenue	LAUDER
3	700000000023114	24-Mar-2003	06-Mar-2004	193 St Andrews Lane	CWMFFRWD
4	700000000023114	07-Mar-2004	07-Sep-2006	193 St Andrews Lane	CYMMER
5	700000000031165	22-Feb-2004	31-Mar-2005	11 Scotsburn Rd	TAMWORTH
6	700000000031165	19-Jun-2006		11 Scotsburn Rd	TACOLNESTON

Experian Pandora actually constructed an automatic filter on the rows for the Table in the above drilldown which can be further edited for more detailed analysis.

Filtering   Reset   Run Query

Pre-filter | Post-filter | Sampling

Add

▲ ✓ 1   ✕
  + ✓ Application Reference is equal to '700000000004752'   ✕
& + ✓ Address Line is equal to '177 Park Avenue' Case Sensitive   ✕

▲ ✓ 2   ✕
  + ✓ Application Reference is equal to '700000000023114'   ✕
& + ✓ Address Line is equal to '193 St Andrews Lane' Case Sensitive   ✕

▲ ✓ 3   ✕
  + ✓ Application Reference is equal to '700000000031165'   ✕
& + ✓ Address Line is equal to '11 Scotsburn Rd' Case Sensitive   ✕

	ΦApplication Reference	Address Start	Address End	ΦAddress Line	Town
1	700000000004752	18-Jul-1995	04-Oct-2003	177 Park Avenue	LAUDER
2	700000000004752	05-Oct-2003	05-May-2006	177 Park Avenue	LAUDER
3	700000000023114	24-Mar-2003	06-Mar-2004	193 St Andrews Lane	CWMFFRWD
4	700000000023114	07-Mar-2004	07-Sep-2006	193 St Andrews Lane	CYMMER
5	700000000031165	22-Feb-2004	31-Mar-2005	11 Scotsburn Rd	TAMWORTH

Manually creating a key

Expand the Table in the explorer and select **Keys** right-click and select **New** from the menu:

Create Key

Name:

Description:

Primary Key? ☐

Select the Key's Columns (no columns selected)

☐ Application Reference
☐ Address Start
☐ Address End
☐ Address Line
☐ Town
☐ Post Code

Include
Exclude
Include All
Exclude All

Save   Cancel

Simply choose the columns that form the multi-column key that you wish to test for.

Global Value & Format Analysis

Experian Pandora provides a unique ability to search across all Columns within all Tables of data. That is searching **every** value in **every** column across **every** table. This capability is due to the novel way that Experian Pandora stores data, as it only ever stores a value once regardless of how many times it appears in columns and tables loaded.

Values by Column Drilldown

This is a special drilldown that allows side-by-side comparison of **all** columns in a table. This drilldown is very unusual as it presents a pair of columns in the drilldown for every column in the table. The first is the value in that column, and the second is the frequency of that value in the column. The column pairs are ordered so that the left-most is related to the **column with the most unique values** in the table. This proceeds to the right such that the rightmost column is the one with the **lowest number of unique values** in the table. Inherently, these are variable length columns, which is unusual in a tabular drilldown!

To view the drilldown, expand a table in the explorer, then right click on the **Columns** folder and select **Multi-Column Value Analysis**. The following drilldown appears.

	Name	Count = 2,219	Height	Count = 871	Range	Count = 96	Country	Count = 3
1	A'Bheinn Bhan	1	153	1	Aberystwyth to Welshpool	15	England	379
2	A'Bhuidheanach Bheag	1	154	1	Altnaharra to Dornoch	14	Scotland	1,703
3	A'Bhuidheanach Bheag - Glas Mheall Mor	1	155	1	Anglesey and the Lleyn Peninsula	10	Wales	285
4	A'Bhuidheanach Bheag - Meall a'Chaorainn	1	159	1	Applecross to Achnasheen	21		
5	A'Chailleach	2	160	1	Arran and Holy Island	11		
6	A'Chailleach - Toman Coinich	1	164	1	Ayr to the River Clyde	9		
7	A'Chralaig	1	167	1	Bala to Welshpool	42		
8	A'Chralaig - A'Chioch	1	168	1	Barmouth to Betws-y-Coed and Bala	27		
9	A'Chralaig - Stob Coire na Cralaig	1	169	1	Barra to Barra Head	9		
10	A'Chruach	2	170	1	Braemar to Montrose	55		
11	A'Ghlas-bheinn	1	171	1	Canna, Rum and Eigg	8		
12	A'Mhaighdean	1	172	2	Cape Wrath Peninsula	8		
13	A'Mharconaich	1	173	1	Carmarthen to the Vale of Neath	9		
14	Abberley Hill	1	178	2	Carrick and Galloway	43		
15	Aberedw Hill	1	179	1	Central Scotland from Dumbarton to Montrose	36		
16	Aconbury Hill	1	180	1	Central and Eastern England	6		
17	Ailsa Craig	1	181	2	Central and Western Fells	100		
18	Ainshval	1	182	1	Cornwall and Devon	8		
19	Airds Hill	1	185	1	Dolgellau to Machynlleth	17		
20	Alhang	1	187	1	Durness to Loch Shin	23		
21	Allen Crags	1	188	1	Eastern Fells	56		
22	Allen Crags North Top	1	189	1	Firth of Forth to the River Tweed	28		
23	Allermuir Hill	1	190	3	Fort William to Loch Treig and Loch Leven	59		
24	Allt Fawr	1	191	1	Fraserburgh to the Dee Valley	13		

You can clearly see how **Country** which only has 3 unique values is a variable length column. Each pair of columns relates to a single column in the table. The heading of column containing the value is the same as the name of the physical column in the table, and the column to the right of it is the number of times the value occurs in that column. The heading of the count column has the overall unique count of values in the column, and is marked **red** if the column also contains nulls.

	Post Code	Count = 27,511 (Nullable)	Addressline	Count = 25,822 (Nullable)	First Order Date	Count = 13,500 (Nullable)
1		192		1		96
2	0	5	1	2	01-Jan-1969	1
3	11	1	119	1	02-Jan-1969	1
4	130	1	122	1	03-Jan-1969	2

This is a very fast process as there is no manipulation required on the part of Experian Pandora. The individual value frequency drilldowns that support this combine drilldown are all instantaneous in operation.

Sorting by reverse frequency

The Values by Column drilldown cannot be manipulated and acts like a static report. However, an alternate drilldown is also provided on the **Columns** folder where the count columns are all sorted in reverse order. This can take significant time to assemble if there a large number of columns and rows as many sorts need to be executed. Experian Pandora does index everything, so repeating the drilldown would be instantaneous the second time around.

To view the drilldown, expand a table in the explorer, then right click on the **Columns** folder and select **Multi-Column Frequency Analysis**. The drilldown that appears is as before, but the values per column pair are sorted in frequency order (most common values per column first).

Heat map Capabilities

The **Multi-Column Value Analysis** drilldown is heat map enabled. In this case, it colours every column using a heat factor that is based on the number of times the value occurs in the column in question.

Data Values Drilldown

The Values item under **Data** on the explorer provides drilldowns showing **every** unique value across **every** column of **every** table loaded into Experian Pandora. This drilldown provides the usual filtering and transformation capabilities found on other data and metadata drilldowns, meaning that it is possible to filter and search every value.

Right-click on the Values folder and select **Values.** Use this to view a breakdown of unique values in use across all tables in the database, by datatype:

	Item	Info	Description
1	Alphanumeric Values	235,071 (71.94% of all Values)	The entire set of unique Alphanumeric Values across all Tables and Columns stored in the Database
2	Date Values	13,316 (4.08% of all Values)	The entire set of unique Date Values across all Tables and Columns stored in the Database
3	Decimal Values	51 (0.02% of all Values)	The entire set of unique Decimal Values across all Tables and Columns stored in the Database
4	Integer Values	78,272 (23.96% of all Values)	The entire set of unique Integer Values across all Tables and Columns stored in the Database
5	Money Values	34 (0.01% of all Values)	The entire set of unique Money Values across all Tables and Columns stored in the Database

Double-click on any of these to see the Values drilldown segmented by datatype.
Right-click on the Values Folder on the table and select **All Values (All Datatypes)** to see all of the unique values for all datatypes:

	Value	Value Length	Total Occurrences	Column Usage Count	Datatype	Simple Format
1	# 256, ShenHe Qu QingNian Street	32	1	1	Alphanumeric	#S9,SASASASA
2	#1 Cooperative Way	18	1	1	Alphanumeric	#9SASA
3	#1000 205 9th Ave. SW	21	2	1	Alphanumeric	#9S9S9ASA.SA
4	#108B 949 W. 3rd St.	20	2	1	Alphanumeric	#9AS9SA.S9ASA.
5	#151/1, OPP IIMB, bannerghatta Road	35	1	1	Alphanumeric	#9/9,SASA,SASA
6	#193, First Floor, R.V. Road	28	2	1	Alphanumeric	#9,SASA,SA.A.SA
7	#2 Meridian Studios	19	2	1	Alphanumeric	#9SASA
8	#200 Ren Min Da Dao	19	1	1	Alphanumeric	#9SASASASA

Finding columns in any table that use a value

Double-clicking on any value in the Values drilldown will show you the columns that contain that value no matter where they reside in the database. We have highlighted **E141184** as it occurs in 2 columns according to the Column Usage Count.

Double-click on it to see the columns where it is used:

	Value Usage Count	Name	Table	Uniqueness	Unique Count	Completeness	Score	Row Count
1	2	Customer Id	Sales Order Header	50.75%	37,707	100%		74,297
2	2	Customer Id	Purchase Order Header	50.75%	37,707	100%		74,298

This provides a filtered version of the Columns drilldown. You can select any of the columns now, right-click on it and choose **View rows for the searched value**:

Multiple values can be searched for in the same manner. This type of navigation is always instantaneous in Experian Pandora.

The standard value and format search capabilities are also available from the Values drilldown.

Data Formats Drilldown

The Formats item under **Data** on the explorer provides drilldowns showing **every** unique format pattern used in **every** column of **every** table loaded into Experian Pandora. This drilldown provides the usual filtering and transformation capabilities found on other data and metadata drilldowns, meaning that it is possible to filter and search every format.

Right-click on the Formats folder and select **Formats** to see a breakdown of unique values by datatype that are in use in the table.

Double-click on any of these to see the Formats drilldown segmented by datatype, or right-click on the Formats Folder on the table and select **All Formats** to see all of the unique formats for all datatypes:

Finding columns in any table that use a format

Double-clicking on any format in the Formats drilldown will show you the columns that contain values matching that format no matter where they reside in the database. In the previous section, we have highlighted **9SAS9AASAA** as it occurs in 4 columns according to the Column Usage Count.

Double-click on it to see the columns where it is used:

Columns (All Versions)
Filter: *Columns using the Format 9SAS9AASAA*

4 Rows (4.49%)

	Format Usage Count	Name	Table	Uniqueness	Unique Count	Completeness	Score	Row Count	Null Count
1	4	Addressline	Customer V1	77.18%	25,822	99.99%		33,455	1
2	4	Addressline	Customers	75.33%	27,259	99.99%		36,184	1
3	4	Addressline	Testing (v1)	77.18%	25,822	99.99%		33,455	1
4	4	Addressline	Testing (v2)	77.18%	25,822	99.99%		33,455	1

This provides a filtered version of the Columns drilldown. You can select any of the columns now, right-click on it and choose **View Rows matching the format 9SAS9AASAA**:

Rows for Customers (Table)
Filter: *Addressline matches the format '9SAS9AASAA'*

12 of 15 Columns

4 Rows (0.01%)

|	Customer Id	Discount Code	Forename	Surname	Company Name	Addressline	City	Country	Post Code
1	D20-E102902	M	Sara	Mittal	Ppl Services	2 N 9th St	Allentown	United States	18101-1139
2	D31-E130184	H	Barbara	Thompson	American Financial Group, Inc.	1 E 4th St	Cincinnati	United States	45202-3717
3	D47-E130134	N	Mike	Barg	Virginia Department of Social Services	7 N 8th St	Richmond	United States	23219-3301
4	D69-E147309	N	William	Cuypers	Hindustan Maize Products Ltd.	7 N 8th St	Ipswich	United Kingdom	IP9 9BS

Multiple formats can be searched for in the same manner. This type of navigation is always instantaneous in Experian Pandora.

The standard format search capabilities are also available from the Formats drilldown.

Global Quick Search

Quick Search provides a Google™-like capability to search all data in the Experian Pandora database, regardless of where it resides. Exact matching is virtually instantaneous. Fuzzy matching takes longer, but a fraction of what time it would take to analyse any number of tables. No standardisation of data is required to do this processing it is a fundamental facet of the Experian Pandora database.

Right-click on **Experian Pandora** in the explorer or **Home** in the navigation bar and select **Search > Quick Search**: The Quick Search dialog is very simple. If you wish to perform fuzzy matching just type one or more words and un-tick the **Exact Match** box if you wish to perform fuzzy matching.
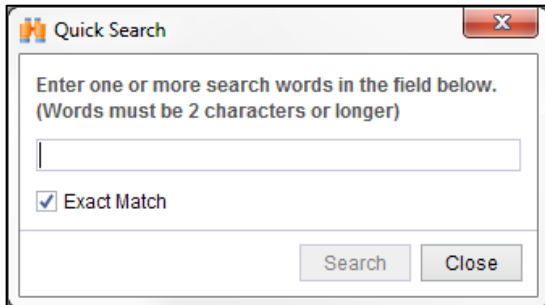
Quick Search is a brief form of Ranked Search and the results are identical. Please refer to the next section on Ranked Searching for more information on the results of Quick Search and Ranked Search.

Global Ranked Search

The Ranked Search allows you to set relative importance and effects of a match for multiple search terms either across all Tables in the database, or just one Table. The Quick Search is a simplified form of the Ranked Search when you are not concerned with relative weightings or labelling of results.
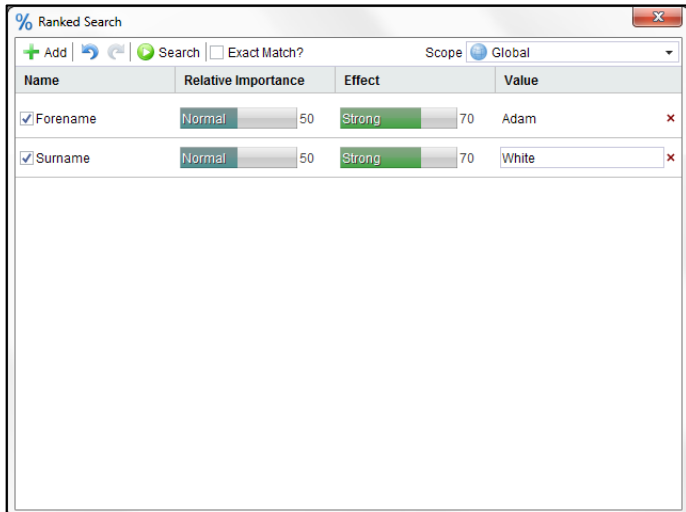
The Ranked Search Window initially is empty of any search terms.

Adding Search Terms

Click the **Add** button to add a new search term. Each term must have a name (which just a documentary label) and the value to search for.

Deleting Search Terms

To delete a search term click on the red **X** on the far right of each term. To disable the search for any term deselect the left tick box for that search term.

Undo and Redo Changes

The two blue arrows next to the Add button are **undo** and **redo** allowing changes to your search criteria to be undone if incorrect.

Controlling Match Strength

The **Exact Match** tick box is disabled by default for Ranked Search. If you want the faster exact match search, select this box. Otherwise Experian Pandora uses a variety of exact and inexact matching techniques including containment and edit distance.

Scoping the Ranked Query

The **Scope** for the search can either be Global (all Tables) or restricted to a specific Table. Use the **Scope** control to set your required search scope.

Controlling Result Ranking

The **Relative Importance** slider controls how much weight this search term adds to the overall rank result. The **Effect** slider controls the impact this search term makes to the result if it is not present. If it is relatively unimportant in its absence, lower this value. If mandatory set it to the right-most position.

Executing the Ranked Search Query

Both Quick Search and Ranked Search produce the same Ranked Search Results drilldown view. Hovering over any result gives you the reason for the match, which is also available in a number of different additional columns that are hidden by default. Coloration of values is also used to highlight which part of the value matches. Matches can be in any table in the database. Essentially only the unique values (either globally or for a table) are used in the match, with the analysis being performed in parallel. This is very fast and scales linearly with the number of unique values being assessed, which is a fraction of the data required to be searched compared to the actual data in all of the tables stored in the database.

Descriptions of the available columns

The Ranked Search Results view contains the follow standard columns:

Column Name	Column Description
Table Name	The name of the Table that the match resides in
Overall Rank	The percentage rank for quality of overall match of the results to the search terms
Rows	The number of rows in the Table that have a match for this search

The following set of columns appears once for **each** defined Search Term.

Column Name	Column Description
<Search Term>	The value matched for that search term highlighted to show how it was matched
<Search Term> Column	The column the match is in for the search term
<Search Term> Function	The match function used to match the search term
<Search Term> Rank	The rank for the match for the search term
<Search Term> Search	The actual value being searched for in the search term

Value Match Highlighting

Depending on how a value was matched, it will be highlighted differently:

Colour	Meaning
Green	Value matches exactly, regardless of case
Partly Red	Some of the characters of the value match the search term exactly ignoring case. Essentially the value contains the search term or contains a word that matches the search term.
Partly Blue	The value matches within a reasonable edit distance or contains a word that matches within a reasonable edit distance.
Orange	Value matches or partially matches the search term phonetically.
Black	Unmatched characters.

Here we show all of the hidden columns providing extending information about which column the value match occurred in and what the actual match rules were:

Hovering over any particular match gives you all the detail in a dynamic popup:

Viewing rows for the match results

Drilling down to one or more results gives you the actual rows in the table where the match was discovered. Multiple results can be selected for drilldown provided the results are from the same Table.

Cross-Table Relationship Discovery

Relationship Discovery Options

Auto Discovery Off (Default)

Before loading data, if you don't want Experian Pandora to discover relationships, right click Relationships in the Explorer under Data, or in the Data drilldown, and select Auto Discovery Options. To completely stop all Relationship discovery at time of load, un-tick the Auto-discover Relationships box.

If relationship discovery has been turned off, the icon on Relationships will be marked with a red line circle.

Auto Discovery On

If you want Experian Pandora to discover relationships at time of load, enable it by right clicking Relationships in the Explorer under Data, or in the Data drilldown, and select Auto Discovery Options. Tick the Auto-discover Relationships box.

If you want to discover relationships and wish to set a quality threshold, either use the slide bar or enter the percentage figure. You can also choose whether you want many-to-many, self-joins and intra-version relationships.

Automatic Relationship Discovery (Auto Discovery Must Be Turned On)

A cross-table relationship is defined by a value domain relationship between two Columns (typically in two Tables) in the database. Experian Pandora automatically identifies every relationship between single Columns among all Tables within the Experian Pandora Database. This relationship can be due to just a single value in common between Columns of the same or different tables.

No user action is required to discover these relationships; it happens automatically and incrementally as part of the data loading process with virtually zero cost. These are called **discovered** relationships. If for any reason you do not want Experian Pandora to discover these relationships when loading your data you will need to select the **Exclude from Relationships** option within the table on the Load and Analyse Data function. We will review the options for Relationship discovery using this method later in the chapter. For now we will look at just the Automatic Relationship Discovery.

You can use the **Relationship Finder** to find the relationships that are of interest and relevance and select to **save** them to the repository. These **saved relationships** are grouped at Column, Table, Project, Schema and Database levels for browsing purposes.

Experian Pandora keeps all relationships of any quality, but will only display those that you choose as being important, by saving them into the main repository. Cartesian product relationships between columns containing constant values are excluded from the results to avoid obviously meaningless relationships, e.g. between columns initialised to 0 or 1.

A wealth of metadata about the Column Relationship is calculated automatically and Experian Pandora can display any type of Relational Join between the Columns (and their owning Tables) in a fraction of the time of a conventional Database as well as unconventional relationships that cannot be easily performed in other types of Database.

When a table is loaded into Experian Pandora, all the Relationships are automatically *discovered*.

Joins that have been saved can be unsaved, which simply relegates them back to a status of discovered. The shortcut to the relationship that is visible on the Explorer disappears. Regardless of whether they have been saved or unsaved, all relationships remain in the repository unless the Tables involved in the relationships are deleted or they are individually deleted from the All Relationships drilldown. Only saved Relationships can have notes attached to them.

Rediscover Relationships

If you wish to change the way relationships have been discovered for tables after a load you need to first ensure that the Auto-discover Relationships option is enabled. At this point you can select the quality threshold either using the slider or entering the percentage above which you want to discover. You can choose to allow many-to-many and/or self-joins and/or intra-version relationships, which are relationships between versions of the same table. If you want to exclude any Schemas, Tables or Columns from the manual discovery you will need to edit the objects in question and tick the box **Exclude from Relationships**. If this already has a grey tick in

there it will mean you have not turned on the Auto Discovery Options or it is disabled at a higher level. E.g. if you exclude a table from relationship discovery then all its columns are implicitly disabled.

To rediscover the relationships you need to do a soft restart. Right-click **Home > System > Rebuild Relationships**, select Yes in the dialog box which asks "Are you sure you want to restart the server and rebuild all discovered relationships". Experian Pandora will show it is Reconnecting. This can take some time to complete.

In the Explorer under Data > Relationships you will see the number of discovered relationships once the rebuild has completed.

Manual Relationship Discovery

If you want to discover a specific relationship between two columns or the same of different tables, this can be done at any time, even if Automatic Relationship discovery is turned off. This manual discovery process will always find a relationship and ignores all filtering options. To do this, open the (All) Columns drilldown. Highlight the two columns you wish to look at the relationship for, right-click and select **Re-analyse Relationship.** This runs as a background job as it could take some time and once this has completed you can right-click again and select **View Relationship Venn.**

Discovered Relationships Drilldown

All of the relationships discovered by Experian Pandora can be viewed in a single drilldown. They can be found in the Explorer under **Data**, right-click **Relationships** and select **Relationships (All)**.

This is a detailed drilldown with a huge array of information. You can filter any aspect to find relationships matching whatever criteria you so wish. It is far easier to use the Relationship Finder for this task.

Colour Coding

Note the colour coding in the drilldown:
- All aspects of the left-hand-side are coloured green.
- All aspects of the right-hand side are in blue, while aspects in common are in yellow. These are identical to the colours on the Venn Diagrams available for relationships.
- All quality measurement columns are variable colour bar graphs and documentary information is in blue italics.
- The name of the relationship if automatically generated is in grey italics.

Saved Relationships Drilldown

This is identical in format to the Discovered Relationships Drilldown and can be accessed from the explorer by going into **Data**, right clicking on **Relationships** and selecting **Relationships**. This drilldown will contain only relationships selectively saved by users.

Profile for a Relationship

Go into a relationship in the explorer and click on **Profile** to bring up the **Profile** drilldown. The drilldown produced contains all of the statistics for the relationship automatically inferred from the data by Experian Pandora.

Info for a Relationship

Go into a relationship in the explorer and click on **Info** to bring up the **Info** drilldown. The summary drilldown encapsulates all the profile information as well as the documented metadata for the relationship.

Relationship Report

Right-click on a relationship and select **Report** from the menu.

Venn Diagram

Right-click on a relationship and select **Venn Diagram**. The Venn diagram will appear on the desktop as a new drilldown with a split pane, the top half showing the Venn diagram itself, and the bottom half showing the profile drilldown for the relationship:
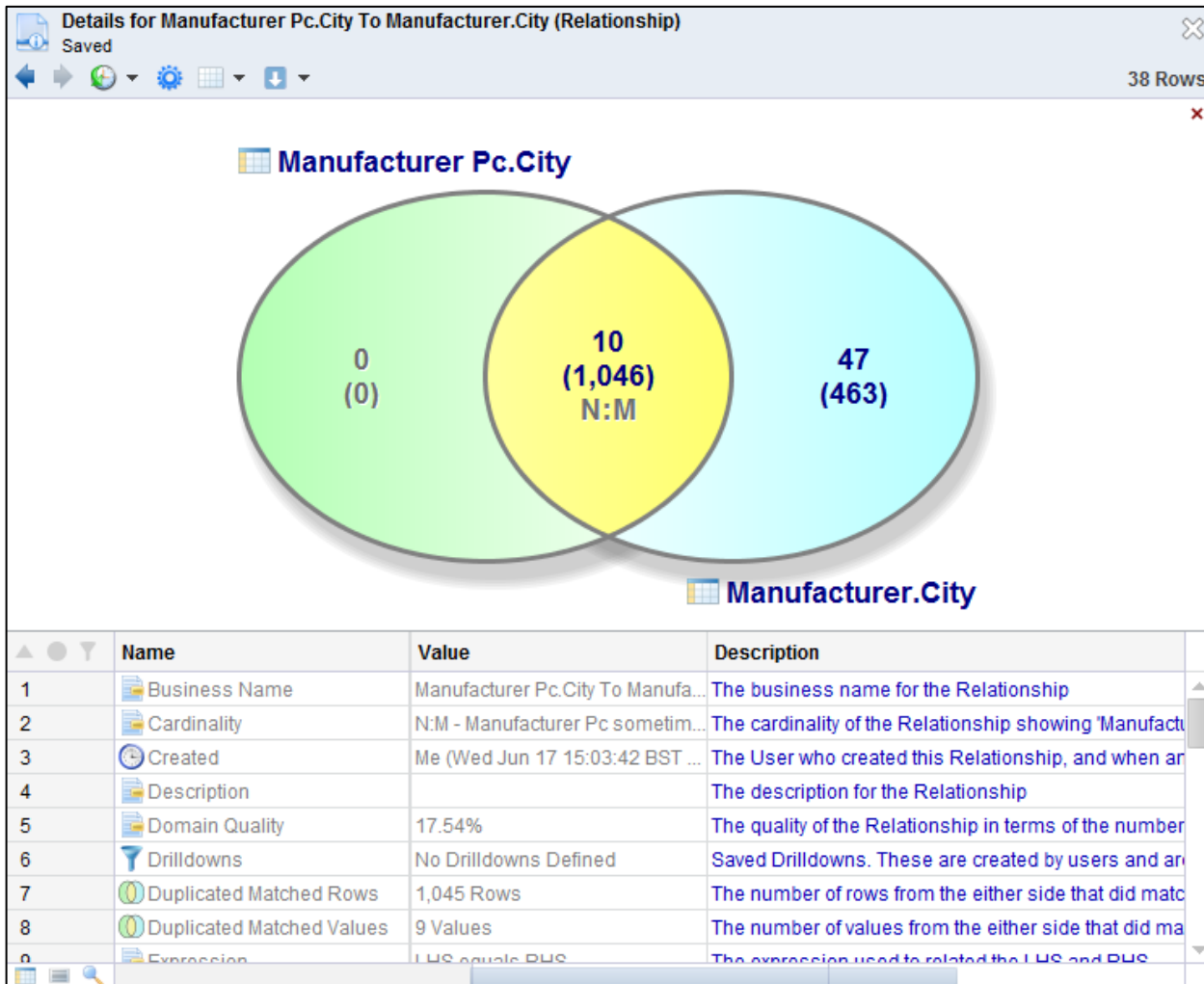
- Right-clicking on the Venn diagram presents the standard relationship right-click menu with all the usual option.

- Single clicking on the numbers will result in a corresponding drilldown occurring.

- The top line numbers are the statistics of the relationship with respect to the value domain relationship between the left and right columns.

- The numbers in brackets are the statistics of the relationship with respect to the table-to-table join relationship between the left and right tables joined on the left and right columns.

- Green indicates the left, yellow the middle (common) and blue denotes the right.

- Where the green and blue circles intersect that denotes what the columns/tables have in common.

- Clicking on the top number in the yellow intersection gives a drilldown of all values in common between the left and right (the **common domain**).

- Clicking on the bottom number in the yellow intersection gives a drilldown of the joined rows that result from the intersection of the left and right side value sets (the **inner join**).

- Clicking on either of the two column symbols presents information about the left or right columns in a drilldown.

- Clicking on the top number in the green section (the left) shows all values that did not match from the left-hand column.

- Clicking on the bottom number in the green section shows all rows that did not match from the left. This may be because of unmatched values or nulls.

- Clicking on the top number in the blue section (the right) shows all values that did not match from the right-hand column.

- Clicking on the bottom number in the blue section shows all rows that did not match from the right. This may be because of unmatched values or nulls.

Value Domain Drilldowns

The Value Domain Drilldowns compare values from the left and/or right columns in a combined frequency drilldown.

Common Values for Manufacturer Pc to Manufacturer on City — 10 Rows

	Value	LHS Rows	RHS Rows	Presence	Value Length
1	Birmingham	1	13	Both	10
2	Weston Rhyn	1	1	Both	11
3	COVENTRY	1	36	Both	8
4	LEICESTER	1	184	Both	9
5	Herefordshire	1	7	Both	13
6	OSWESTRY, Salop	1	6	Both	15
7	WOLVERHAMPTON	2	62	Both	13
8	WORCESTER	1	15	Both	9
9	West Midlands	3	199	Both	13
10	Worcestershire	1	63	Both	14

Different types of Value Domain Drilldowns

There are a number of different value domain related drilldowns for a relationship:

Left Side Drilldowns

Unmatched Values

Right-click on a relationship and select **Left > Values Unmatched.** This is a drilldown of all values from the left hand column that did not match those of the right.

All Values

Right-click on a relationship and select **Left > All Values.** This is a drilldown of all values from the left hand column.

Middle Drilldowns

Matched Values

Right-click on a relationship and select **Right > Values Matched.** This is a drilldown of all values that matched from both the left and the right hand columns.

Unmatched Values

Right-click on a relationship and select **Right > Values Unmatched.** This is a drilldown of all values that didn't match from both the left and the right hand columns.

All Values

Right-click on a relationship and select **Right > All Values.** This is a drilldown of all values from the right hand column.

Right Side Drilldowns

Unmatched Values

Right-click on a relationship and select **Right > Values Unmatched.** This is a drilldown of all values from the right hand column that did not match those of the left.

All Values

Right-click on a relationship and select **Right > All Values.** This is a drilldown of all values from the right hand column.

Joined Row Drilldowns

The joined row drilldown is the join between two tables. It contains all of the columns of the left-hand-side table and those of the right-hand-side table.

- The value in common is only ever included once in the drilldown, with a heading marked in yellow.
- The columns from the left side have their headings marked in green, with those of the right side in blue.
- Name clashes from the left and right are handled by automatically renaming clashing right-hand columns with an increasing number. For example, if there was a Price column on both sides, the second instance of it would be automatically renamed **Price1**.
- Hovering over the column header provides the full canonical name of the column, including its parent table:

- The Edit dropdown menu includes a wide variety of available drilldowns and reports:
  - Table rows for the left hand and right tables
  - The relationship report
  - Table Quality Reports for the left and right hand tables
  - Quick mapping report to show the details behind the join relationship between the tables

Relationship for Manufacturer Pc to Manufacturer on City
Saved

	Manufacturer Id	Name	Company Id	Addressline1	Addressline2	City	Zip
1	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
2	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
3	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
4	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
5	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
6	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
7	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
8	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
9	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
10	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
11	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
12	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34
13	19911697-C	Dilo Engineering	DLN28	94 Middle Leaford	Stechford	Birmingham	B34

Different Types of Joined Row Drilldowns

There are a number of different joined row related drilldowns for a relationship:

Left Side Drilldowns

Unmatched Rows

Right-click on a relationship and select **Left > Rows Unmatched.** This is a drilldown of all rows from the left hand table that did not join with the corresponding right hand columns left blank.

All Rows

Right-click on a relationship and select **Left > All Rows.** This is a drilldown of all rows from the left hand table.

Middle Drilldowns

Matched Rows

Right-click on a relationship and select **Right > Rows Matched.** This is a drilldown of all rows that matched from both the left and the right hand columns.

Unmatched Rows

Right-click on a relationship and select **Right > Rows Unmatched.** This is a drilldown of all rows that didn't match from both the left and the right hand columns.

All Rows

Right-click on a relationship and select **Right > All Rows.** This is a drilldown of all rows from the right hand column.

Right Side Drilldowns

Unmatched Rows

Right-click on a relationship and select **Right > Rows Unmatched.** This is a drilldown of all rows from the right hand column that did not match those of the left.

All Rows

Right-click on a relationship and select **Right > All Rows.** This is a drilldown of all rows from the right hand column.

Saving Relationships

From a relationship drilldown, right-click on the relationship and select **Save**. The status of the Relationship will change from **Discovered** to **Saved**. Highlight more than one relationship to be able to save all in one go.

Adding Notes

From a relationships drilldown, you have the option to add notes to a relationship if it has been saved. You can also add notes to the left or right hand side tables or columns. From the explorer you may only add notes to the join directly.

Unsave Relationships

To unsave a relationship, right-click on it in the explorer and select **Unsave,** or from a relationships drilldown select **Unsave**. Its status will change from **Saved** to **Discovered** and it will be removed from the explorer. Any notes attached to the relationship will be removed.

Deleting Relationships

Any discovered relationships that you are not interested in can be deleted via the All Relationships drilldown. Highlight the relationships you do not want to keep, right-click select **Delete Relationships.** A warning message will appear to confirm you wish to delete these relationships. Select Yes if you want to go ahead.

Relationship Cardinalities

Cardinality describes the relationship from the point of view of the cardinality of rows joined from the left hand table to the right hand table. It is formed of a left-hand moniker and a right-hand moniker separated by a colon.

Cardinality

The possible monikers are:-

Moniker	Description
0	Matches occur at most once with some unmatched rows
1	Matches occur at most once with no unmatched rows
M	Matches occur one or more times with some unmatched rows
N	Matches occur one or more times with no unmatched rows

The cardinality can tell you a lot about a relationship without even looking at the data. For example, matching Customer to Address, you may have expected a customer to always have an address, and to only have one address. In this case the expected cardinality would be **1:1**. If the actual cardinality was, say, **0:M** it would tell you straight away that some customers have more than one address, some customers have no address and worst of all there are addresses with no corresponding customers!

In general, joins tend to be asymmetrical, in that one side does not feature duplicates. For example, the relationship between Purchase Order and Purchase Order Line Item would be a 1:N (One to Many with no missing records). Many-to-many relationships (M:M, M:N, N:N or N:M) can occur just through coincidence and should never be modelled in any real database. Note the use of the word **should**!

Finding Relationships

Although it is possible to filter the various relationships drilldowns to find what you are looking for, Experian Pandora provides a number of much faster and friendlier ways to navigate the relationships that are automatically discovered.

The relationship finder is a wizard that gives you the ability to search for relationships within a specific scoped area. This can be:

- Tables in a Schema
- Tables in a Project
- Columns in a Table
- A specific Table
- A specific Column
- All versions of a Table
- An arbitrary selection of Tables from a Tables drilldown
- An arbitrary selection of Columns from a Columns drilldown

You will find the option to **Find Relationships** in the explorer by right clicking on the item you want to search from, or by right clicking on a **Relationships** folder**.**

In this example, we right click on the Purchase Order Header table and select **Find Relationships.** The relationship finder will appear as a child window:
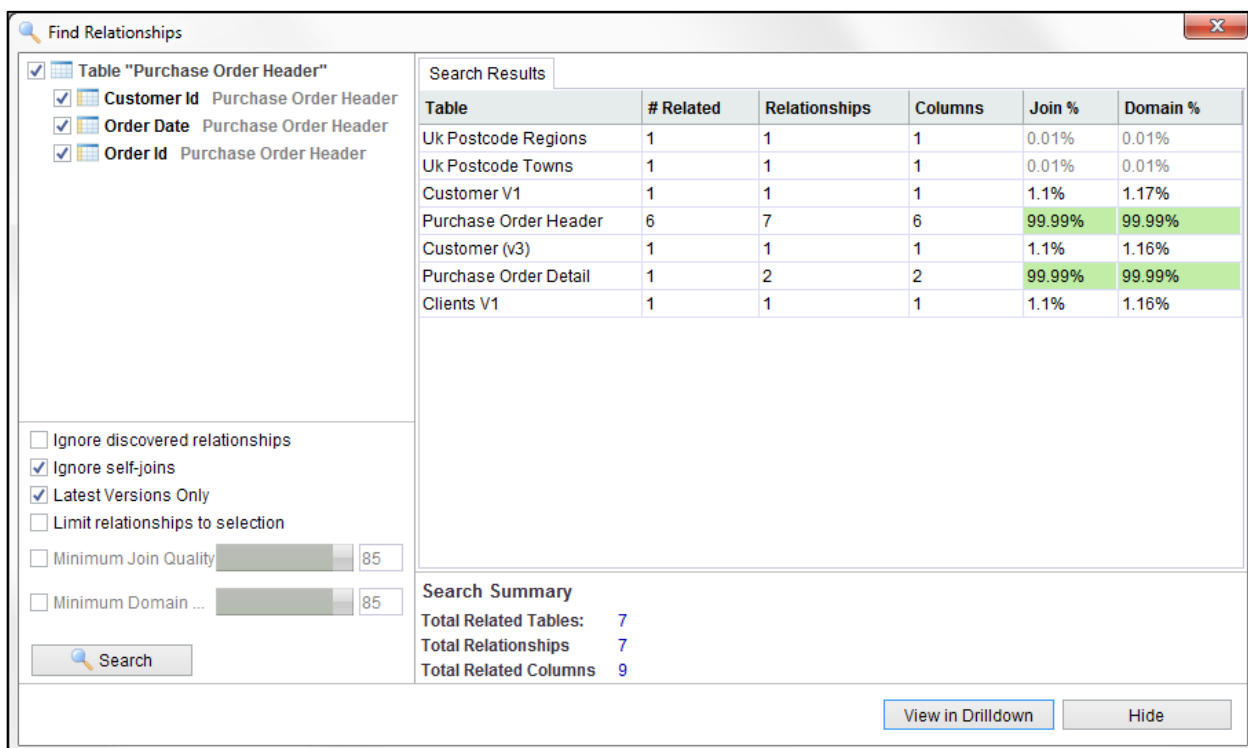

The left-hand panel provides controls for limiting the search scope, while the right-hand-side provides the results of the search and a summary of the search.

The key features of the Relationship Finder dialog are as follows:

Selection Tree

Use this tree to select the scope of the relationship search. You can check or uncheck any item that you do not want to include in the search, for example eliminating certain columns.

Ignore Discovered Relationships

This limits searching to only the relationships that have been permanently saved to the repository, and eliminates all discovered relationships.

Ignore Self-Joins

This eliminates any joins where the left and right tables are the same (i.e. a join to ones-self)

Latest Versions Only

This limits searching to the latest versions of the tables that you are using.

Limit Relationships to Selection

If this is selected then the finder only looks for relationships between the selected items; otherwise it will find all relationships across all the tables in the entire Experian Pandora repository where the items in the selection tree appear on one side of the relationship.

Minimum Join Quality

This allows you to specify minimum threshold for the Join Quality. This is a measure of the number of rows joined between the left and right tables.

Minimum Domain Quality

This allows you to specify minimum threshold for the Domain Quality. This is a measure of the number of values joined between the left and right tables.

Search Button

Clicking **Search** searches all relationships using the supplied filtration rules, and produces a table of results in the Search Results panel.
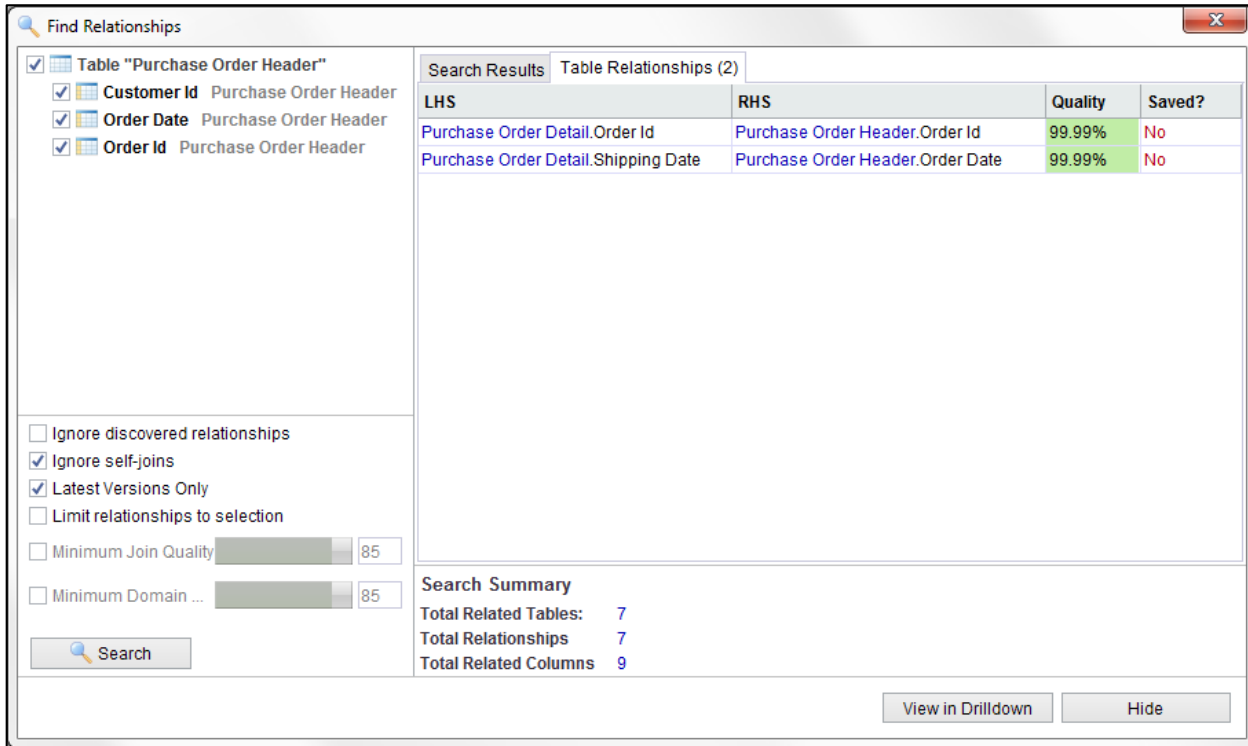
Search Results

This is a list of all tables found that participate in the relationships found based on the search criteria. The columns in the search results are as follows:

Name	Description
Table	The name of the Table
# Related	The number of other tables related to it based on the relationships found
Relationships	The number of relationships the table is involved in from the relationships found
Columns	The number of columns in the table involved in relationships from the set of relationships found
Join %	The best quality of table join from the set of relationships found
Domain %	The best quality of domain relationships from the set of relationships found

Table Relationships

This tab only appears after a table search result has been double-clicked. In this case we have double-clicked on Purchase Order Detail. The result is the list of relationships from the set of relationships found that correspond to the table that was double-clicked on. To return to the search results, click on the Search Results tab. The number of relationships displayed in the Table Relationships tab will be the same as the number in the **Relationships** column of the Search Results table.

The Table Relationships table has 4 columns:

Column	Description
LHS	The full hierarchical name of the left column including Table name
RHS	The full hierarchical name of the right column including Table name
Quality	The quality of the relationship
Saved?	Whether the relationship has been saved or not

Right-clicking on a relationship gives a pop-up menu to:

- Save the relationship if it is not yet saved
- Unsave the relationship if it has already been saved
- Delete the relationship
- View the Relationship Report for the relationship

Selecting multiple relationships and right-clicking on them gives a menu to:

- Save all selected relationships if not saved
- Unsave the selected relationships if they have already been saved
- Delete all selected relationships

Relationship Summary

The summary provides information about the number of related tables discovered from the search criteria, the number of relationships found and the total number of columns involved in the relationships.

Search Summary	
Total Related Tables:	7
Total Relationships	7
Total Related Columns	9

View in Drilldown Button

This button filters the entire relationships drilldown to show the same relationships listed in the Table Relationships tab to allow you to do further filtering to narrow down selection outside of the capabilities of the Relationship Finder

Hide Button

This hides the Relationship Finder by minimising to the Experian Pandora Desktop taskbar. Clicking on the red close button will also hide it.

Venn Diagram

Double-clicking on a relationship in the Table Relationships tab will display the Venn Diagram for that relationship within the Relationship Finder window:

This is identical to the Venn Diagrams documented earlier in this section of the manual with the same drilldown capabilities and access to all aspects of the relationship by right-clicking on the Venn Diagram to see the relationship menu.

Relationship Statistics

A second tab appears next to the Venn Diagram tab, which is the Relationship Statistics tab, which contains the complete summary of profile statistics and metadata for the relationship. These are active items with drilldowns that can be invoked by double-clicking them.

Arbitrary Table Relationship Searching

You can find the relationships between one or more specific tables by simply selecting them from any Tables drilldown and from the right-click menu selecting **Find Relationships**. This invokes the Relationship Finder with just the specifically selected tables and the option to **Limit Relationships to Selection** being set.

Arbitrary Column Relationship Searching

You can find the relationships between one or more specific columns by simply selecting them from any Columns drilldown and from the right-click menu selecting **Find Relationships**. This invokes the Relationship Finder with just the specifically selected tables and the option to **Limit Relationships to Selection** being set.

Relationships in the Explorer

All saved relationships

In the explorer, go to **Data > Relationships.** The relationships item can be expanded if there are any relationships that have been saved:
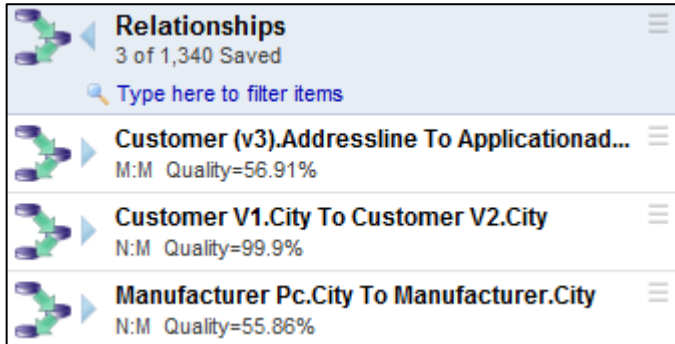
We can now see all of the saved joins in the Experian Pandora repository. The relationships tell you the left and right tables and columns, as well as the cardinality and the overall quality of the saved relationships. Right-clicking on these relationships provides the standard relationship menu of capabilities. Expanding the folder for relationship displays folders for the profile and info of the relationship, as well as an item to access the number of rows that result in the Inner Join (the rows in common).
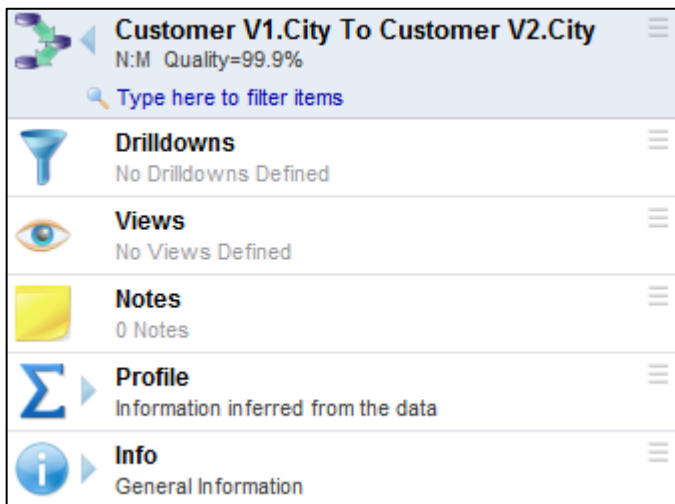
Expanding the **Profile** folder presents all of the inferred information about the relationship discovered by Experian Pandora from the actual data.

Expanding the **Info** folder provides general information about the relationship.

Relationships for a Table

If there are saved relationships, the tables involved in the relationships will have their **Relationships** folder populated.
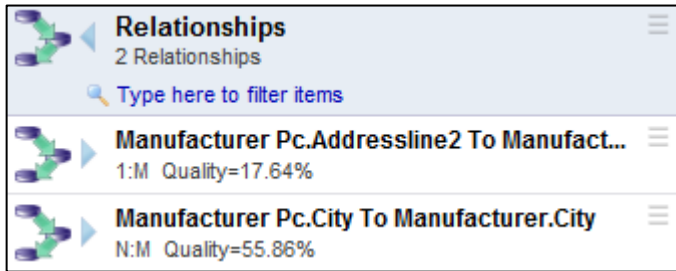
 We can see that there is a relationship between Manufacturer Pc and Products. The relationship to Products is excellent as it is at 100%. We can see a list of the relationships there are on a table, the cardinality of the relationship and the join quality.

Relationships for a Column

If there are saved relationships, the columns involved in the relationships will have their **Relationships** folder populated. Expanding the **Relationships** folder, we can now see a list of all the currently saved relationships related to that column.
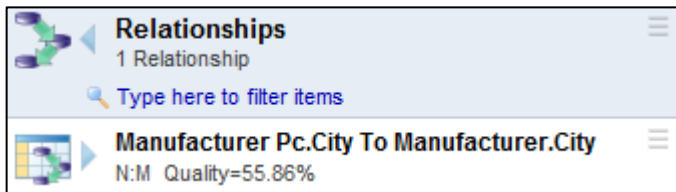
Expanding the relationship provides the usual information for a relationship:
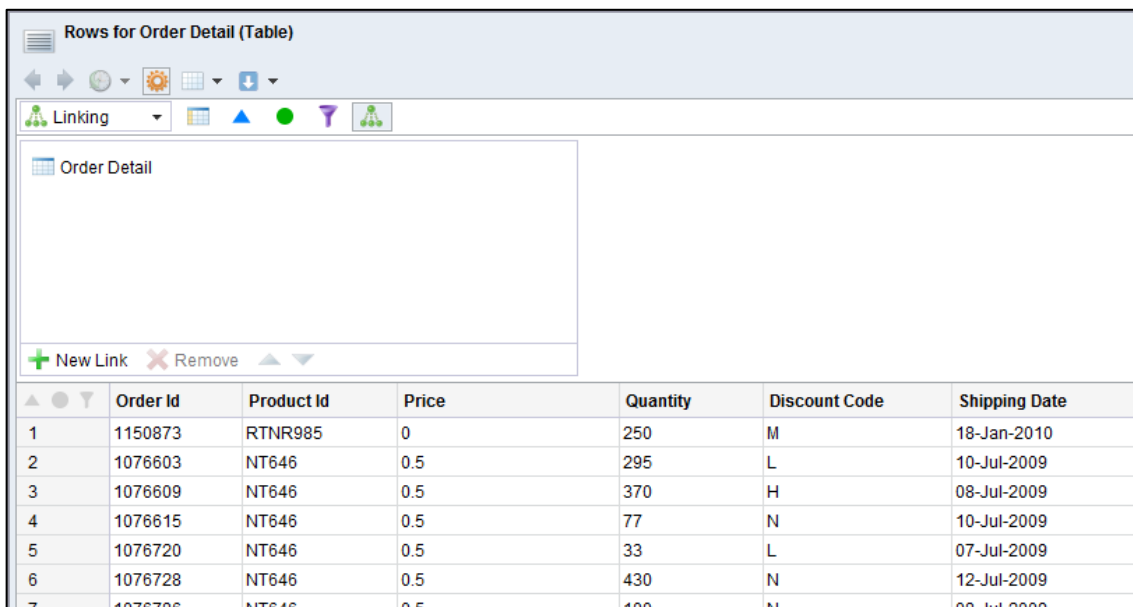
Table Linking

The table linking functionality offers a convenient way to create a multi-table join in a single step without the need for saving intermediate tables.

For example, to join tables ORDER_DETAIL, ORDER_HEADER and CUSTOMERS on the columns ORDER_DETAIL.ORDER_ID <-> ORDER_HEADER.ORDER_ID and ORDER_HEADER.CUSTOMER_ID <-> CUSTOMERS.CUSTOMER_ID using Discovered Relationships would require the Relationship between ORDER_DETAIL and ORDER_HEADER to be saved as a Table (DETAIL_TO_HEADER) and then a Relationship between DETAIL_TO_HEADER.CUSTOMER_ID and CUSTOMERS.CUSTOMER_ID to then be discovered. The step of saving the Relationship between A and B could introduce significant additional time to the process, depending on the respective row counts of the tables concerned. Table linking offers an alternative to this approach.

Creating a Table Link Hierarchy

Using the above example as a reference, the first step to creating the table link hierarchy is to drilldown to the rows of the first table, **Order Detail** and then open the Drilldown Manager on the **Linking** tab:

The objective in this example is, to use the ORDER_DETAIL table as the starting point, and join it to ORDER_HEADER and then CUSTOMERS in order to obtain the order date, customer id, forename, surname and postcode associated with each sales order.

Adding Links

Clicking the **New Link** button or right-clicking on Order Detail and selecting **New Link**, will present a dropdown which allows you to select a table to link with the **Order Detail** table:

You can also drag tables from the Explorer onto the link hierarchy in the Linking tab by clicking and holding on the table icon beside the table name and dragging it into the Drilldown Manager and holding it over the table for the link. The insert position will be shown as an orange line on the link tree:

Selecting the **Order Header** table will add it to the link hierarchy. A panel will appear on the right-hand-side:

Editing Link Criteria

The link criteria panel appears on the right-hand-side of the Linking tab whenever you create a new link or click on an existing one on the link tree on the left. This panel allows you to specify:

- The LHS and RHS columns to be joined
- The criteria that determine which rows to return

The LHS/RHS columns are automatically populated with the best discovered relationship between the two tables (if any exist). In the above example, the highest-quality relationship happens to be between ORDER_DETAIL.SHIPPING_DATE and ORDER_HEADER.ORDER_DATE, which is not what we need in this case. The join criterion can be changed by clicking the small dropdown arrow to the right which will present alternative discovered joins or by clicking each column label and choosing manually from a dropdown list of available columns.

The three checkboxes above the column join criteria determine the specific rows from the RHS table that are returned:

Singletons – If you want the first match of a 1 to many match AND any matches that occur only once, just turn Singletons on (this is a normal relational join, with duplicates removed)

Duplicates – If you want the all the matches of things that are 1 to many only, just turn duplicates on (this is a normal relational join, less things that match only once)

No Matches – if the LHS and RHS columns do not match, a blank record will be returned AND the first match of any 1 to many matches.

If you want the first match of a 1 to many match, have nothing selected (this is a normal relational join, with duplicates removed and things that match only once removed)

If you want anything that matches one or more times, put singletons and duplicates on (this is a normal relational join)

If you want anything that matches one or more times AND things that don't match, turn everything on (this is a Left-Outer join effectively)

NB: you **cannot** do right-outer join operations with Table Linking as the driver is the LHS always by definition.

In this example, we will leave all three selected.

The next step is to link the **Customers** table to **Order Header** in the same way described above.
If Experian Pandora can't discover any relationships you will see a red exclamation mark over the table name and in the right hand panel you will see Undefined in red.   Click on Undefined and make your selections manually:

We can now select the Columns tab on the Drilldown Manager to choose the columns to be displayed in the final result:

526,374 Rows

| Columns | | | | | |

Hidden Columns

Search

Visible Columns

Search

Hidden Columns		Visible Columns	
Row Id		Order Id	
Source Row Id		Product Id	
Customer Id	Customers	Price	
Discount Code	Customers	Quantity	
Company Name	Customers	Discount Code	
Addressline	Customers	Shipping Date	
City	Customers	Customer Id	Order Header
Country	Customers	Order Date	Order Header
Telephone	Customers	Forename	Customers
Email	Customers	Surname	Customers
Credit Limit	Customers	Post Code	Customers
Order Id	Order Header		

Clicking the **Run Query** button will run the final query:

527,786 Rows

	Order Id	Product Id	Price	Quantity	Discount Code	Shipping Date	Customer Id Order Header	Order Date Order Header	Forename Customers	Surname Customers	Post Code Customers
1	1150873	RTNR985	0	250	M	18-Jan-2010	E117212	16-Jan-2010	Kenneth	Hopkins	DH7 8DT
2	1076603	NT646	0.5	295	L	10-Jul-2009	E114373	10-Oct-2009	Roger	Tuck	PA20 0SA
3	1076609	NT646	0.5	370	H	08-Jul-2009	E117055	06-Jul-2009	John	Godwin	60610-2932
4	1076615	NT646	0.5	77	N	10-Jul-2009	E132421	08-Jul-2009	Geoffrey	Hughes	CF45 6PD
5	1076720	NT646	0.5	33	L	07-Jul-2009	E107576	08-Jul-2009	Ann	Woodfield	CW6 7JC
6	1076728	NT646	0.5	430	N	12-Jul-2009	E103627	03-Jul-2009	Arron	Culshaw	EX37 6WN
7	1076786	NT646	0.5	100	N	08-Jul-2009	E130825	02-Jul-2009	Gail	Saks	PA60 9SX
8	1076796	STSCRW587	0.5	6	M	08-Jul-2009	E138237	14-Jul-2009	Nicholas	Leighton	PE32 1UP
9	1076919	NT646	0.5	2400	H	04-Nov-2009	E150915	05-Jul-2009	William	Seaman	AB54 1DF
10	1076968	NT646	0.5	3750	N	07-Jul-2009	E132192	16-Nov-2009	Edward	Do	BH20 1YF
11	1076970	STSCRW587	0.5	47	L	06-Jul-2009	E116453	02-Jul-2009	Greg	Davis	6902
12	1076981	NT646	0.5	365	L	08-Jul-2009	E109107	02-Jul-2009	Eugene	Orchard	TD15 7AG
13	1077001	STSCRW587	0.5	1050	L	09-Jul-2009	E135826	28-Oct-2009	Leslie	Mayer	TN26 9NK
14	1077013	NT646	0.5	315	H	05-Jul-2009	E143283	18-Jul-2009	Alex	Lowe	M1S 4E7
15	1077044	NT646	0.5	8	N	09-Jul-2009	E129628	29-Nov-2009			
16	1077129	NT646	0.5	6	N	15-Jul-2009	E127701	17-Nov-2009	David	Toxidis	BA16 6GG
17	1077179	STSCRW587	0.5	4100	M	07-Jul-2009	E130979	08-Jul-2009	Katherine	O connell	DN1 6DX
18	1077181	NT646	0.5	395	N	05-Jul-2009	E102413	07-Jul-2009	Wayne	Kennedy	FK8 5ES

Note the blank fields on row number 15, which indicates that there is no matching Customer Id = E129628 in the CUSTOMERS table.

If the **No Matches** checkbox on the link between Order Header and Customers had been disabled, these blank records would not appear in the final result:

499,953 Rows

	Order Id	Product Id	Price	Quantity	Discount Code	Shipping Date	Customer Id Order Header	Order Date Order Header	Forename Customers	Surname Customers
1	1150873	RTNR985	0	250	M	18-Jan-2010	E117212	16-Jan-2010	Kenneth	Hopkins
2	1076603	NT646	0.5	295	L	10-Jul-2009	E114373	10-Oct-2009	Roger	Tuck
3	1076609	NT646	0.5	370	H	08-Jul-2009	E117055	06-Jul-2009	John	Godwin
4	1076615	NT646	0.5	77	N	10-Jul-2009	E132421	08-Jul-2009	Geoffrey	Hughes
5	1076720	NT646	0.5	33	L	07-Jul-2009	E107576	08-Jul-2009	Ann	Woodfield
6	1076728	NT646	0.5	430	N	12-Jul-2009	E103627	03-Jul-2009	Arron	Culshaw
7	1076786	NT646	0.5	100	N	08-Jul-2009	E130825	02-Jul-2009	Gail	Saks
8	1076796	STSCRW587	0.5	6	M	08-Jul-2009	E138237	14-Jul-2009	Nicholas	Leighton
9	1076919	NT646	0.5	2400	H	04-Nov-2009	E150915	05-Jul-2009	William	Seaman
10	1076968	NT646	0.5	3750	N	07-Jul-2009	E132192	16-Nov-2009	Edward	Do
11	1076970	STSCRW587	0.5	47	L	06-Jul-2009	E116453	02-Jul-2009	Greg	Davis
12	1076981	NT646	0.5	365	L	08-Jul-2009	E109107	02-Jul-2009	Eugene	Orchard
13	1077001	STSCRW587	0.5	1050	L	09-Jul-2009	E135826	28-Oct-2009	Leslie	Mayer
14	1077013	NT646	0.5	315	H	05-Jul-2009	E143283	18-Jul-2009	Alex	Lowe
15	1077129	NT646	0.5	6	N	15-Jul-2009	E127701	17-Nov-2009	David	Toxidis
16	1077179	STSCRW587	0.5	4100	M	07-Jul-2009	E130979	08-Jul-2009	Katherine	O connell
17	1077181	NT646	0.5	395	N	05-Jul-2009	E102413	07-Jul-2009	Wayne	Kennedy
18	1077182	NT646	0.5	45	M	15-Sep-2009	E119375	12-Jul-2009	Stephen	Hodgson

Deleting Links

Links can be deleted by selecting on the link tree and clicking the **Remove** button or right-clicking on the link and choosing the Remove menu action:
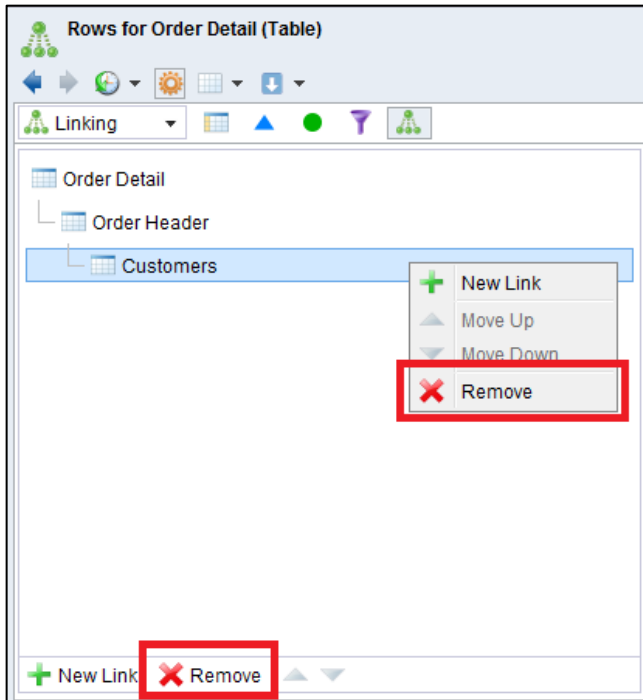
Re-arranging Links

Links at a particular level in the hierarchy can be re-positioned by selecting and clicking the up or down buttons on the toolbar, or by right-clicking on the link and choosing the **Move Up** or **Move Down** menu actions:
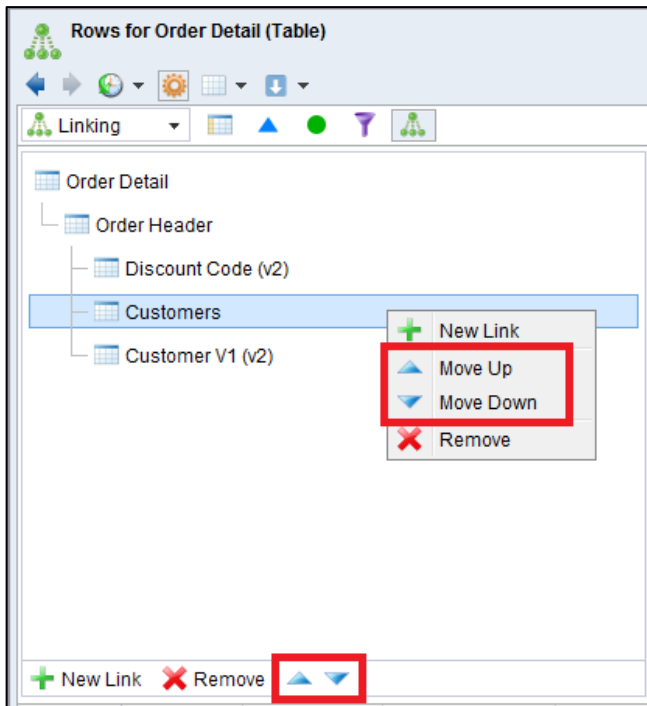
Object Triggers

You can set up triggers on tables and drilldowns. Triggers are actions to be performed when an object that a table is referring to updates.
There are four types to choose from:
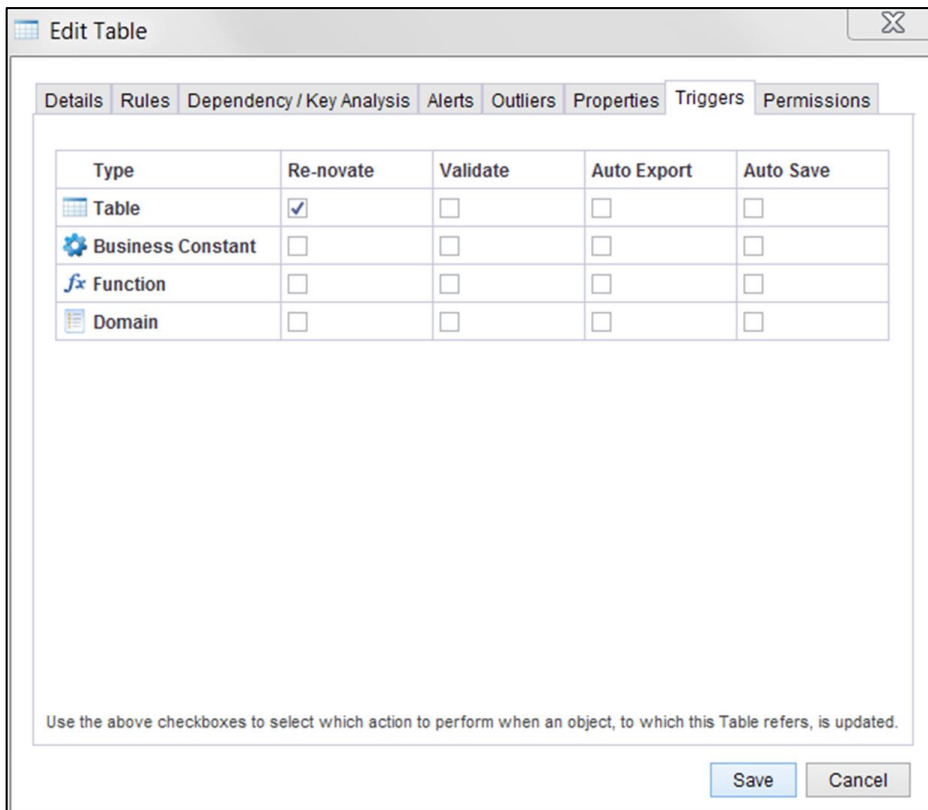
- Table
- Business Constant
- Function
- Domain

Each of these can trigger one or more jobs from the following options: Re-novate, Validate, Auto Export and Auto Save.

To use Triggers, you must have a referenced object of the allowed type. When the referenced object changes it will trigger the respective job to run.

Based on the first linking example above, create a drilldown where **Order Detail** is linked to **Order Header** and save it.
This saved Drilldown is parented by Order Detail and references Order Header, thus creating a link between the two tables. **Right click > Edit > Settings** on the Order Detail table, and go to the Triggers tab.

Then tick the checkbox for **Table > Re-novate**.

By selecting this option, you are telling Order Detail to re-novate when Table objects to which Order Detail refers to (in this case Order Header) are updated.

In the example, Order Detail refers to Order Header via the link in the saved Drilldown. So when you load a new version of Order Header, this causes Order Detail to re-novate, which in turn novates the linked Drilldown.

You can also set Triggers on Drilldowns.

For instance you may want the drilldown from the previous example to auto-save as a table whenever the tables referenced by the drilldown reloads. To do this, **Right click > Edit** on the saved drilldown, and go to the Triggers tab.

Tick the checkbox for **Table > Auto Save**
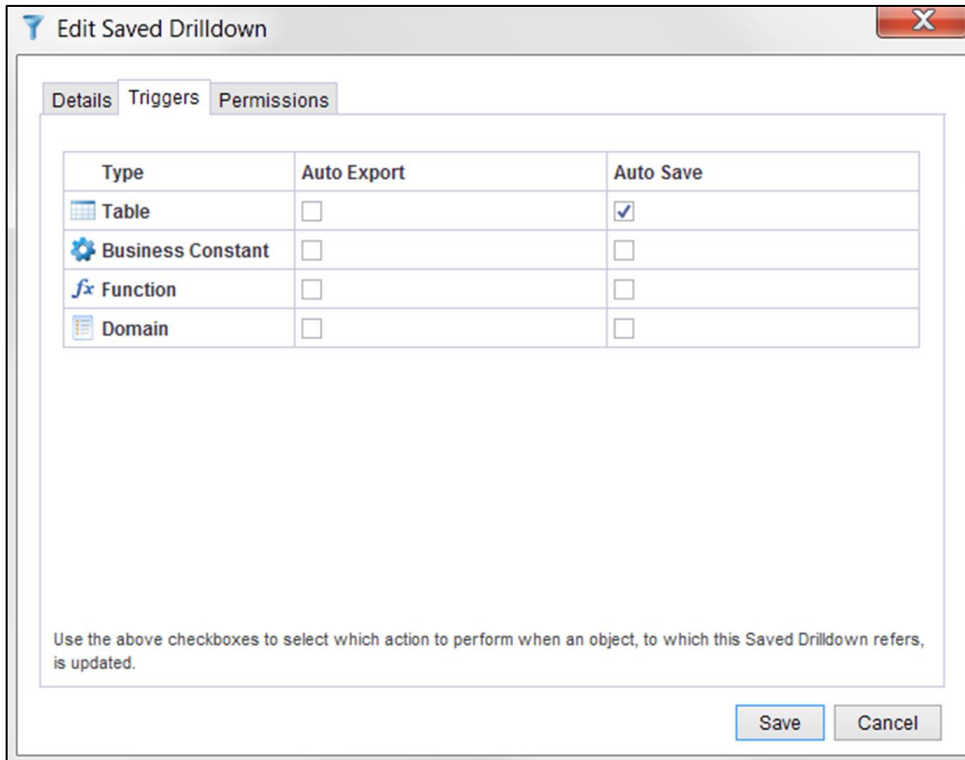
Order Header is being referenced by the drilldown, so this setting means that whenever Order Header is reloaded, the Drilldown will auto-save.

Combining this with the standard "Latest Version" novation option and "Auto Save-As-Table" setting on the saved Drilldown, means that the Drilldown will auto-save as a table whenever either Order Detail or Order Header is reloaded.

Data Transformation

The Expression editor allows for the simple, graphical specification of rules to *transform* data, in the broadest sense of the word. Existing columns can be transformed and new columns of data can be added to any drilldown.

Starting the Expression Editor

There are two ways to start the Expression Editor in order to create a new column. The first way is using the right-click menu option from the column header in the drilldown and the second is via a right-click on the Indicator panel.

Invoking the Expression Editor from a Column

Right-click on any column header in a drilldown and select **Insert Column.** This will present you with multiple options with a sub menu for each category of quick-pick transformation functions. In the example below we are looking at the Text options.
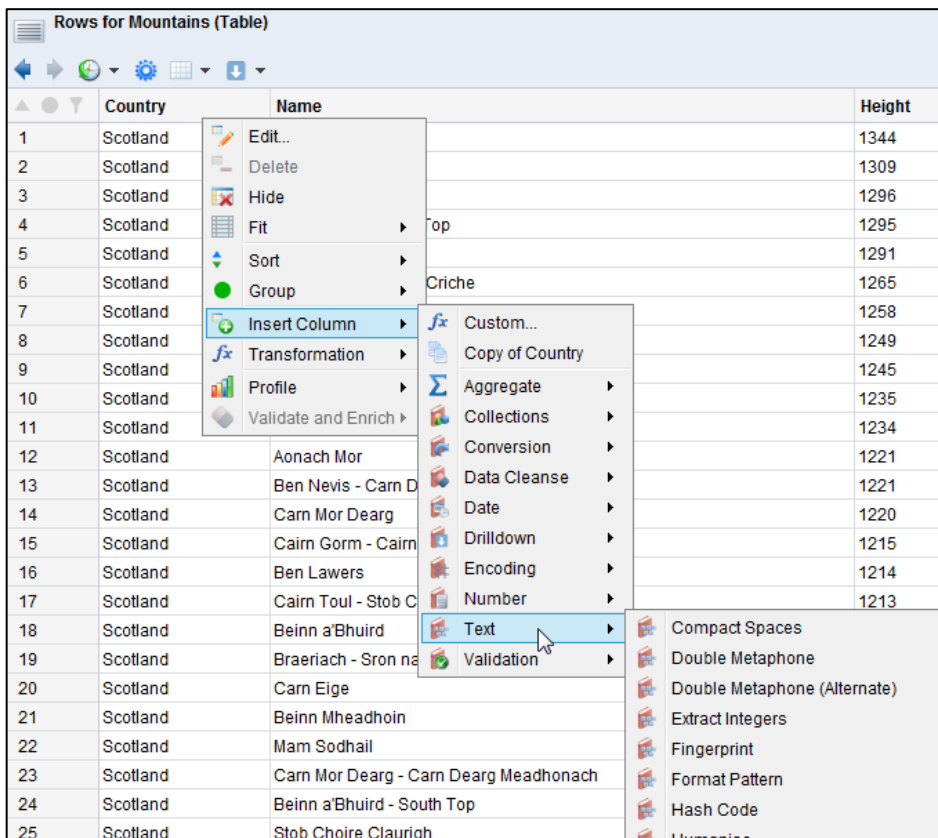
Custom Expressions

Custom allows you to use more complex transformation functions via the Expression Editor which will open when this option is selected. We will cover this on the next page.

Copying Columns and Expressions

Copy of *<Column_Name>* creates a new column which is a copy of the column you have selected. This works whether the column is an original one or an expression already created.

Name	Height	Height (Copy) *f*	Range
Ben Nevis	1344	1344	Fort William to Loch Treig and Loch Leven
Ben Macdui	1309	1309	The Cairngorms
Braeriach	1296	1296	The Cairngorms
Ben Macdui - North Top	1295	1295	The Cairngorms
Cairn Toul	1291	1291	The Cairngorms
Braeriach - Carn na Criche	1265	1265	The Cairngorms
Sgor an Lochain Uaine	1258	1258	The Cairngorms

Expression Editor Layout

The Expression Editor allows flexible and powerful data transformations to be edited or created visually. The following diagram illustrates the key features of the Expression Editor dialog:

Auto Arrange
Automatically lays out the functions and scales the expression to the visible desktop size.

Drag
When selected, the desktop view will move when an empty part of it is dragged.

Selection Box
When selected, dragging on the desktop will control a selection box.

Delete, Cut, Copy & Paste
Basic editing functions that are applied to selected functions or connections.

Function Library
Allows the list of available functions to be filtered by library

Function Search Field
Allows the list of available functions to be dynamically filtered by those containing text typed in this field. This is only displayed when the **All Functions** category is selected.

Available Functions List
Functions are added to the function desktop by double-clicking items on this list, or dragging them onto the desktop.

Function Desktop
This area displays the functions that make up the expression currently being created or edited. The function desktop represents a view of a large area that can be zoomed (using the mouse wheel or the **Auto Arrange** and **Zoom to All** toolbar buttons) and moved around by dragging an empty area when the **Drag & Zoom** toolbar button is enabled.

Function Box
Represents a single function. A function box can be individually-named and its inputs can be constant values, columns or the result of other functions. A red border around a function indicates that it is not complete (e.g. missing input values). A yellow title bar indicates that it is selected. A blue title bar indicates that it is unselected.

Column Pick-List
A list of the columns that can be used as input into functions or conditions in this expression.

Overview Navigator
Allows you to see where you are currently looking on the expression editor

Expression Editor Examples

Transforming a column and creating a new custom transformed column will follow the same process, so the examples given below will concentrate on creating new columns.

The following section will illustrate how to build transformations using examples based upon a simple table. The example table we will use is Mountains which contains information about all mountains in the United Kingdom, and consists of 4 columns:

|	
Country	the UK country in which the mountain is located
Name	the name of the mountain
Height	the height of the mountain (in metres)
Range	the name of the mountain range in which the mountain is located

A Simple Calculation Expression

Objective: Create a column that contains the height of the mountain in feet

Open the drilldown which shows the rows of the Mountains table. The height of the mountains in this table is given in metres.

Creating a New Column with a Custom Expression

Right-click on the **Height** column and from the drop down menu, select **Insert Column > Custom**. The Expression Editor dialog will open, with the title New Transformed Column with Custom Rule.
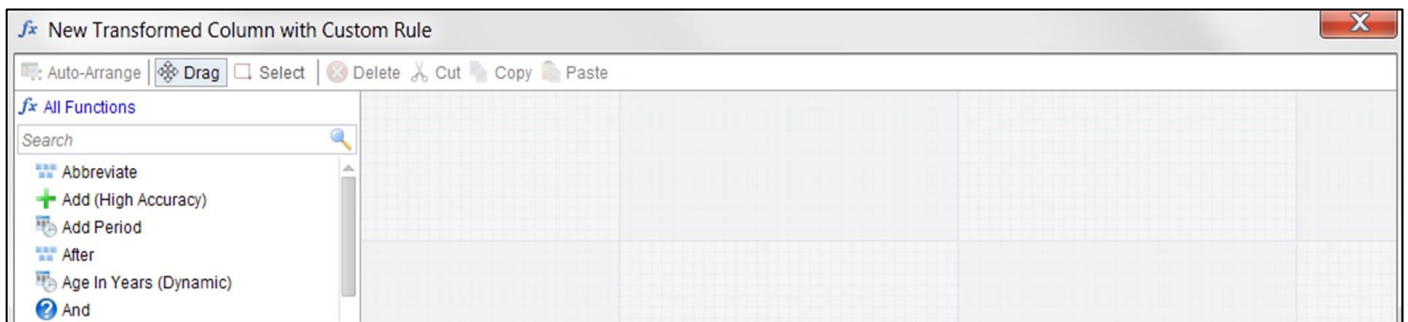

We will define the transformation rule then name the new column which contains it.

Note: In **Settings > My Preferences** the user can choose to change the view of the expression editor to be like the old behaviour, you will be able to zoom in and out to look around instead of having the new overview navigator in the bottom right corner.

Finding the Required Function

Select the **Calculate** function in the function search field. This can be found by typing into the function search field **cal**. The function search field will quickly filter the full function list to only show the functions whose names contain **cal**:
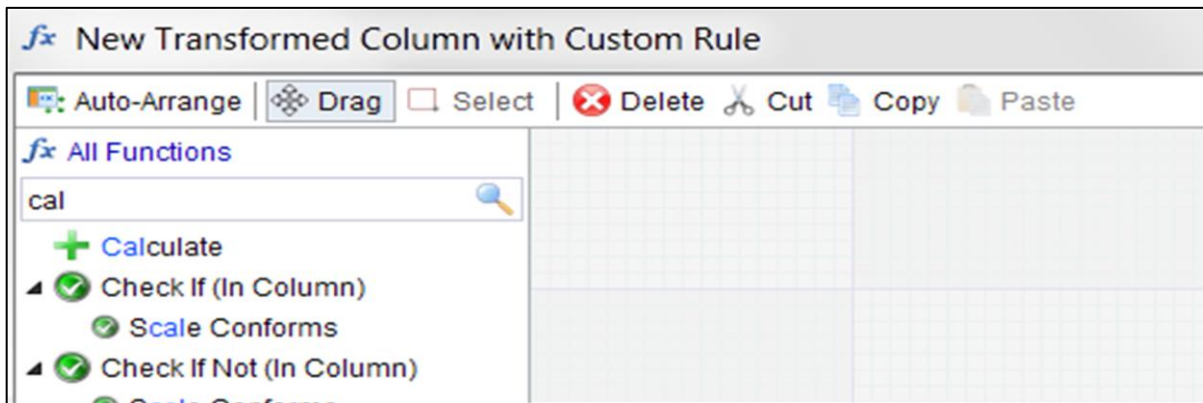
Applying a Calculate function

Double-click **Calculate** in the list to add the function to the function desktop on the right. It is also possible to drag the function from the list onto the function desktop. In **Settings > My Preferences** the user can choose whether the next function is added to the left or the right of the previous function when adding additional functions to the Expression Editor.
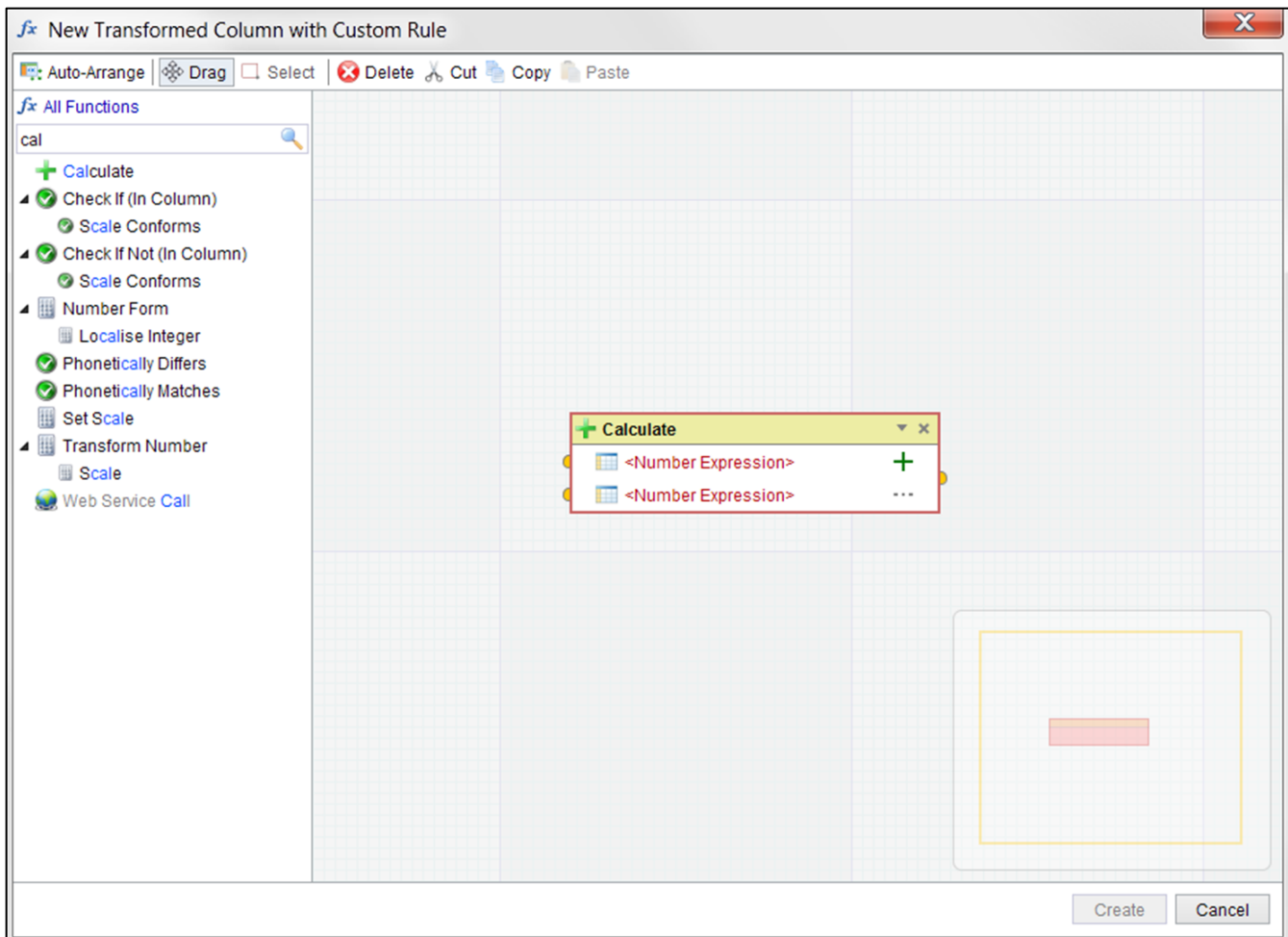
There are a few things to note about how a function is represented in the main function area:

Red Border on Functions or Function Input Fields

The red border around a function indicates that it is not yet complete. In this case, all of the input parameters are both mandatory and empty i.e. empty values or they are not connected to another function.
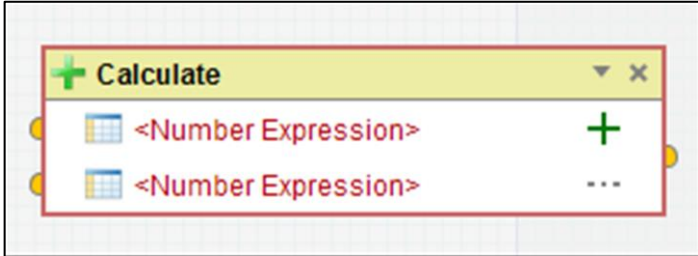

Editing the Name and Description of a Function

Clicking on the small downward arrow on the top-right of the function will add a name and description field, which allows the function to be given a custom name or description within the context of the current expression. Clicking the arrow again closes the name/description panel. In this example we have used the **Average** function to rename. The original function name remains along with your chosen name.
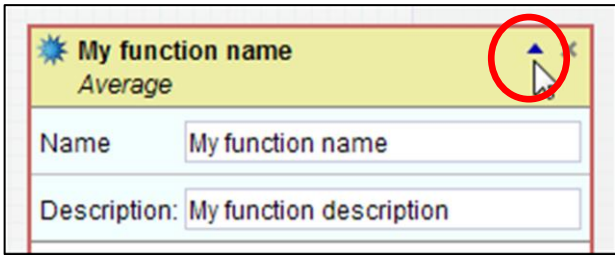
Removing a Function from the Expression

The function may be removed by clicking on the small cross located at the top-right of the function.
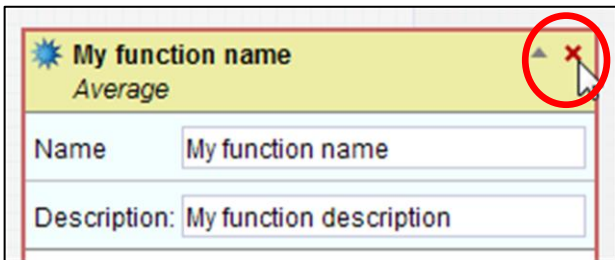
Changing the Input Type to a Function

Clicking on the icons to the left of the input fields will show a menu allowing you to change the type of the input value:

Different Input Types to a Function

Type	Description
Value	A constant text or numeric value that you may enter in the input field
Column	A value taken from another column in the drilldown
Function Input	A connection from the output node of another function to the input node of this input.
Business Constant	A Business Constant which is a value defined in the repository and referenced here by its name

Function Input Selection

By clicking the highlighted text you will open the input selection for the selected input type. For example, if the input type is set to **Business Constant** then this will be a Business Constant selector. If it was set to **Column** then this will be a list of Columns in the drilldown. Some inputs have defined values, and may not offer any options for populating the input other than the fixed set of values available in this drop down.
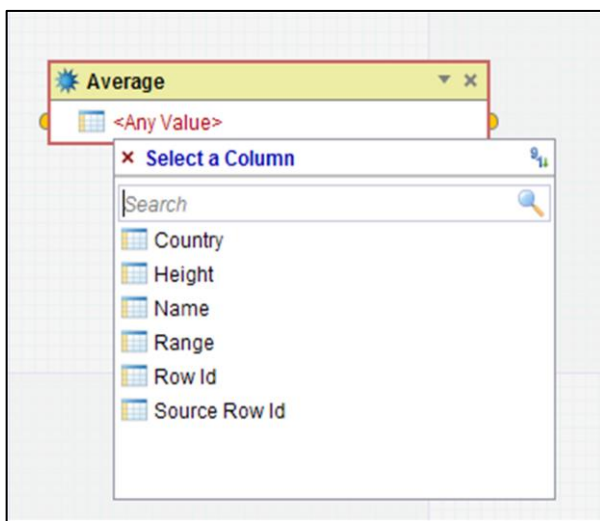
Completing the Calculate function

For this example, we need to derive the height in feet from the height in metres. This is achieved by multiplying the height in metres, which we already have access to via the existing **Height** column, by the constant numeric value 3.28

Select the **Height** column from the input selection list by clicking the highlighted text.
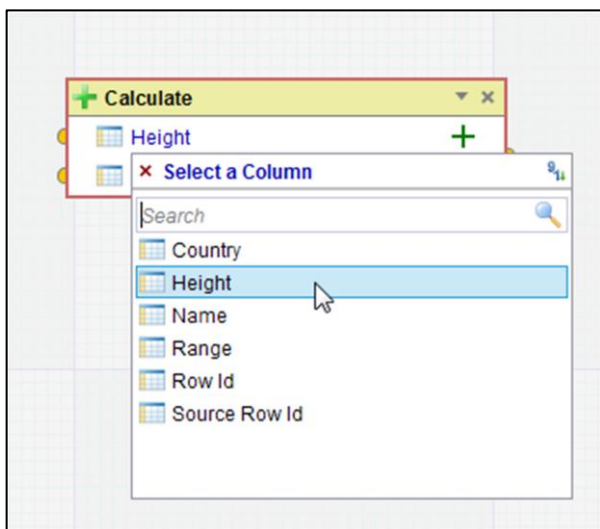
Now type the value **3.28** in the field for the second input value. Then click on the large green plus sign at the right side of the function box and select the Multiplication operator from the list.
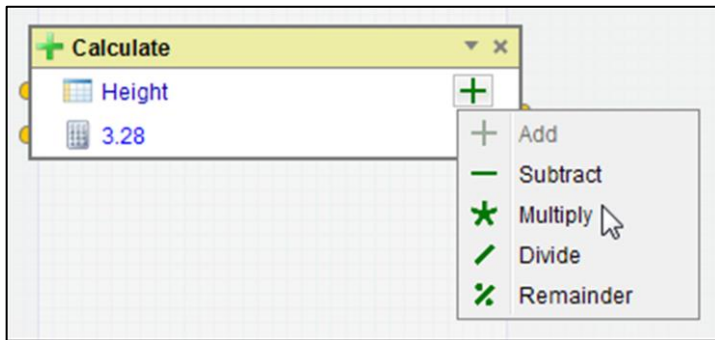
Saving the new Column

Now that the expression is complete and valid, the **Create** button is enabled in the lower right hand corner. Clicking the **Create** button will now show a dialog that allows you to enter the alias (name) and description of the new column:

The **Alias** field will initially contain a suggested unique column name. For this example, we will change it to **Height in Feet**. Clicking on the **Save** button will close the dialog.

The drilldown immediately shows the newly-created column:

	Country	Name	ƒHeight in Feet	Height	Range
1	Scotland	Ben Nevis	4408.32	1344	Fort William to Loch Treig ar
2	Scotland	Ben Macdui	4293.5199999999995	1309	The Cairngorms
3	Scotland	Braeriach	4250.88	1296	The Cairngorms
4	Scotland	Ben Macdui - North Top	4247.599999999999	1295	The Cairngorms
5	Scotland	Cairn Toul	4234.48	1291	The Cairngorms
6	Scotland	Braeriach - Carn na Criche	4149.2	1265	The Cairngorms
7	Scotland	Sgor an Lochain Uaine	4126.24	1258	The Cairngorms
8	Scotland	Ben Macdui - Stob Coire Sputan Dearg	4096.719999999999	1249	The Cairngorms
9	Scotland	Cairn Gorm	4083.6	1245	The Cairngorms
10	Scotland	Braeriach - Stob Coire an Lochain	4050.7999999999997	1235	The Cairngorms
11	Scotland	Aonach Beag	4047.52	1234	Fort William to Loch Treig ar
12	Scotland	Aonach Mor	4004.8799999999997	1221	Fort William to Loch Treig ar
13	Scotland	Ben Nevis - Carn Dearg North	4004.8799999999997	1221	Fort William to Loch Treig ar
14	Scotland	Carn Mor Dearg	4001.6	1220	Fort William to Loch Treig ar
15	Scotland	Cairn Gorm - Cairn Lochan	3985.2	1215	The Cairngorms
16	Scotland	Ben Lawers	3981.9199999999996	1214	Glen Lyon to Glen Dochart a
17	Scotland	Cairn Toul - Stob Coire an t-Saighdeir	3978.64	1213	The Cairngorms
18	Scotland	Beinn a'Bhuird	3926.16	1197	The Cairngorms
19	Scotland	Braeriach - Sron na Lairige	3883.52	1184	The Cairngorms
20	Scotland	Carn Eige	3880.24	1183	Loch Duich to Cannich
21	Scotland	Beinn Mheadhoin	3876.9599999999996	1182	The Cairngorms
22	Scotland	Mam Sodhail	3873.60	1181	Loch Duich to Cannich

f Rows for Mountains (Table)   5 of 7 Columns   2,367 Rows

Combining functions

Objective: Modify the **Height in Feet** column to show only 2 decimal places

In the previous example, a simple maths expression was used to convert a mountain's height in metres to feet. The data produced by the new column is accurate but appears ugly and inconsistent due to the varying number of decimal places representing the number.

In this example, we will modify the **Height in Feet** column to show a more pleasing representation of the numeric value.

Editing an Expression Column

Right-click on the header for the **Height in Feet** column and select the **Transformation > Edit** menu item:
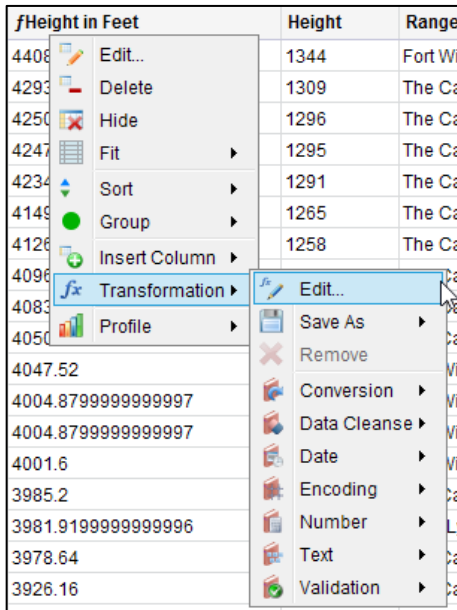
The Expression Editor dialog will now appear with the title **Edit Transformation for 'Height in Feet'** and will show the expression that was created in Example 1 above.

Adding Functions to an Expression

Search for the **Set Scale** function (e.g. type *set scale* in the function list search field), locate it on the function list and drag it to the right-hand side of the existing **Calculate** function:

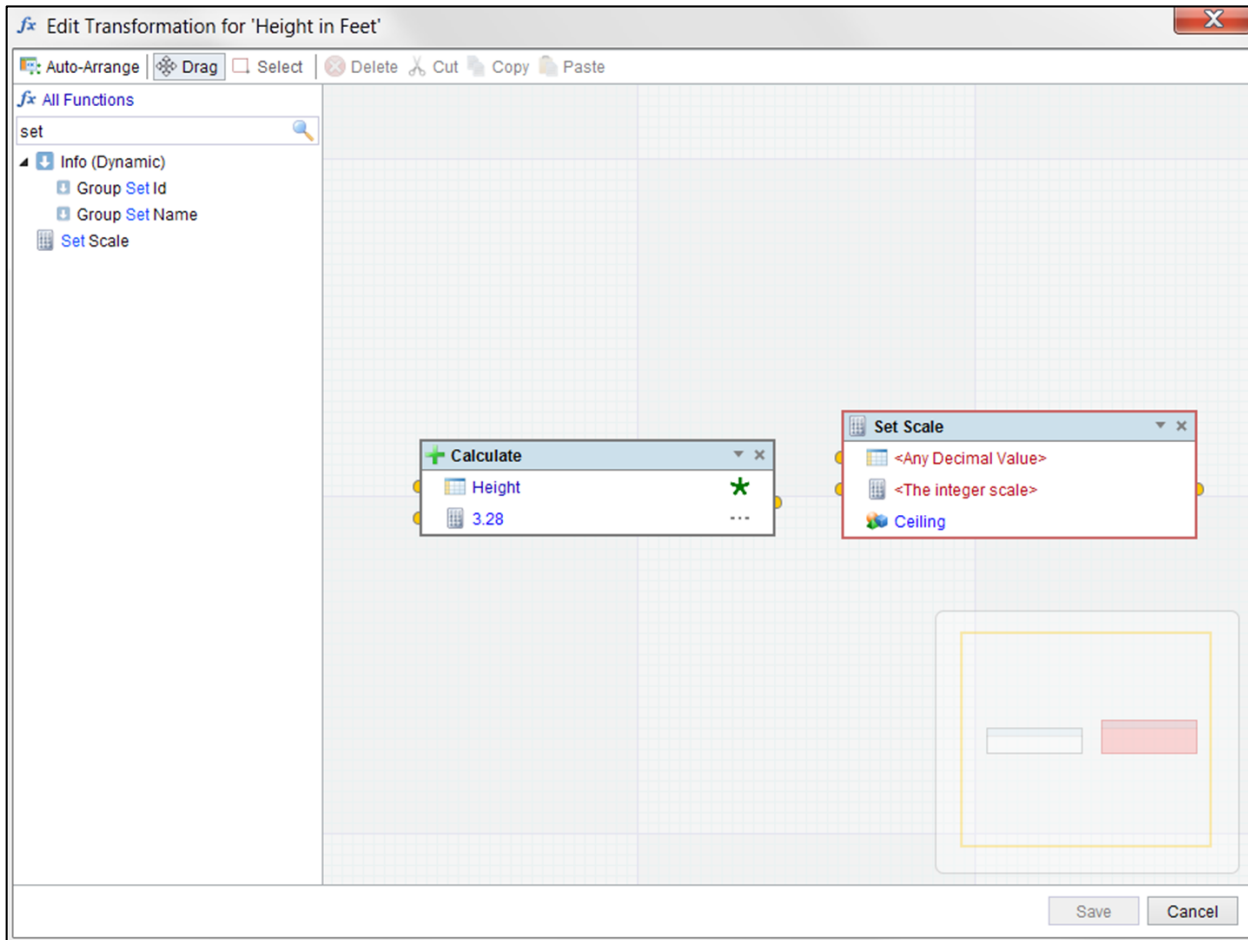Note: Option for the new function to go to the left or the right of the previous one when using double-click to add a function can be set in **Settings > My Preference**

Connecting Functions Together

Click on the output node on the right side of the **Calculate** function shown above as the small orange semi-circle on the right-hand side of the function box; then drag a line by holding down the left mouse button from this node to the first input node of the new **Set Scale** function. As the line is dragging, it will show as a blue bold line. When the line is over a node to which it can be connected, it will show as a bold green line. When the line has been connected, it will be shown as grey. The two connected boxes are now green in colour:
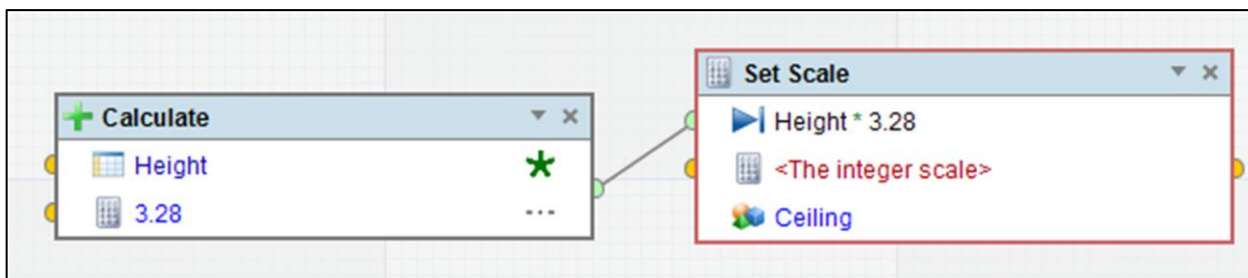
The first input node of the **Set Scale** input value now indicates that it is a function input i.e. its value comes from the output of another function and shows the expression that contributes to its value.

The **Set Scale** function takes a decimal value as input, and sets the scale of that value, with your choice of rounding methods, and outputs a new decimal or potentially integer value.
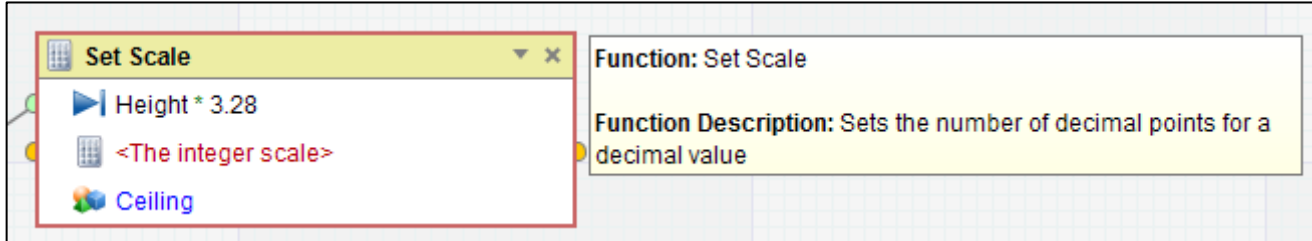
Position the mouse over the yellow function heading to see hover over help.

For this example, we will use 2 decimal places and will round half-down.

Type **2** into the *Integer Scale* field of the **Set Scale** function and set the rounding method by clicking on the blue highlighted text. Set it to **Half Down**.
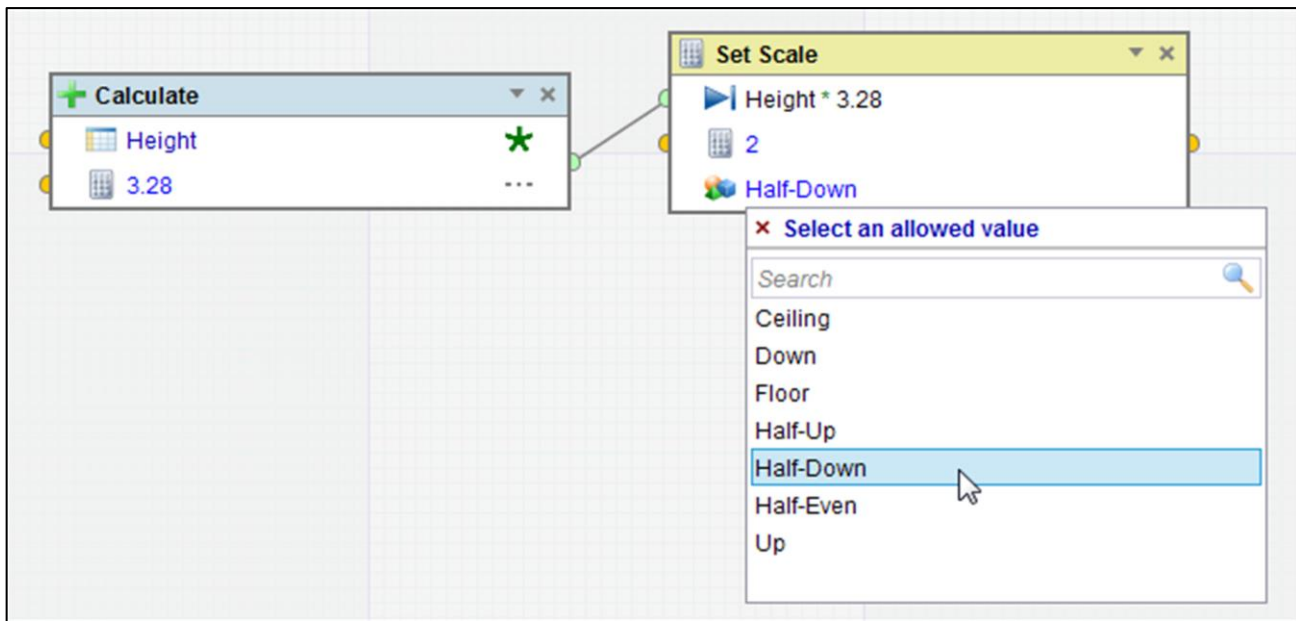
Now that the **Set Scale** function has all input values populated, its border changes from red to grey to indicate that it is complete and ready to **Save.**

Saving an Edited Expression Column

The expression is complete and therefore the **Save** button is enabled in the lower right hand corner.

Clicking the **Save** button will close the Expression Editor dialog and the drilldown will now show the result of the updated expression:

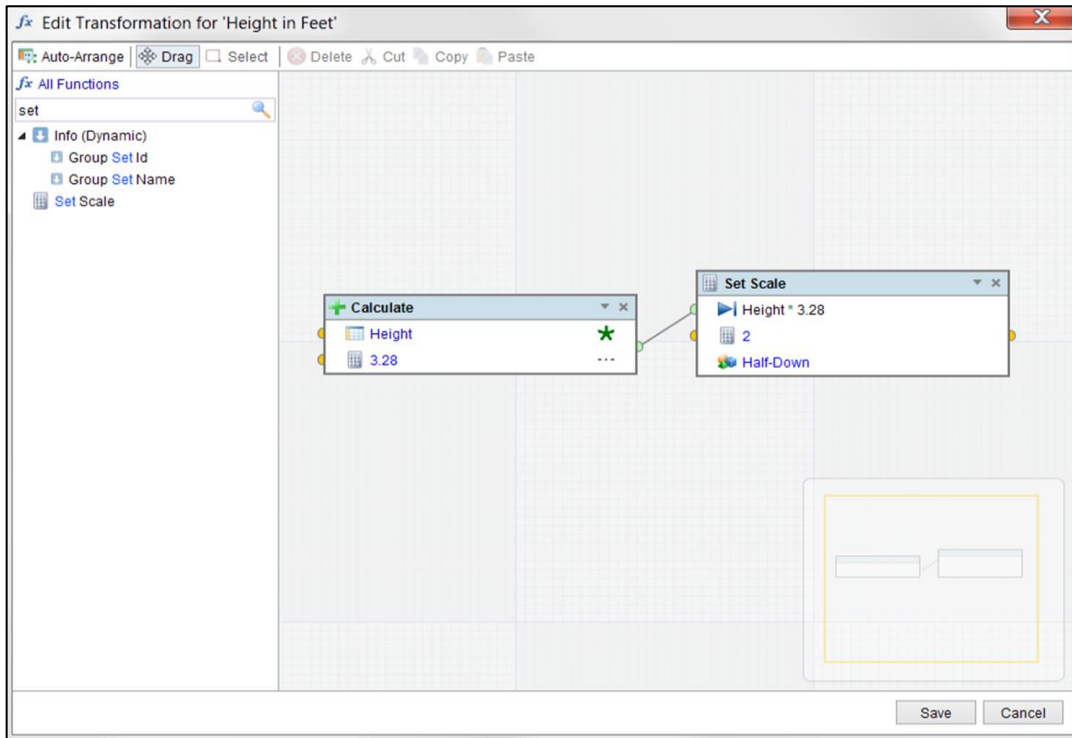|	Country	Name	ƒHeight in Feet	Height	Range
1	Scotland	Ben Nevis	4408.32	1344	Fort William to Loch Treig and Loch Leven
2	Scotland	Ben Macdui	4293.52	1309	The Cairngorms
3	Scotland	Braeriach	4250.88	1296	The Cairngorms
4	Scotland	Ben Macdui - North Top	4247.6	1295	The Cairngorms
5	Scotland	Cairn Toul	4234.48	1291	The Cairngorms
6	Scotland	Braeriach - Carn na Criche	4149.2	1265	The Cairngorms
7	Scotland	Sgor an Lochain Uaine	4126.24	1258	The Cairngorms
8	Scotland	Ben Macdui - Stob Coire Sputan Dearg	4096.72	1249	The Cairngorms
9	Scotland	Cairn Gorm	4083.6	1245	The Cairngorms
10	Scotland	Braeriach - Stob Coire an Lochain	4050.8	1235	The Cairngorms
11	Scotland	Aonach Beag	4047.52	1234	Fort William to Loch Treig and Loch Leven
12	Scotland	Aonach Mor	4004.88	1221	Fort William to Loch Treig and Loch Leven
13	Scotland	Ben Nevis - Carn Dearg North	4004.88	1221	Fort William to Loch Treig and Loch Leven
14	Scotland	Carn Mor Dearg	4001.6	1220	Fort William to Loch Treig and Loch Leven
15	Scotland	Cairn Gorm - Cairn Lochan	3985.2	1215	The Cairngorms
16	Scotland	Ben Lawers	3981.92	1214	Glen Lyon to Glen Dochart and Loch Tay
17	Scotland	Cairn Toul - Stob Coire an t-Saighdeir	3978.64	1213	The Cairngorms
18	Scotland	Beinn a'Bhuird	3926.16	1197	The Cairngorms
19	Scotland	Braeriach - Sron na Lairige	3883.52	1184	The Cairngorms
20	Scotland	Carn Eige	3880.24	1183	Loch Duich to Cannich
21	Scotland	Beinn Mheadhoin	3876.96	1182	The Cairngorms

Using Conditional Logic

Objective: Create a column that categorises a mountain's height

In this example, our objective is to create a new column that indicates which of the following three height ranges applies to a particular mountain:

Low   - less than 400 metres
Medium   - greater than or equal to 400 metres and less than 1000 metres
High   - greater than or equal to 1000m

In order to achieve this, we need to evaluate an expression conditionally, by using the **Decision** function.

Create a Decision function

As before, we right-click on the header for the **Height** column in the Mountains Rows drilldown and select the **Insert Column > Custom** which will open the Expression Editor dialog.

As an alternative to searching for a function by name, it is possible to click on the function category drop-down and, in this example, selecting the function category **Conditional**. This action will filter the function list to show only those functions in the selected category:
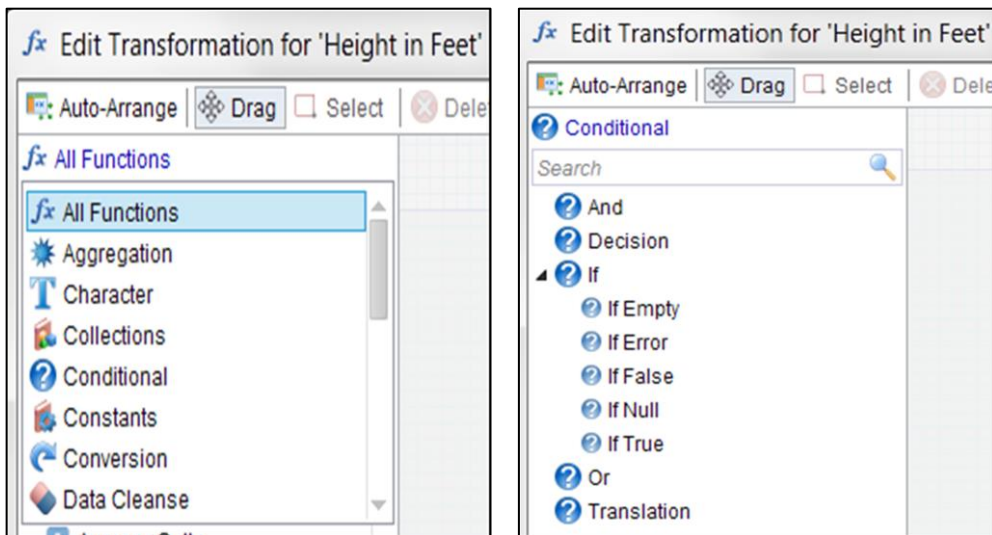
Double-clicking the **Decision** function will place the new function on the function desktop:
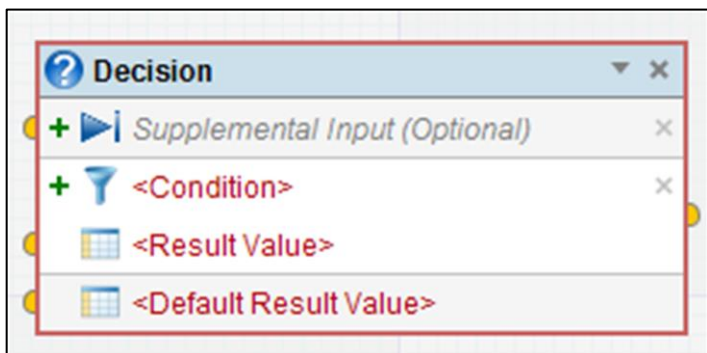
The **Decision** function is structured as follows:

Optional Supplemental Input

A supplemental input allows you to use values derived in the current expression, so must be connected to the output of another function in the same expression. Once connected, the parameter may be treated as an input column when defining the conditions of the Decision (see below).

Clicking on the small green plus icon on any supplemental input will add another one to the list. Clicking on the small cross icon will remove the associated supplemental input.

Condition/Result pairs

During the evaluation of the Decision function, true/false conditions are evaluated in order, i.e. from top to bottom as they appear in the function box and the first condition to evaluate to true will lead to its result value being used as the output value for the function.

In the diagram above, the condition is still undefined.
Clicking on the "<Condition>" label will open a condition editor at the bottom of the Expression Editor dialog.

Condition/Result pairs may be created and deleted in the same way as supplemental inputs, i.e. by using the plus and cross icons.

The order of Condition/Result pairs may be changed by dragging them from one position to another.

The Default Result value

If none of the conditions is true, the output value of the Decision function will be the value specified in the Default Result.
All of the above concepts will be shown in the remainder of this worked example.

Create values for the Low and High thresholds

In this example, we will create two **Constant Value** functions, which will represent the maximum height of a **Low** range and the minimum height of a **Tall** range.

Change the function category to **All Functions**, search for **Constant Value** and drag the Constant Value function across to the function desktop twice. Change the name of each constant function box to **Low Threshold** and **High Threshold** – this is done by clicking the small down arrow which opens the show name/description panel:
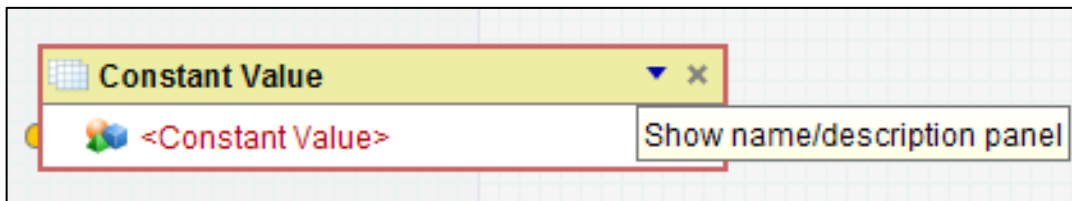
Type in the values **400** in the Low Threshold and **1000** in the High Threshold:
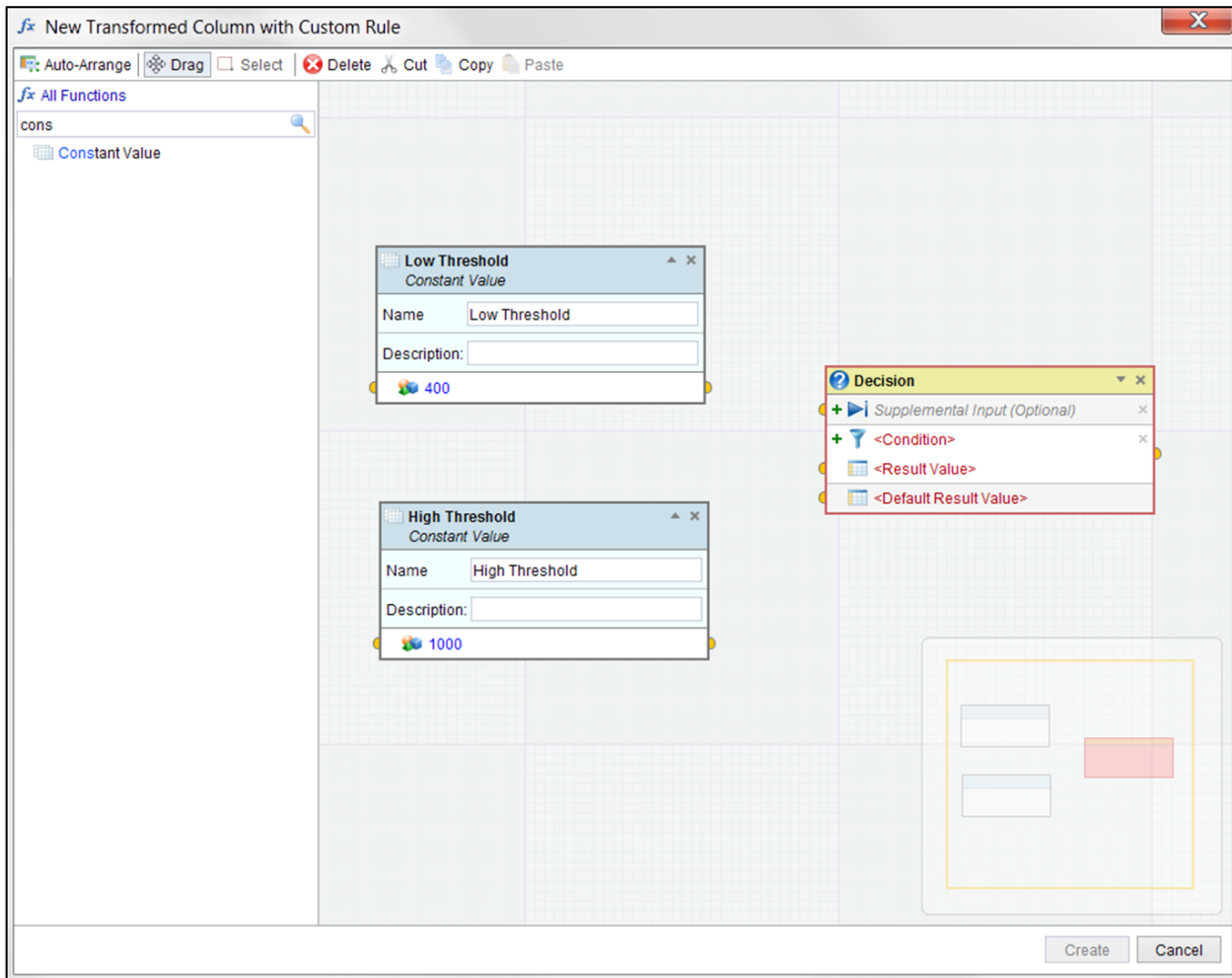
In order for the Constant Values to be seen as integers, we need to pass it through the Convert To Integer function before adding them to the decision function. It can be found by searching for **Convert Int**. Add two, one for each threshold.
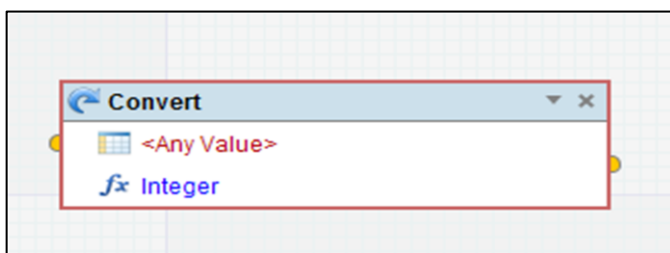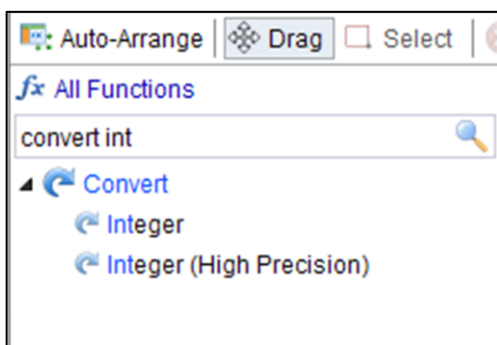
We now need to create an additional condition in the Decision function by clicking on the small green *plus* icon to the left of the initial condition and an additional Supplemental input again by clicking on the appropriate green *plus* icon.

Next we connect the **Low Threshold** value to a **Convert To Integer** function and then to the first supplemental parameter. Do the same with the **High Threshold** value, connecting it to the second **Convert To Integer** function and then to the second supplemental parameter:

Define the Conditions and Results

We have defined a **Low** mountain as one whose height is less than 400 metres, and a **High** mountain as one whose height is greater than or equal to 1000 metres. Anything else must be **Medium**.

Our Decision function therefore needs to evaluate the following conditions in order:

If Height is less than 400 then the result is **Low**
If Height is greater than or equal to 1000 then the result is **High**
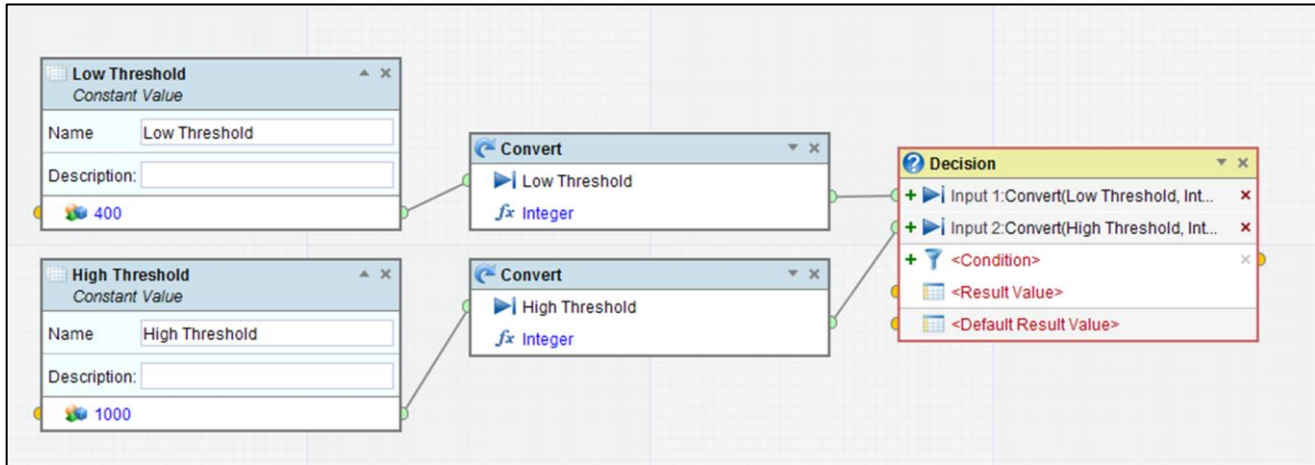If neither condition is true then the result is **Medium**

The method of creating conditions is identical to that used in the Drilldown Manager.

Now we need to click on the **<Condition>** Decision function to open the filter condition editor at the bottom of the expression editor window. Either click Add or Double-click the <Click here to create a new filter condition>:

We then use the condition editor to define the first condition as follows:
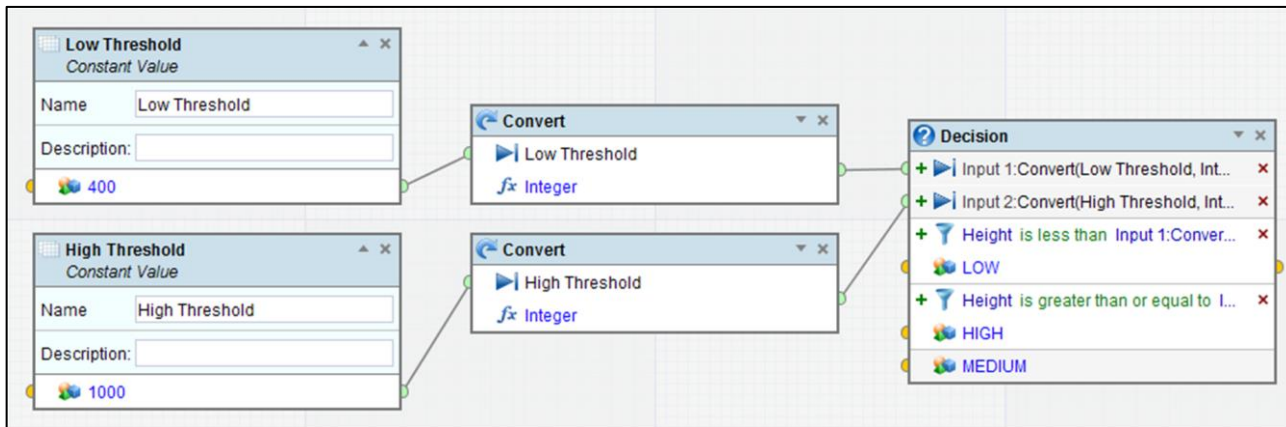

After clicking on the second condition, we define it as follows:
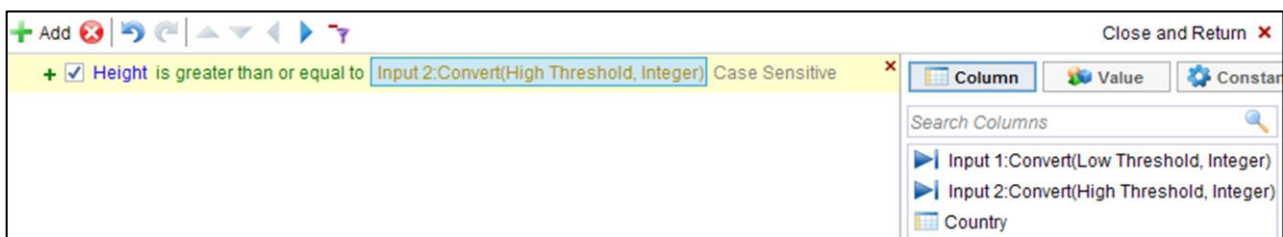
Note the cross icon in the top-right of the condition editor that can be used to hide the editor.

Next we fill in the results that we want to see for each of the two conditions and the default result for the Decision:

Save the Expression Column

The expression is complete so we click **Create** and then give it a name of **Height Range**.

This results in the following drilldown. Note that we have sorted by Name to show the different Heights and their corresponding **Height Range.**

Row Id	Country	Name	ƒ Height in Feet	Height	ƒ Height Range	Range
1852	Scotland	A'Bheinn Bhan	1564.56	477	MEDIUM	Sunart and Ardgour
485	Scotland	A'Bhuidheanach Bheag	3070.08	936	MEDIUM	Loch Ericht to Glen Tromie and Glen Garry
517	Scotland	A'Bhuidheanach Bheag - Glas Mheall Mor	3043.84	928	MEDIUM	Loch Ericht to Glen Tromie and Glen Garry
587	Scotland	A'Bhuidheanach Bheag - Meall a'Chaora...	3004.48	916	MEDIUM	Loch Ericht to Glen Tromie and Glen Garry
273	Scotland	A'Chailleach	3270.16	997	MEDIUM	The Fannaichs
507	Scotland	A'Chailleach	3050.4	930	MEDIUM	Glen Albyn and the Monadh Liath
490	Scotland	A'Chailleach - Toman Coinich	3066.8	935	MEDIUM	The Fannaichs
61	Scotland	A'Chralaig	3673.6	1120	HIGH	Glen Affric to Glen Moriston
438	Scotland	A'Chralaig - A'Chioch	3106.16	947	MEDIUM	Glen Affric to Glen Moriston
235	Scotland	A'Chralaig - Stob Coire na Cralaig	3306.24	1008	HIGH	Glen Affric to Glen Moriston
1772	Scotland	A'Chruach	1685.92	514	MEDIUM	Arran and Holy Island
2086	Scotland	A'Chruach	1200.48	366	LOW	Oban to Loch Fyne
573	Scotland	A'Ghlas-bheinn	3011.04	918	MEDIUM	Loch Duich to Cannich
368	Scotland	A'Mhaighdean	3171.76	967	MEDIUM	Loch Maree to Loch Broom
341	Scotland	A'Mharconaich	3198	975	MEDIUM	Loch Ericht to Glen Tromie and Glen Garry
2218	England	Abberley Hill	934.8	285	LOW	West Gloucestershire and Hereford & Worcester
1919	Wales	Aberedw Hill	1479.28	451	MEDIUM	Welshpool to Hay-on-Wye
2239	England	Aconbury Hill	905.28	276	LOW	West Gloucestershire and Hereford & Worcester
2127	Scotland	Ailsa Craig	1108.64	338	LOW	Carrick and Galloway
878	Scotland	Ainshval	2561.68	781	MEDIUM	Canna, Rum and Eigg
2348	Scotland	Airds Hill	593.68	181	LOW	Loch Linnhe to Loch Etive
1402	Scotland	Alhang	2105.76	642	MEDIUM	The Glenkens to Annandale
867	England	Allen Crags	2574.8	785	MEDIUM	Central and Western Fells

Getting Function Help

A full list of available categories and functions are available from within Experian Pandora. From the top level of the **explorer**, select **Libraries,** right click on the **Standard Function Libraries** item and select the drilldown for **All Standard Functions**. This produces a drilldown which has the same information in drilldown format that is in this section of the manual. You can also view help for individual functions from within the Expression Editor itself by hovering over the functions in the pick list and the individual function parameters once a function has been dragged on the Expression Editor workspace. Like all drilldowns in Experian Pandora, the functions drilldown can be filtered to find the information you need. Hovering over any function produces a detailed popup with all available information for that function.

Date & Time formats for the Format Date Function

Using the **Format Date** function, date and time formats are specified by *date and time pattern* strings. Within date and time pattern strings, unquoted letters from **A** to **Z** and from **a** to **z** are interpreted as pattern letters representing the components of a date or time string. Text can be quoted using single quotes (') to avoid interpretation. All other characters are not interpreted and are simply copied to the output during formatting or matched against the input string during parsing. The following pattern letters are defined (all other characters from **A** to **Z** and from **a** to **z** are reserved):

Letter	Date or Time Component	Presentation	Examples
G	Era designator	Text	AD
y	Year	Year	1996; 96
M	Month in year	Month	July; Jul; 07
w	Week in year	Number	27
W	Week in month	Number	2
D	Day in year	Number	189
d	Day in month	Number	10
F	Day of week in month	Number	2 (Saturday is "1")
E	Day in week	Text	Tuesday; Tue
a	Am/pm marker	Text	PM
H	Hour in day (00-23), 2 digits	Number	01
k	Hour in day (0-23)	Number	1
K	Hour in am/pm (0-11)	Number	0
h	Hour in am/pm (12-11)	Number	12
m	Minute in hour	Number	30
s	Second in minute	Number	55
S	Millisecond	Number	978
z	Time zone	General time zone	Pacific Standard Time; PST; GMT-08:00
Z	Time zone	RFC 822 time zone	-0800

Pattern letters are usually repeated, as their number determines the exact presentation:

Pattern	Description
Text	For formatting, if the number of pattern letters is 4 or more, the full form is used; otherwise a short or abbreviated form is used if available.
Number	For formatting, the number of pattern letters is the minimum number of digits, and shorter numbers are zero-padded to this amount.
Year	For formatting, if the number of pattern letters is 2, the year is truncated to 2 digits; otherwise it is interpreted as a number.
Month	If the number of pattern letters is 3 or more, the month is interpreted as text; otherwise, it is interpreted as a number.

Formats for the *Format String* Function

The **Format String** function transforms input values according to a format specification.

The input values can be of any type, although the possible formats are dependent on data type of each value. The format specification can apply to all the input values and/or specify a format for individual input values. Each format specification is made up of several components, some of which are optional.

Format string : [text] **Format** [text] **Input_string** [text] {Input string [text]}

Format : %[**input_index**$][**flags**][**width**][.**precision**] **conversion**

Optional components

Component	Description
Text	Any text you wish. See example below.
input_index	An integer corresponding to the position of the input string to which this format will be applied. The first input string is referenced by **1$**, the second by **2$**, etc.
Flags	A set of characters that modify the output format. The set of valid flags depends on the conversion.
Width	A non-negative integer indicating the minimum number of characters to be written to the output.
Precision	A non-negative integer usually used to restrict the number of characters. The exact behaviour depends on the conversion. Precision cannot be specified for datatypes which are used to represents dates and times.

Mandatory components

Component	Description
Conversion	Characters indicating how the argument should be formatted. The set of valid conversions for a given input string depends on its data type. For datatypes which are used to represents dates and times the conversion is the letter **T** followed by the format to be used.

Conversions

These are divided into the following categories:

Conversion	Description
Character	May be applied to any data type
Numeric - Integer	May be applied to any number resulting in an integer
Numeric – Decimal	May be applied to any number resulting in a decimal
Date, Time, Timestamp	May be applied to any date/time value
Percent	A number followed by the percent sign %
Line Separator	The platform specific line separator

The following table summarizes the supported conversions. Conversions denoted by an upper-case character (i.e. **B, H, S, C, X, E, G, A** and **T**) are the same as those for the corresponding lower-case conversion characters except that the result is converted to upper case.

Conversion	Argument Category	Description
B or **b**	Character	If the input is NULL or **false**, then the result is **false**. Otherwise, the result is **true**.
H or **h**	Character	If the input is NULL, then the result is NULL. Otherwise, the result is the Hexadecimal representation of the input
S or **s**	Character	If the input is NULL, then the result is NULL. Otherwise, the result is the string representation of the input
C or **c**	Character	The result is a Unicode character
D or **d**	Integer	The result is formatted as an integer
o	Integer	The result is formatted as an octal integer
X or **x**	Integer	The result is formatted as a hexadecimal integer
E or **e**	Decimal	The result is formatted as a number in scientific notation
f	Decimal	The result is formatted as number
G or **g**	Decimal	The result is formatted using scientific notation or decimal format, depending on the precision and the value after rounding.
A or **a**	Decimal	The result is formatted as a hexadecimal floating-point number with a significand and an exponent
T or **t**	Date/time	Prefix for date and time conversion characters. See Date/Time Conversions below.
%	Percent	The result is a literal **%**
n	Line separator	The result is the platform-specific line separator

Date/Time Conversions

The following date and time conversion suffix characters are defined for the **t** and **T** conversions.

The following conversion characters are used for formatting times:

Conversion	Description
H	Hour of the day for the 24-hour clock, formatted as two digits with a leading zero as necessary i.e. 00 - 23.
I	Hour for the 12-hour clock, formatted as two digits with a leading zero as necessary, i.e. 00 - 11.
k	Hour of the day for the 24-hour clock, i.e. 0 - 23.
l	Hour for the 12-hour clock, i.e. 0 - 11.
M	Minute within the hour formatted as two digits with a leading zero as necessary, i.e. 00 - 59.
S	Seconds within the minute, formatted as two digits with a leading zero as necessary, i.e. 00 - 59 ("60" is a special value required to support leap seconds).
L	Millisecond within the second formatted as three digits with leading zeros as necessary, i.e. 000 - 999.
N	Nanosecond within the second, formatted as nine digits with leading zeros as necessary, i.e. 000000000 - 999999999.
p	Language specific morning or afternoon marker in lower case, e.g. **am** or **pm**. Use of the conversion prefix **T** forces this output to upper case.
z	RFC 822 style numeric time zone offset from GMT, e.g. -0800.
Z	A string representing the abbreviation for the time zone.
s	Seconds since the beginning of the epoch starting at 1 January 1970 00:00:00 UTC
Q	Milliseconds since the beginning of the epoch starting at 1 January 1970 00:00:00 UTC

The following conversion characters are used for formatting dates:

Conversion	Description
B	Language specific full month name, e.g. **January**, **February**.
h	Same as **b**.
A	Language specific full name of the day of the week, e.g. **Sunday**, **Monday**
a	Language specific short name of the day of the week, e.g. **Sun**, **Mon**
C	Four-digit year divided by 100, formatted as two digits with leading zero as necessary, i.e. 00 - 99
Y	Year, formatted as at least four digits with leading zeros as necessary, e.g. 0092 equals 92 CE for the Gregorian calendar.
y	Last two digits of the year, formatted with leading zeros as necessary, i.e. 00 - 99.
j	Day of year, formatted as three digits with leading zeros as necessary, e.g. 001 - 366 for the Gregorian calendar.
m	Month, formatted as two digits with leading zeros as necessary, i.e. 01 - 13.
d	Day of month, formatted as two digits with leading zeros as necessary, i.e. 01 - 31
e	Day of month, formatted as two digits, i.e. 1 - 31.

The following conversion characters are used for formatting common date/time compositions.

Conversion	Description
R	Time formatted for the 24-hour clock as **%tH:%tM**
T	Time formatted for the 24-hour clock as **%tH:%tM:%tS**
r	Time formatted for the 12-hour clock as **%tI:%tM:%tS %Tp** The location of the morning or afternoon marker (**%Tp**) may be language dependent.
D	Date formatted as **%tm/%td/%ty**
F	ISO 8601 complete date formatted as **%tY-%tm-%td**
c	Date and time formatted as **%ta %tb %td %tT %tZ %tY** e.g. **Sun Jul 20 16:17:00 EDT 1969**

Flags

The following table summarizes the supported flags. *y* means the flag is supported for the indicated argument types.

Flag	Character	Integer	Decimal	Date/Time	Description
-	Yes	Yes	Yes	Yes	The result will be left-justified.
+		Yes[2]	Yes		The result will always include a sign
<space>		Yes[2]	Yes		The result will include a leading space for positive values
0		Yes	Yes		The result will be zero-padded
,		Yes[1]	Yes[3]		The result will include locale-specific grouping separators
(		Yes[2]	Yes[3]		The result will enclose negative numbers in parentheses

[1] **d** conversion only

[2] **d, o, x,** and **X** conversions applied to Integers

[3] **e, E, f, g,** and **G** conversions only

Width

The width is the <u>minimum</u> number of characters to be written to the output.

For the line separator conversion, width is not applicable; if it is provided, an error will result.
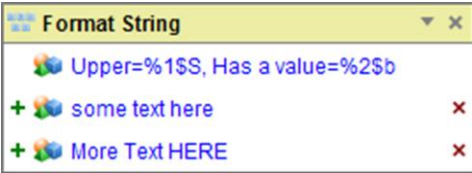
Precision

For decimal conversions **e**, **E**, and **f** the precision is the number of digits after the decimal separator. If the conversion is **g** or **G**, then the precision is the total number of digits in the resulting magnitude after rounding. If the conversion is **a** or **A**, then the precision must not be specified.

For character, integer, and date/time argument types and the percent and line separator conversions, the precision is not applicable; if a precision is provided, an error will be the result.

Examples

The following examples are simplified as they are using a literal value as input, instead either an input column or the result of some other function.

Example	Result
Format String ▼ ✕ 🔷 %S ➕ 🔷 some text hERe ✕	**SOME TEXT HERE**
Format String ▼ ✕ 🔷 Upper=%1$S, Has a value=%2$b ➕ 🔷 some text here ✕ ➕ 🔷 More Text HERE ✕	**Upper=SOME TEXT HERE, Has a value=true** *Note: the comma and text were inserted into the result simply by including them in the format.*

Operators in the Filter panel of the Drilldown Manager

Experian Pandora provides a rich array of operators for use both in filters within the Drilldown Manager, and also within conditions in the **Decision** function. Many of these will be familiar if you have used query tools before, but many are very specific to Experian Pandora, with some only possible because of the advanced database techniques employed within the Experian Pandora server.

Operator panel prompt	Filter panel text	Operator Category	Has case insensitive option?	Allows multiple values?
alphabetic	is alphabetic	Type	No	No
alphanumeric	is alphanumeric	Type	No	No
between	is in the range *<start value>* to *<end value>*	Simple	**Yes**	No
constant length	has length *<length>*	Advanced	No	No
contains	contains *<value>*	Containment	**Yes**	**Yes**
contains format	contains the format *<Format>*	Advanced	No	**Yes**
contains similar word	contains a word similar to *<Value>* edit distance *<Edit Distance>* length variation *<Variation>*	Advanced	No	No
contains soundalike	contains a word sounding like *<Value>*	Pattern	No	Yes
contains word	contains the word *<Value>*	Containment	No	Yes
date	is a date	Type	No	No
decimal	is a decimal	Type	No	No
differs	does not match regular expression *<Value>*	Advanced	**Yes**	**Yes**
empty	is empty	Simple	No	No
ends with	ends with *<Value>*	Containment	**Yes**	**Yes**
equals	is equal to *<Value>*	Simple	**Yes**	**Yes**
error	has a transform error	Advanced	No	No
false	is false	Type	No	No
formatted like	is formatted like	Pattern	No	**Yes**
greater than	is greater than	Simple	**Yes**	No
greater than or equal	is greater than or equal to	Simple	**Yes**	No
in edit distance	is within edit distance of *<Value>* edit distance *<Edit Distance>* length variation *<Variation>*	Advanced	No	No
in jaro distance	is within Jaro distance of *<Value>* edit distance *<Distance>*	Advanced	No	No
in jaro-winkler distance	is within Jaro-Winkler distance of *<Value>* edit distance *<Distance>* prefix length *<Length>* distance weight *<Weight>*	Advanced	No	No
in length range	has length within the range *<Start Value>* to *<End Value>*	Advanced	No	No
integer	is integer	Type	No	No
less than	is less than *<Value>*	Simple	**Yes**	No
less than or equal	is less than or equal to *<Value>*	Simple	**Yes**	No
matches expression	matches the regular expression *<Value>*	Advanced	**Yes**	**Yes**
money	is money	Type	No	No
not alphabetic	is not alphabetic	Type	No	No
not alphanumeric	is not alphanumeric	Type	No	No
not between	is not in the range *<Start Value>* to *<End Value>*	Simple	**Yes**	No

Operator panel prompt	Filter panel text	Operator Category	Has case insensitive option?	Allows multiple values?
not constant length	does not have the constant length *<Length>*	Advanced	No	No
not contains	does not contain *<Value>*	Containment	**Yes**	**Yes**
not contains format	does not contain the format *<Format>*	Advanced	No	**Yes**
not contains similar word	does not contain a word similar to *<Value>* edit distance *<Edit Distance>* length variation *<Variation>*	Advanced	No	No
not contains soundalike	does not contain a word sounding like *<Value>*	Pattern	No	Yes
not contains word	does not contain the word *<Value>*	Containment	No	**Yes**
not date	is not a date	Type	No	No
not decimal	Is not a decimal	Type	No	No
not empty	is not empty	Simple	No	No
not ends with	does not end with *<Value>*	Containment	**Yes**	**Yes**
not equals	is not equal to *<Value>*	Simple	**Yes**	**Yes**
not error	does not have a transform error	Advanced	No	No
not false	is not false	Type	No	No
not formatted like	is not formatted like	Pattern	No	Yes
not in edit distance	is not within edit distance of *<Value>* edit distance *<Edit Distance>* length variation *<Variation>*	Advanced	No	No
not in jaro distance	is not within Jaro distance of *<Value>* edit distance *<Distance>*	Advanced	No	No
not in jaro-winkler distance	is not within Jaro-Winkler distance of *<Value>* edit distance *<Distance>* prefix length *<Length>* distance weight *<Weight>*	Advanced	No	No
not in length range	does not have a length within the range *<Start Value>* to *<End Value>*	Advanced	No	No
not integer	is not integer	Type	No	No
not money	is not money	Type	No	No
not null	is not NULL	Type	No	No
not number	is not a number	Type	No	No
not original	is not an original value	Type	No	No
not sounds like	sounds different to *<Value>*	Pattern	No	**Yes**
not space filled	is not space filled	Type	No	No
not starts with	does not start with *<Value>*	Containment	**Yes**	**Yes**
not the error	does not have the transform error message *<Value>*	Advanced	No	No
not the warning	does not have the warning message	Advanced	No	No
not true	is not true	Type	No	No
not used	does not exist in the database	Advanced	No	No
not warning	is not a user defined warning	Advanced	No	No
null	is NULL	Type	No	No
number	is a number	Type	No	No
original	is an original value	Type	No	No
sounds like	sounds like *<Value>*	Pattern	No	**Yes**
space filled	is space filled	Type	No	No
starts with	starts with *<Value>*	Containment	**Yes**	**Yes**
the error	has the transform error message *<Value>*	Advanced	No	No

Operator panel prompt	Filter panel text	Operator Category	Has case insensitive option?	Allows multiple values?
the warning	has the warning message *<Value>*	Advanced	No	No
true	is true	Type	No	No
used	exists in the database	Advanced	No	No
warning	is a user defined warning	Advanced	No	No

Operators by applicability

In this section we view operations by specific areas of applicability, referencing them to the operators as you would see presented on the filter panels within the Experian Pandora user interface. We have also included the SQL syntax for the operators as ultimately they will end up being part of a **WHERE** or **HAVING** clause, or part of a **WHEN** clause in a **CASE** statement (implemented as a **Decision** function within the Experian Pandora user interface).

Equality & Inequality

Action	SQL	Operator Panel Prompt
Is exactly equal	*expression = expression*	equals
Is equal ignoring case	*expression* **EQ** *expression*	equals
Is equal to any one of a set of values	*expression* **IN**(*value1,value2...*)	equals
Is equal to any one of a set of values ignoring case	*expression* **IN INSENSITIVE** (*value1,value2...*)	equals
Is not equal	*expression <> expression*	not equals
Is not equal ignoring case	*expression* **NE** *expression*	not equals
Is not equal to any one of a set of values	*expression* **NOT IN** (*value1,value2...*)	not equals
Is not equal to any one of a set of values ignoring case	**NOT IN INSENSITIVE** (*value1,value2...*)	not equals

Null & Empty check

Action	SQL	Operator Panel Prompt
Is either NULL or space filled	*expression* **IS EMPTY**	empty
Is not NULL or space filled	*expression* **IS NOT EMPTY**	not empty
Is NULL	*expression* **IS NULL**	null
Is not NULL	*expression* **IS NOT NULL**	not null
Is space filled	*expression* **IS SPACES**	space filled
Is not space filled	*expression* **IS NOT SPACES**	not space filled

Special Operators

Action	SQL	Filter Operator Prompt
Value exists in the database	*expression* **IS USED VALUE**	used
Value does not exist in the database	*expression* **IS NOT USED VALUE**	not used

Range Checking

Action	SQL	Operator Panel Prompt
Is less than	*expression* **<** *expression*	less than
Is less than ignoring case	*expression* **LT** *expression*	less than
Is less than or equal to	*expression* **<=** *expression*	less than or equal
Is less than or equal to ignoring case	*expression* **LE** *expression*	less than or equal
Is greater than	*expression* **>** *expression*	greater than
Is greater than ignoring case	*expression* **GT** *expression*	greater than
Is greater than or equal to	*expression* **>=** *expression*	greater than or equal
Is greater than or equal to ignoring case	*expression* **GE** *expression*	greater than or equal
Is in a specified value range	*expression* **BETWEEN** *min* **AND** *max*	between
Is in a specified value range ignoring case	*expression* **BETWEEN INSENSITIVE** *min* **AND** *max*	between
Is not in a specified value range	*expression* **NOT BETWEEN** *min* **AND** *max*	not between
Is not in a specified value range ignoring case	*expression* **NOT BETWEEN INSENSITIVE** *min* **AND** *max*	not between

Type Checking

Action	SQL	Filter Operator Prompt
Is an alphabetic string	*expression* **IS TYPE ALPHABETIC**	alphabetic
Is not an alphabetic string	*expression* **IS NOT TYPE ALPHABETIC**	not alphabetic
Is an alphanumeric string	*expression* **IS TYPE ALPHANUMERIC**	alphanumeric
Is not an alphanumeric string	*expression* **IS NOT TYPE ALPHANUMERIC**	not alphanumeric
Is a number	*expression* **IS TYPE NUMBER**	number
Is not a number	*expression* **IS NOT TYPE NUMBER**	not number
Is a decimal number	*expression* **IS TYPE DECIMAL**	decimal
Is not a decimal number	*expression* **IS NOT TYPE DECIMAL**	not decimal
Is an integer number	*expression* **IS TYPE INTEGER**	integer
Is not integer	*expression* **IS NOT TYPE INTEGER**	not integer
Is a monetary value	*expression* **IS TYPE MONEY**	money
Is not a monetary value	*expression* **IS NOT TYPE MONEY**	not money
Is a date/time value	*expression* **IS TYPE DATE**	date
Is not a date/time value	*expression* **IS NOT TYPE DATE**	not date
Is a word	*expression* **IS TYPE TOKEN**	token
Is not a word	*expression* **IS NOT TYPE TOKEN**	not token
Is an original value	*expression* **IS TYPE ORIGINAL**	original
Is not an original value	*expression* **IS NOT TYPE ORIGINAL**	not original

Length Checking

Action	SQL	Filter Operator Prompt
Is of a constant length	*expression* **LENGTH IS CONSTANT** *length*	constant length
Is not of a constant length	*expression* **LENGTH IS CONSTANT** *length*	not constant length
Is within length range	*expression* **LENGTH BETWEEN** *min* **AND** *max*	in length range
Is not within length range	*expression* **LENGTH NOT BETWEEN** *min* **AND** *max*	not in length range

Approximate Matching

Action	SQL	Filter Operator Prompt
Matches a regular expression	*expression* **MATCHES** *expression* (can also use **NOT DIFFERS**)	matches expression
Matches one of many regular expressions	*expression* **MATCHES IN** (*value1, value2…*) (can also use **NOT DIFFERS IN**)	matches expression
Does not match a regular expression	*expression* **DIFFERS** *expression* (can also use **NOT MATCHES**)	differs
Does not match any of many regular expressions	*expression* **DIFFERS IN** (*value1, value2…*) (can also use **NOT MATCHES IN**)	differs
Has the same format as a specific value	*expression* **FORMATTED LIKE** *expression*	formatted like
Has the same format as any one of a specific set of values	*expression* **FORMATTED LIKE**(*value1, value2…*)	formatted like
Does not have the same format as a specific value	*expression* **FORMATTED DIFFERENTLY** *expression* (can also use **NOT FORMATTED LIKE**)	not formatted like
Does not have the same format as a specific set of values	*expression* **FORMATTED DIFFERENTLY**(*value1, value2…*) (can also use **NOT FORMATTED LIKE**)	not formatted like
Sounds like (matches phonetically) another value	*expression* **SOUNDS LIKE** *expression*	sounds like
Sounds like (matches phonetically) one of many values	*expression* **SOUNDS LIKE**(*value1, value2…*)	sounds like
Does not sound like another value	*expression* **SOUNDS DIFFERENT** *expression* (can also use **NOT SOUNDS LIKE**)	not sounds like
Does not sound like any one of many values	*expression* **SOUNDS DIFFERENT**(*value1, value2…*) (can also use **NOT SOUNDS LIKE**)	not sounds like
Is within edit distance of another value	*expression* **IN EDIT DISTANCE**(*expression, edit_distance, length_variation*)	in edit distance
Is within Jaro distance of another value	*expression* **IN JARO DISTANCE**(*expression, distance_measure*)	in jaro distance
Is within Jaro-Winkler distance of another value	*expression* **IN JARO WINKLER DISTANCE**(*expression, distance_measure, prefix_length, weight*)	in jaro-winkler distance
Is not within edit distance of another value	*expression* **NOT IN EDIT DISTANCE**(*expression, edit_distance, length_variation*)	not in edit distance
Is not within Jaro distance of another value	*expression* **NOT IN JARO DISTANCE**(*expression, distance_measure*)	not in jaro distance
Is not within Jaro-Winkler distance of another value	*expression* **NOT IN JARO WINKLER DISTANCE**(*expression, distance_measure, prefix_length, weight*)	not in jaro-winkler distance

Containment

Action	SQL	Filter Operator Prompt
Contains a specific value	*expression* **CONTAINS** *expression*	contains
Contains a specific value ignoring case	*expression* **CONTAINS INSENSITIVE** *expression*	contains
Contains any one of many values	*expression* **CONTAINS IN** (*value1,value2…*)	contains
Contains any one of many values ignoring case	*expression* **CONTAINS INSENSITIVE IN**(*value1,value2…*)	contains
Does not contain a specific value	*expression* **NOT CONTAINS** *expression*	not contains
Does not contain a specific value ignoring case	*expression* **NOT CONTAINS INSENSITIVE** *expression*	not contains
Does not contain any one of many values	*expression* **NOT CONTAINS IN** (*value1,value2…*)	not contains
Does not contain any one of many values ignoring case	*expression* **NOT CONTAINS INSENSITIVE IN**(*value1,value2…*)	not contains
Starts with a specific prefix	*expression* **STARTS WITH** *expression*	starts with
Starts with a specific prefix ignoring case	*expression* **STARTS WITH INSENSITIVE** *expression*	starts with
Starts with one of many prefixes	*expression* **STARTS WITH IN**(*value1,value2…*)	starts with
Starts with one of many prefixes ignoring case	*expression* **STARTS WITH INSENSITIVE IN**(*value1,value2…*)	starts with
Ends with a specific suffix	*expression* **ENDS WITH** *expression*	ends with
Ends with a specific suffix ignoring case	*expression* **ENDS WITH INSENSITIVE** *expression*	ends with
Ends with one of many suffixes	*expression* **ENDS WITH IN**(*value1,value2…*)	ends with
Ends with one of many suffixes ignoring case	*expression* **ENDS WITH INSENSITIVE IN**(*value1,value2…*)	ends with
Does not start with a specific prefix	*expression* **NOT STARTS WITH** *expression*	not starts with
Does not start with as specific prefix ignoring case	*expression* **NOT STARTS WITH INSENSITIVE** *expression*	not starts with
Does not start with one of many prefixes	*expression* **NOT STARTS WITH IN**(*value1,value2…*)	not starts with
Does not start with one of many prefixes ignoring case	*expression* **NOT STARTS WITH INSENSITIVE IN**(*value1,value2…*)	not starts with
Does not end with a specific suffix	*expression* **NOT ENDS WITH** *expression*	not ends with
Does not end with a specific suffix ignoring case	*expression* **NOT ENDS WITH INSENSITIVE** *expression*	not ends with
Does not end with one of many suffixes	*expression* **NOT ENDS WITH IN**(*value1,value2…*)	not ends with
Does not end with one of many suffixes ignoring case	*expression* **NOT ENDS WITH INSENSITIVE IN**(*value1,value2…*)	not ends with
Contains a specified token	*expression* **CONTAINS TOKEN** *expression*	contains word
Contains one of many tokens	*expression* **CONTAINS TOKEN IN** (*value1,value2…*)	contains word
Does not contain a specified token	*expression* **NOT CONTAINS TOKEN** *expression*	not contains word

Action	SQL	Filter Operator Prompt
Does not contain any of the specified tokens	*expression* **NOT CONTAINS TOKEN IN** (*value1,value2…*)	not contains word
Contains a format	*expression* **CONTAINS FORMAT** *expression*	contains format
Does not contain a format	*expression* **CONTAINS FORMAT IN** (*value1,value2…*)	contains format
Contains one of many formats	*expression* **NOT CONTAINS FORMAT** *expression*	not contains format
Does not contain any one of many formats	*expression* **NOT CONTAINS FORMAT IN** (*value1,value2…*)	not contains format
Contains a similar word	*Expression* **CONTAINS_EDITED_TOKEN**(*value, edit_distance, length_variation*)	contains similar word
Does not contain a similar word	*Expression* **NOT CONTAINS_EDITED_TOKEN**(*value, edit_distance, length_variation*)	not contains similar word
Contains a word that matches a value phonetically	*expression* **CONTAINS PHONETIC TOKEN** *expression*	contains soundalike
Does not contain a word that matches a value phonetically	*expression* **NOT CONTAINS PHONETIC TOKEN** *expression*	not contains soundalike

Result Checking

Action	SQL	Filter Operator Prompt
Is true	*expression* **IS TRUE**	true
Is not true	*expression* **IS NOT TRUE**	not true
Is false	*expression* **IS FALSE**	false
Is not false	*expression* **IS NOT FALSE**	not false
Is a warning	*expression* **IS WARNING**	warning
Is not a warning	*expression* **IS NOT WARNING**	not warning
Is an error	*expression* **IS ERROR**	error
Is not an error	*expression* **IS NOT ERROR**	not error
Is an error with a particular message	*expression* **IS ERROR TEXT** *expression*	the error
Is not an error with a particular message	*expression* **IS NOT ERROR TEXT** *expression*	not the error
Is a warning with a particular message	*expression* **IS WARNING TEXT** *expression*	the warning
Is not a warning with a particular message	*expression* **IS NOT WARNING TEXT** *expression*	not the warning

Data Mapping and Prototyping

Experian Pandora offers functionality that significantly eases the task of mapping source data to target data. The approach with Experian Pandora is to join, transform, aggregate and filter the available data sources until the result on-screen contains the data necessary for loading into the target application; effectively, building a prototype of the necessary rules and data.

Once the content of the drilldown is correct, you can use **Mapping** to rename and reposition columns to suit the final target table layout.

The available functionality can be broadly split into three key areas:

Data Prototype	Transforming data that has been previously loaded into Experian Pandora
Mapping	Associating source fields to an exact target layout
Generating a Mapping Report	Producing a report that shows exactly how source (loaded) data was transformed into a target view

Prototyping in Stages

There are several reasons for building a data prototype in stages, and to support this Experian Pandora allows drilldowns to be saved as Views and Tables.

Building a data prototype in stages:

- Assists in understanding
- Provides re-usable intermediate results
- Allows for logically incompatible operations such as requiring two different levels of aggregation to reach a result.
- Validates intermediate results, using a full data profile
- Repair incompatible keys on tables prior to joining them. This step requires the drilldown to be saved as a **Table**, whereas the previous stages can be simply saved as a **View**.

If a drilldown can be saved, the **Save As** button on the Edit dropdown menu will be enabled, and clicking it will give you one or more options as illustrated below:
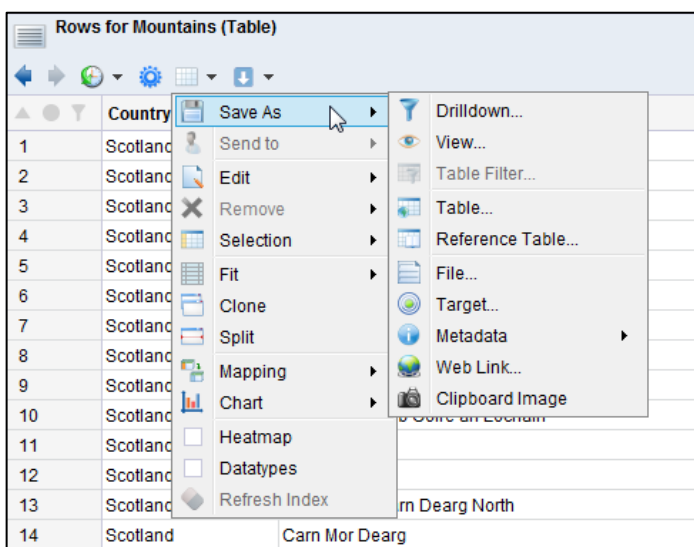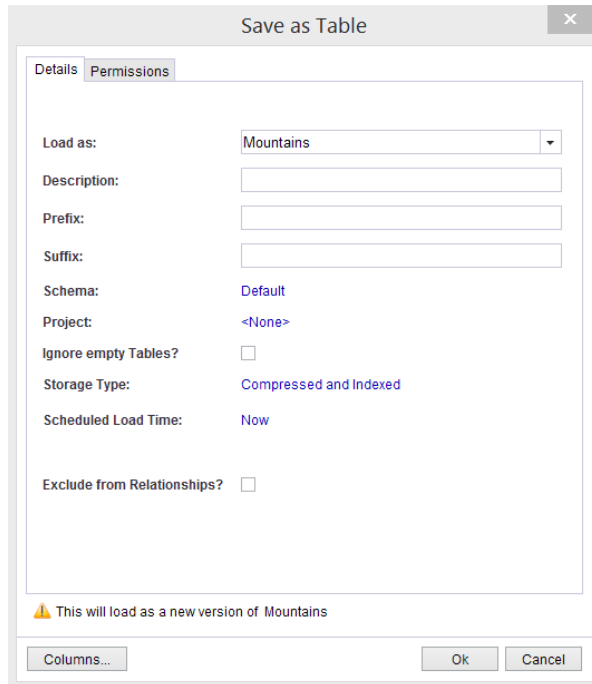
Creating a Saved Table

Selecting the option **Save As > Table** from a drilldown window will turn the data in the drilldown into a fully analysed and loaded table in Experian Pandora complete with quality statistics. The saved table retains a reference to the drilldown from which it was created, which can be accessed at any time in the future by drilling back to the original source.
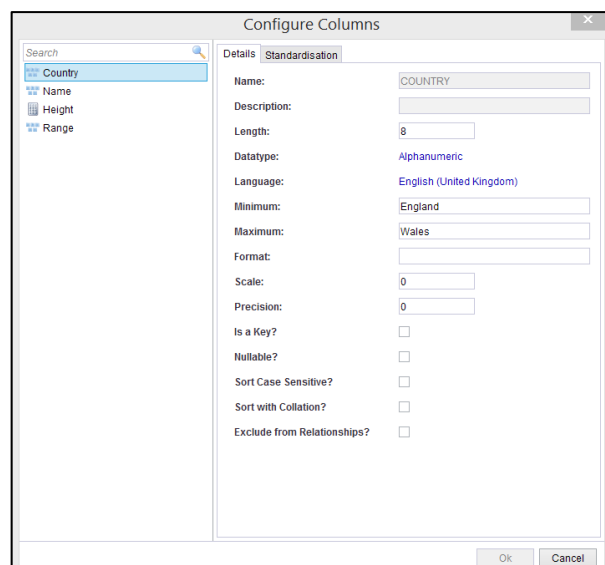
It should be noted that only data saved in this way is then subject to **relationship analysis** with other loaded tables.

Upon selecting this option the Save as Table dialog is displayed:

A warning will display as long as the table name is the same as the name of the table you are saving. As soon as you change the table name this warning will disappear.

The columns can be configured in the same way as when loading a table from new:

Clicking **Save** will start the load process (unless you have scheduled the load for some time in the future). The newly-created table may be accessed from the Experian Pandora explorer in the usual way.

Creating a Saved Drilldown

Selecting the menu option **Save As > Drilldown** from a drilldown window will save the drilldown exactly as you see it, which is actually saving the query used to create the drilldown. It is best thought of as a shortcut mechanism to get back a specific drilldown of interest, or to share it with others as Saved Drilldowns are potentially available to all users of Experian Pandora depending on their permissions settings. Drilldowns will be saved under the relevant parent object.

A drilldown can be parented by any of the following:
- Database
- Project
- Schema
- Tables
- Columns
- Relationships
- Views

Any info drilldown cannot be saved.

If the drilldown is created from a table then it will be saved in the Drilldowns section within the table where it has been created. NB: If it has been created from data that is not related to a specific table, e.g. a drilldown of all columns, then it will be saved in the global drilldowns under Database.

You can view a flat list of all drilldowns at Project, Schema or Global level by clicking on the drilldowns icon at the relevant level. This will also show the parent of the drilldown and when it was created and the owner

Selecting the **Save As > Drilldown** option presents a dialog where you can set the name and description of the Saved Drilldown, Novation options, and Export Definition.

When creating or editing a Saved Drilldown, the user has the option to automatically save the data as a table when a new version of the parent is loaded. This is done by ticking the 'Auto Save-as-Table?' option and providing the name to load it as. If the user gives a name of a table that does not currently exist, a new table is loaded rather than a new version. If it is a new table, the user is also given the option to choose which schema and project the new table should reside in.

Creating a View

Selecting the menu option **Save As > View** from a drilldown window will save the data in the drilldown in a dynamic way, i.e. without statistics and without the ability to infer relationships with other loaded tables. Selecting this option will show a simple dialog that allows you to specify the name and description of the View. The reference to the table from which the View was derived remains available.
If the underlying data in the original table changes, then the saved View will also change. For more information about this, see Novation.
The query from which a View is saved can always be accessed via the **Original Source** menu item (on the dropdown menu for the explorer item for the View), or via the Edit dropdown menu button when you drilldown to the rows of the View.
A View in many respects looks and feels just like a Table. All of the transformations used to create that data are hidden from you, allowing you to apply completely new transformations, groupings and filters.

Data Prototype

In this example, we will use the **Expression Editor** as well as filtering, grouping and joining to transform some source data until we achieve the required result which is fit for the new intended purpose.

In the following example we will combine **Purchase Order Headers** with **Purchase Order Details** and **Customers** tables, augment it with information from a reference table of **Discount Code** and perform some calculations and cleansing.

The reference table we will use contains discount codes and corresponding discount rates. We will also use a project to encapsulate our work.

Creating a project

Firstly, we create a new project. Right click on the **Projects** top-level element in the **Explorer,** select **New** from the menu, and call it **Mapping**.

Next we select the tables that are to be members of the project. Select from the Available Tables on the left by double clicking the individual tables you require or select the tables and use the add button at the bottom. This will move them to the Project Table.
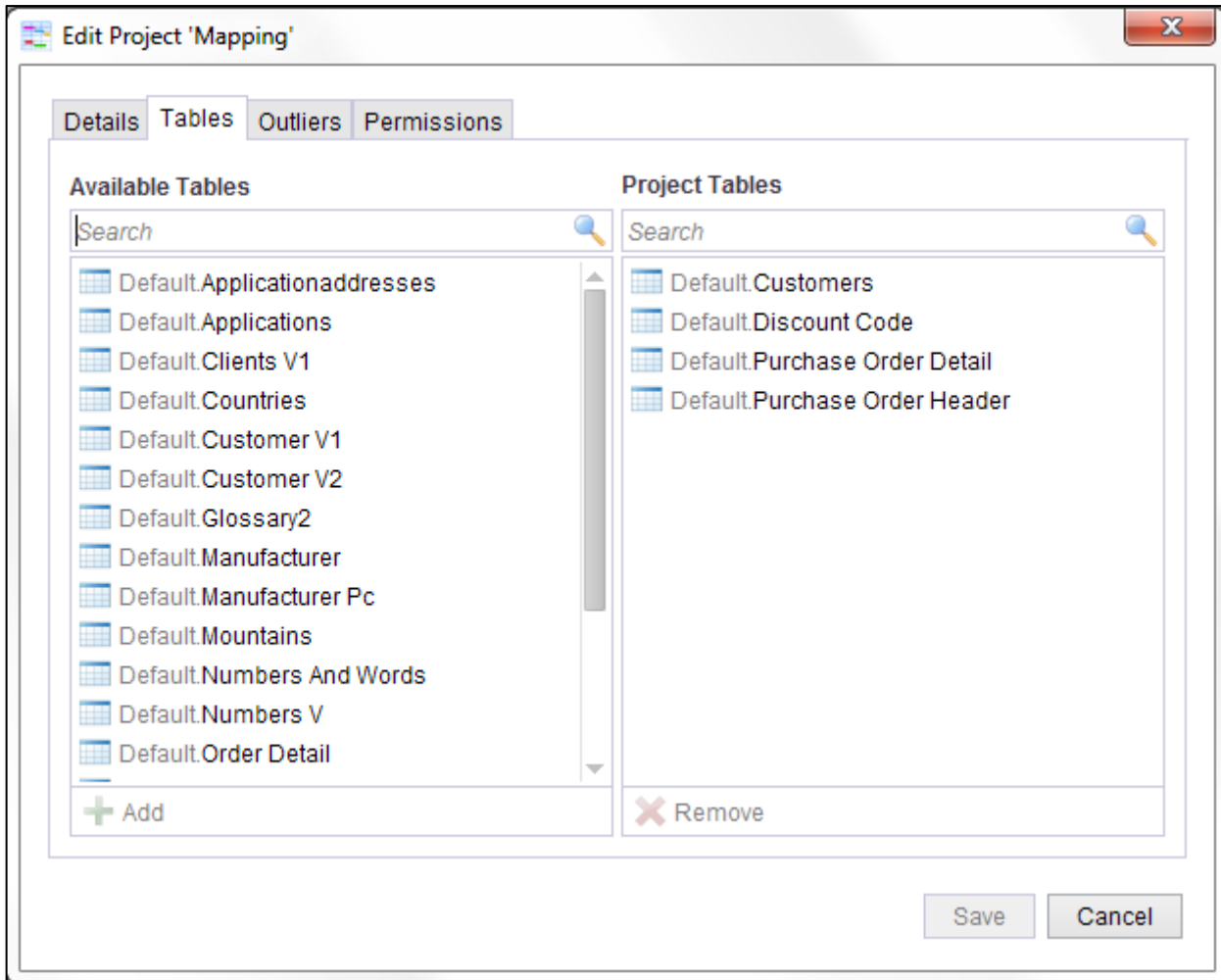
Click **Save** to create the project, which can now be seen in the explorer. We now right-click on the project and select **Set as Default**. This makes it our current working project.
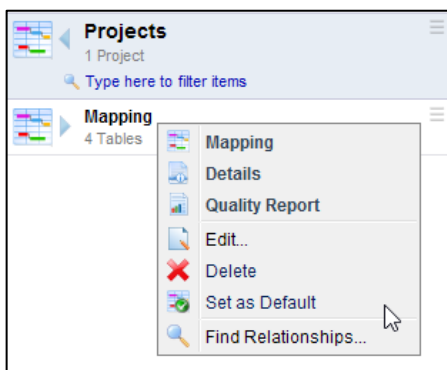
Join the tables together

We now want to combine the purchasing tables and the customer table using the relationship finder. Open up the project in the explorer, right-click on the explorer item **Tables** and select the **Tables** menu option.

This presents a drilldown of all tables in the project. We now select the individual purchasing and customer tables, using CTRL and left-click to select them individually, then right-click on the selection and choose the **Find Relationships** menu option.
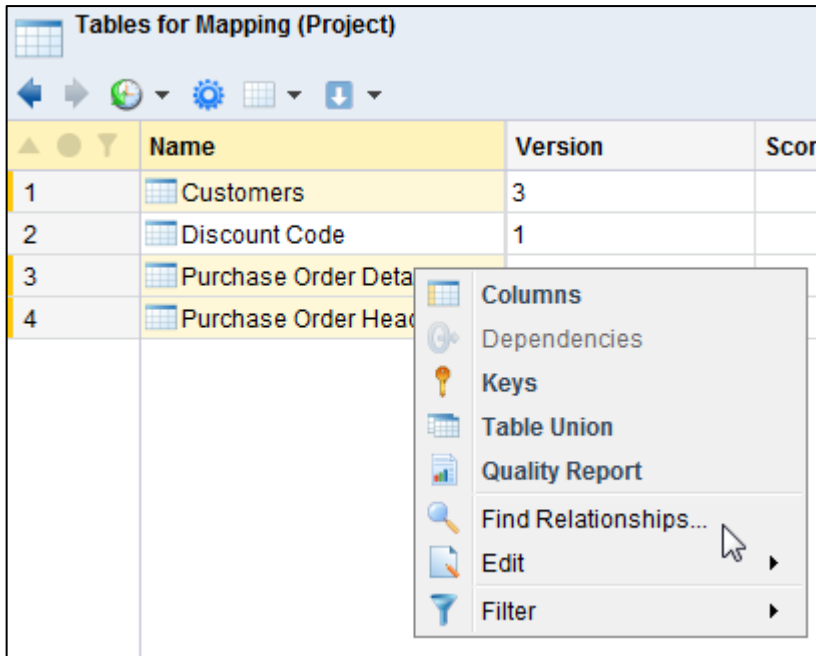
This displays the Relationship Finder:
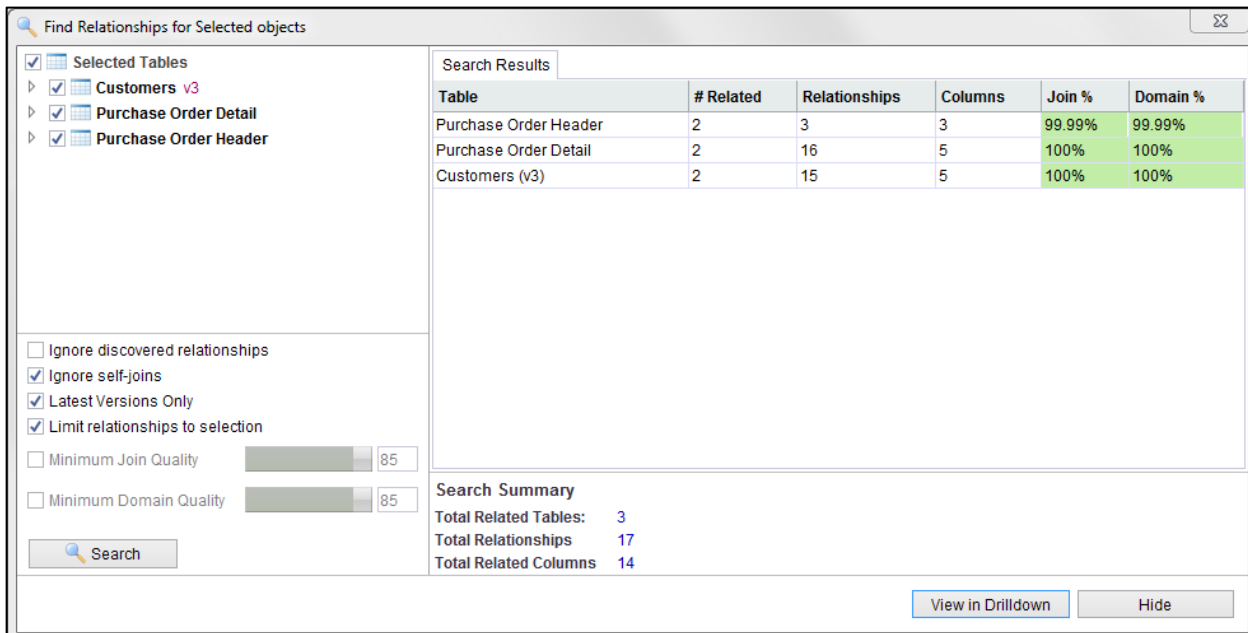
If we double-click on the Purchase Order Header row we can then see the relationships with the other tables in a new tab. Now double-click any of the rows to see the Venn Diagram showing the join.

We see that there is a strong relationship between Purchase Order Header and Purchase Order Detail on **Order Id** to **Order Id**. Since the Customer information comes from a different system, we discover that the values of **Customer Id** need to be standardised to allow the information to be joined together because the expected relationship between **Customers** and **Purchase Order Header** on **Customer Id** and **Customer Number** is not valid according to the Relationship Finder.

These relationships are 100% accurate as they are determined by Experian Pandora during load of the tables using the inherent structure of the Experian Pandora database.

The approach will therefore be to join Purchase Order Header and Purchase Order Detail and then to join the result to the Customer table on a modified version of the Customer ID column.

Drilling down to the Inner Join

Firstly, we select the relationship between **Purchase Order Header.Order Id** and **Purchase Order Detail.Order Id**, as shown in the relationship finder window above. If we click on the yellow part of the Venn diagram, on the number in brackets, this will produce a drilldown showing the rows in common between the two tables (the inner join).

Adding a Table Lookup translation

We can now start to modify this drilldown by adding our lookup to retrieve the discount percentage from the reference table. We will add this next to the **Discount Code** column. Right-click on the column header and select **Insert Column > Custom.**

This will display the **expression editor**. We will use the Experian Pandora Table Lookup functionality to achieve our translation. In the *Search* box on the left side of the expression editor, type *Table*. This brings up a filtered list of functions. Select the Table Lookup function and drag the function onto the expression editor desktop.
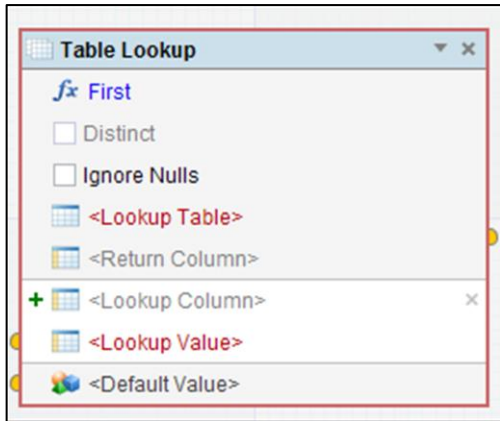
We can now fill in the fields in the function by clicking the text:

Lookup Type	First	The lookup type (first match)
Distinct	n/a	Whether or not to return distinct values
Ignore Nulls	<unchecked>	Whether or not to ignore null return values
Lookup Table	Discount Code	The table we want to lookup
Return Column	Discount Code.Rate	The column in the lookup table we want to return
Lookup Column	Discount Code.Discount Code	The column in the lookup table we want to match
Lookup Value	Discount Code	The value or column to match against the match column in the lookup table
Default Value	0	The value to return if no match is found

You may have noticed the **+** sign next to the Lookup Column and Lookup Value fields (which are also highlighted in white compared to other fields in the function). This is because you can add additional lookup columns to do a multi-column match.

See the **Lookup Types** section below for a more detail description of the available lookup types.
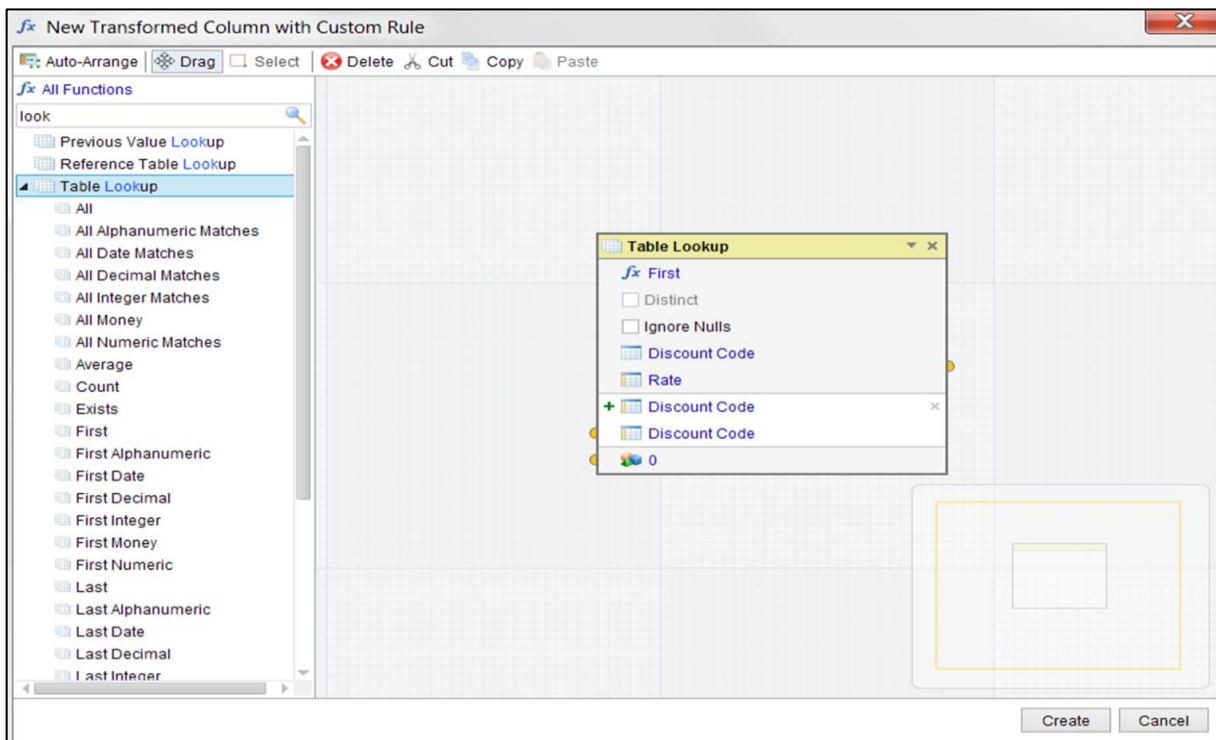
Now click **Create** to save the new column and call it **Discount Rate.**

Table Lookup Types

The Table Lookup function supports a number of possible return types:

First – returns the value of the return column on the first row that matches
Last - returns the value of the return column on the last row that matches

First Integer – returns the first Integer value of the return column on matching rows
First Decimal – returns the first Decimal value of the return column on matching rows
First Numeric – returns the first numeric (Integer, Decimal or Money) value of the return column on matching rows
First Date – returns the first Date value of the return column on matching rows
First Money – returns the first Money value of the return column on matching rows
First Alphanumeric – returns the first Alphanumeric value of the return column on matching rows

Last Integer – returns the last Integer value of the return column on matching rows
Last Decimal – returns the last Decimal value of the return column on matching rows
Last Numeric – returns the last numeric (Integer, Decimal or Money) value of the return column on matching rows
Last Date – returns the last Date value of the return column on matching rows
Last Money – returns the last Money value of the return column on matching rows
Last Alphanumeric – returns the last Alphanumeric value of the return column on matching rows

All – returns all return values for rows that match (as a comma-separated list)
All Integer Matches - returns all Integer return values for rows that match (as a comma-separated list)
All Decimal Matches - returns all Decimal return values for rows that match (as a comma-separated list)
All Numeric Matches - returns all numeric (Integer, Decimal or Money) return values for rows that match (as a comma-separated list)
All Date Matches - returns all Date return values for rows that match (as a comma-separated list)

All Money Matches - returns all Money return values for rows that match (as a comma-separated list)
All Alphanumeric Matches - returns all Alphanumeric return values for rows that match (as a comma-separated list)

Minimum – returns the minimum value of the return column in all rows that match
Maximum – returns the maximum value of the return column in all rows that match
Exists – returns true/false to indicate whether or not there is any match
Not Exists – the opposite of Exists
Count – returns the number of matches
Sum – returns the sum of the return column values in all rows that match
Average – returns the average of the return column values in all rows that match
Standard Deviation (Sample) – returns the sample standard deviation of the return column values in all rows that match
Standard Deviation (Population) – returns the population standard deviation of the return column values in all rows that match
Variance (Sample) – returns the sample variance of the return column values in all rows that match
Variance (Population) – returns the population variance of the return column values in all rows that match

In addition to the above, the following filters may also be applied to the above where applicable:

Distinct – for lookup types returning lists, enabling this option ensures that the list contains distinct (i.e. not duplicated) values
Ignore Nulls – for all lookup types, enabling this option ensures that null return values are not returned or processed.

Extending the table lookup to perform a calculation

Now that we have the basic functionality to translate our discount code into a rate, we need to calculate the net value of each order line using the formula **Price** x **Quantity** x **Discount %**

In our case, the discount is not the percentage number, so will need dividing by 100. It is also a discount rate, so if we want to get the discounted price we have to subtract that from the whole price – most easily achieved by subtracting the discount rate percentage from 100%. So the actual calculation implemented will be:

Price x **Quantity** x (100 – **Discount Rate**) / 100

To achieve this we use two calculate function boxes. The first one does the subtraction of **Discount Rate** from 100, and the second one performs all the multiplication and division for the rest of the calculation. Using multiple calculate boxes like this is how bracketing of formulas is achieved.
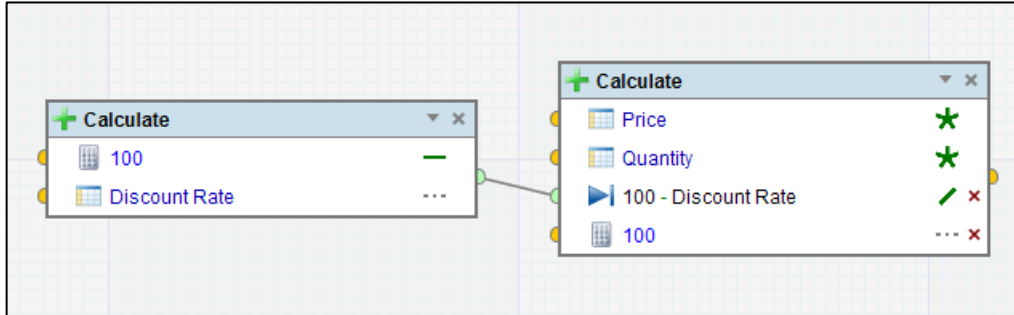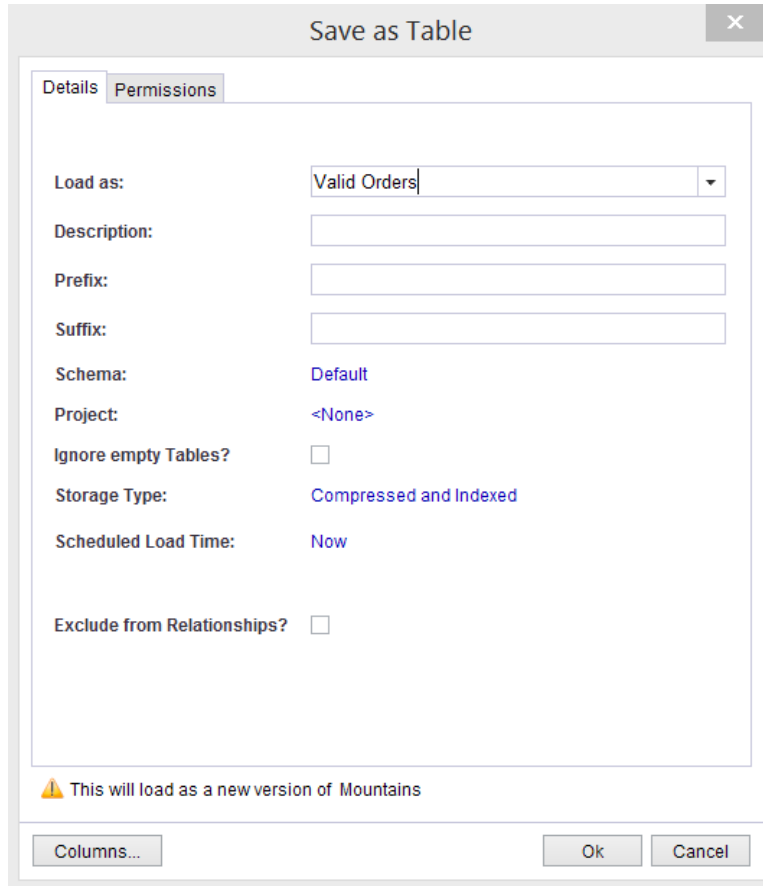
Save the new column as **Net Value**

ƒ Discount Rate	ƒ Net Value	Sh
11	37380	06-
11	194.6875	09-
11	35.6	01-
11	10.68	05-
11	37.38	06-

Save the join as a Table

We can now save the inner join of Purchase Order Header and Purchase Order Detail complete with new Discount Rate and Net Value columns as a new table. This is required in order to join it with the Customers table, which we will first need to modify.

Our project has automatically been selected. We will call the new table **Valid Orders**. We simply accept the default parameters and select **Save**. This will save in the background. Once it has saved, it will be available in the project.

Modify the key of a table to enable joining

We now turn our attention to the Customers table which requires the key to be modified (Customer Id) in order for it to match to Customer Number on the Purchase Order Header table. We also need to clean up the telephone number column.

The first thing we need to do is understand why the Customers and Purchase Order Header tables do not join. The easiest way to do this is to perform a quick visual comparison. We already know that there is no relationship listed at all in the Relationship Finder which means that there is not a single value in common between these two columns, and hence not a single row in common between the two tables.

From the **Mapping** project, open up the **Columns** item on the explorer. This is a list of the columns across all the tables in the project.

By right-clicking in turn on the **Customer Id.Valid Orders** and **Customer Id.Customers** columns, we can select the right-click menu option **Values** to see the unique values in each of those columns.
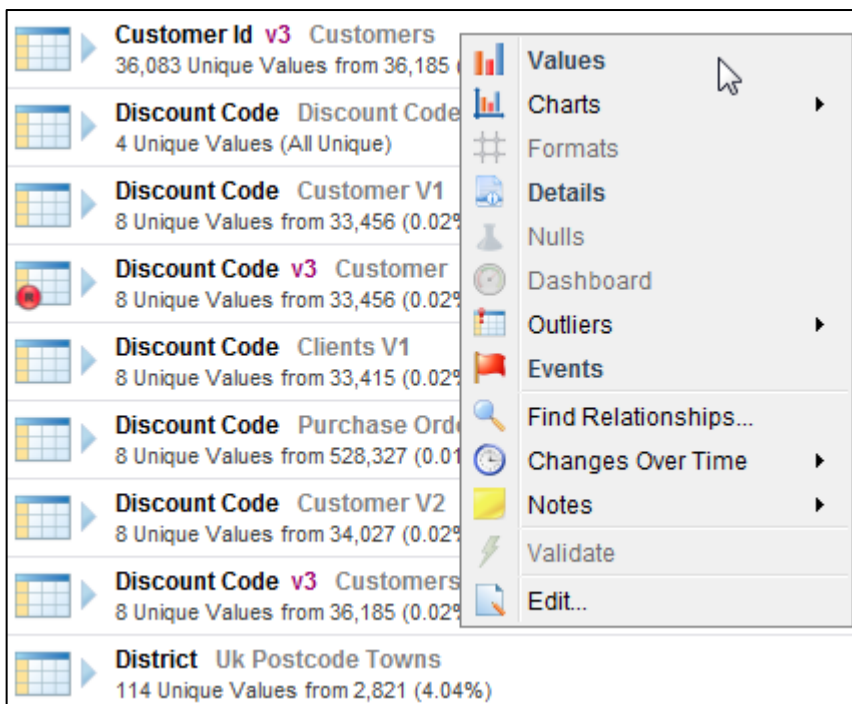
You should ensure you are not in Single Drilldown Mode to do this comparison. You can then use the window arrangement facilities to move these windows side by side. By dragging the window past the edge of the Experian Pandora desktop, it will be offered to you as docked to either the left or right, splitting the desktop in half:
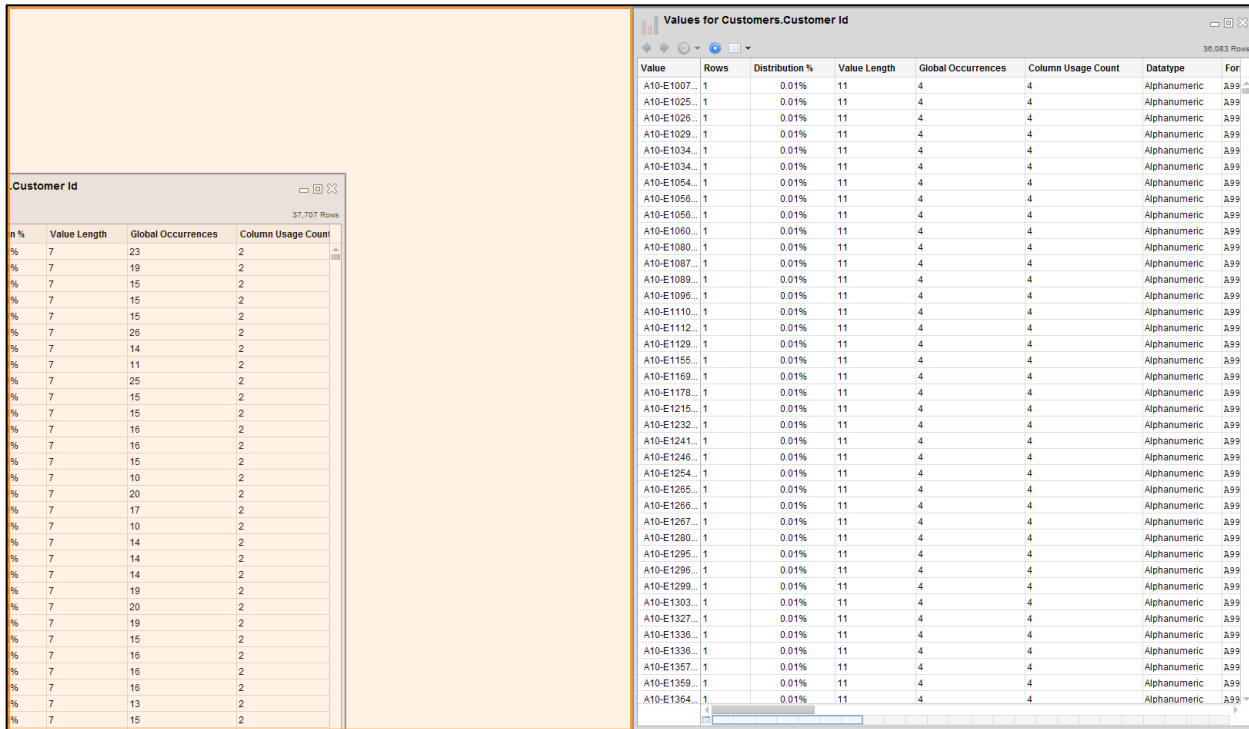
Release the mouse to stop dragging and the window will dock into place. You now have a useful side-by-side comparison.

It can be immediately seen that the **Customers.Customer Id** column appears to have the same value as that of **Valid Orders.Customer Id,** but is prefixed. We can quickly test-transform **Customers.Customer Id** to strip the header whilst visually comparing by creating a new column that repairs the **Customer Id** Value. To do this we use the **After** function, and simply take everything in the **Value** column after the "-"

Let's save the new column as Fixed Value. Once we have the new column, we need to sort the Customer Id by Fixed Value.

Values for Valid Orders.Customer Id — 37,707 Rows

Value	Rows	Distribution %	Value Length	Global Occurrences	Column Usage Count	Datatype
E100001	21	0.01%	7	23	2	Alphanum
E100003	17	0.01%	7	19	2	Alphanum
E100005	13	0.01%	7	15	2	Alphanum
E100006	13	0.01%	7	15	2	Alphanum
E100008	13	0.01%	7	15	2	Alphanum
E100009	24	0.01%	7	26	2	Alphanum
E100011	12	0.01%	7	14	2	Alphanum
E100012	9	0.01%	7	11	2	Alphanum
E100016	23	0.01%	7	25	2	Alphanum
E100017	13	0.01%	7	15	2	Alphanum
E100019	13	0.01%	7	15	2	Alphanum
E100020	14	0.01%	7	16	2	Alphanum
E100022	14	0.01%	7	16	2	Alphanum
E100023	13	0.01%	7	15	2	Alphanum
E100024	8	0.01%	7	10	2	Alphanum
E100025	18	0.01%	7	20	2	Alphanum
E100026	15	0.01%	7	17	2	Alphanum
E100027	8	0.01%	7	10	2	Alphanum
E100028	12	0.01%	7	14	2	Alphanum
E100030	12	0.01%	7	14	2	Alphanum
E100031	12	0.01%	7	14	2	Alphanum
E100033	17	0.01%	7	19	2	Alphanum
E100034	18	0.01%	7	20	2	Alphanum
E100035	17	0.01%	7	19	2	Alphanum
E100036	13	0.01%	7	15	2	Alphanum
E100038	14	0.01%	7	16	2	Alphanum
E100040	14	0.01%	7	16	2	Alphanum
E100042	14	0.01%	7	16	2	Alphanum
E100044	11	0.01%	7	13	2	Alphanum
E100045	13	0.01%	7	15	2	Alphanum
E100046	13	0.01%	7	15	2	Alphanum
E100047	14	0.01%	7	16	2	Alphanum
E100048	13	0.01%	7	15	2	Alphanum
E100049	15	0.01%	7	17	2	Alphanum
E100051	13	0.01%	7	15	2	Alphanum
E100052	11	0.01%	7	13	2	Alphanum
E100053	12	0.01%	7	14	2	Alphanum
E100055	16	0.01%	7	18	2	Alphanum

Values for Customers.Customer Id — 36,083 Rows

Value	Rows	f Fixed Value	Distribution %	Value Length	Global Occurrences	Col...
B17-E1000...	1	E100001	0.01%	11	4	4
C84-E1000...	1	E100003	0.01%	11	4	4
D29-E1000...	1	E100005	0.01%	11	4	4
D80-E1000...	1	E100006	0.01%	11	4	4
B64-E1000...	1	E100008	0.01%	11	4	4
D72-E1000...	1	E100009	0.01%	11	4	4
C85-E1000...	1	E100011	0.01%	11	1	1
C17-E1000...	1	E100012	0.01%	11	4	4
D76-E1000...	1	E100016	0.01%	11	4	4
D45-E1000...	1	E100017	0.01%	11	4	4
D83-E1000...	1	E100019	0.01%	11	2	2
B96-E1000...	1	E100020	0.01%	11	4	4
B90-E1000...	1	E100022	0.01%	11	4	4
A57-E1000...	1	E100023	0.01%	11	4	4
A68-E1000...	1	E100024	0.01%	11	4	4
D24-E1000...	1	E100025	0.01%	11	4	4
B53-E1000...	1	E100027	0.01%	11	1	1
D75-E1000...	1	E100028	0.01%	11	4	4
D93-E1000...	1	E100030	0.01%	11	4	4
D30-E1000...	1	E100031	0.01%	11	4	4
D38-E1000...	1	E100033	0.01%	11	1	1
D50-E1000...	1	E100034	0.01%	11	4	4
B33-E1000...	1	E100035	0.01%	11	4	4
C44-E1000...	1	E100036	0.01%	11	4	4
D73-E1000...	1	E100038	0.01%	11	4	4
A97-E1000...	1	E100040	0.01%	11	4	4
C53-E1000...	1	E100042	0.01%	11	4	4
D35-E1000...	1	E100044	0.01%	11	4	4
C14-E1000...	1	E100045	0.01%	11	4	4
D45-E1000...	1	E100046	0.01%	11	4	4
D73-E1000...	1	E100047	0.01%	11	4	4
C33-E1000...	1	E100048	0.01%	11	4	4
D36-E1000...	1	E100049	0.01%	11	4	4
D59-E1000...	1	E100051	0.01%	11	4	4
D57-E1000...	1	E100052	0.01%	11	4	4
D34-E1000...	1	E100053	0.01%	11	4	4
B17-E1000...	1	E100055	0.01%	11	4	4
D90-E1000...	1	E100056	0.01%	11	4	4

From a simple visual inspection we can see that the fixed value does indeed look to be identical to that of **Valid Orders.Customer Id**. This is not conclusive, and we could examine this further, but for our purposes we will go ahead and use this transformation to fix the **Customers** table for real.

From the explorer select **Projects > Mapping > Tables > Customers** click on the **Rows** item to see all the rows in the **Customers** table. We now can now create a column that is the cleansed key, which will match the column in **Valid Orders**.

Right-click on the **Customer Id** column header and select **Transformation > Create** to bring up the expression editor. Use the **After** function in the same way as we trialled previously. Save the transformation and the original column will automatically hide, leaving us with the transformed **Customer Id** which matches that of our **Valid Orders** table. We now have the following result:

f Rows for Customers

Row Id	f Customer Id	Discount Code
1	E100750	M
2	E102508	H
3	E102659	H
4	E102982	M
5	E103408	M
6	E103489	H
7	E105402	N
8	E105686	N
9	E105688	M
10	E106043	N
11	E108032	N
12	E108768	N
13	E108908	M
14	E109669	M
15	E111093	N

Cleanse the Telephone Number column

We can now cleanse the **Telephone** column. From the explorer, select and open:
Projects > Mapping > Tables > Customers > Columns > Telephone > Profile

This gives you the profile for the Telephone column. Click on **Unique Formats** item in the explorer (or right-click and select **Unique Formats**):

We can see immediately from the statistics on the explorer that there are 83 different formats for telephone number, and a visual inspection of the drilldown of formats shows a myriad of different formats, some of which are complete nonsense.

After some inspection, we decide that we need to:

- **Blank out any numbers containing text**
- **Remove all punctuation characters**
- **Strip leading zeros as some numbers have them and others do not**
- **Strip the country code at the front if the numbers have it**
- **Add the country code as a final step by looking up into another reference table**

We can go back to the rows of Customers we were working on by clicking on the **Customers** button on the desktop toolbar. In this instance we are going to transform the Telephone number in place. We right-click on Telephone and select **Transformation > Create.**

This brings up the expression editor once again.
The first function we apply is the **Remove Noise** function. This strips out all white space and punctuation.

We feed the output of this into a function to remove leading zeroes, the **Remove Leading Zeroes** function.

We now need to concatenate the correct dial code with this resultant function. To do this, we use the built-in ISO country content to translate the **Country** column in the Customers table to get the valid dial code for that country. We use the **Convert ISO Country** function to do this, passing **Country**, from Country Name to Dial Code.

We now can use the **Concatenate** function to concatenate our translated dial code and clean telephone number. However, we have not accounted for short numbers or those containing letters.

We now add a **Decision** function. This has a single test, a condition to check whether the original telephone number had either letters in it (**contains format 'A'**) or is not of the correct length (**not within length range 11 to 13**). The cleanse telephone number is fed into the Decision function as the value to be output if the conditional test is true. If it is not true, we want to return nothing at all (Null). The default of the Decision function is to return an error message. We will instead feed the function **Null** into the false return value of the Decision function, so that we get a Null if the telephone number was garbage.

When the function is complete, we can also edit the individual function box names to better describe what is going on.

The whole end-to-end expression, complete with annotated functions, looks like this:
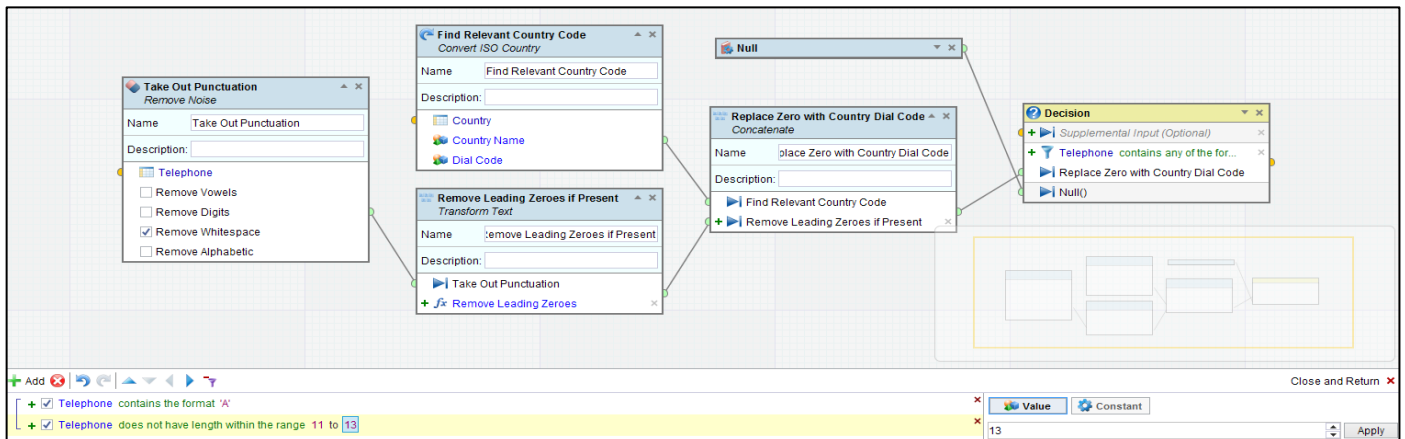
Click **Create** and our transformation will replace the original column on the visible columns drilldown.
It is fairly self-explanatory viewed this way. We have renamed the functions to give more information to other viewers of this transformation as to our intent. Viewing the transformed Telephone field reveals the type of telephone number data we were expecting:

We could further enhance this to place a "**+**" on the front of the number, bracket the country code etc. This is fine for this example as it is to be used in an automated dialling system.

Saving the modified Customers table

Now that we have cleansed the telephone number, we need to filter out any records that do not have telephone numbers (i.e. ones where the telephone number was garbage). We can simply do this by filtering out all those records that are NULL as our transformation returned NULL if the telephone number could not be cleansed.

There are two ways to do this. One is to use the filter dialog in the Drilldown Manager. The other is to simply Quick Filter it. An easy way to see if a column contains nulls is to sort it as column values containing null will always appear at the top of a sort.

Right-click on the Telephone column header on the drilldown and select **Sort > Only By This (Ascending)**

We can clearly see the null values in Telephone. Simply right-click on one of them and select **Filter > Remove Selection**
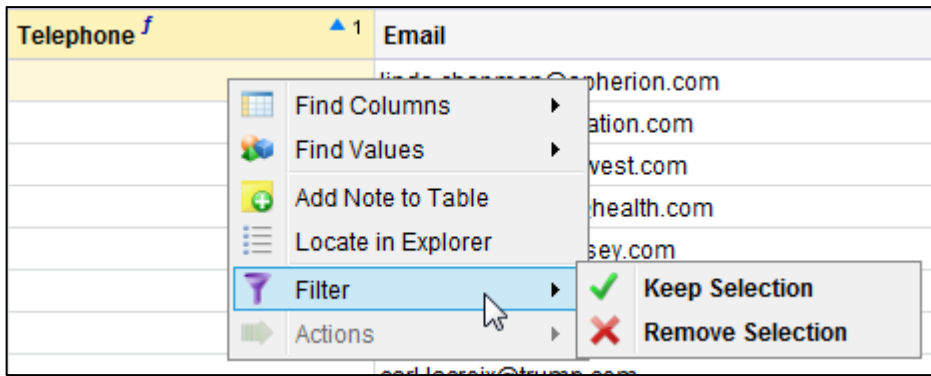
This builds a filter for you to remove the selected values, in this case, all Null values in that column.

The result is a filter where **Telephone is not NULL**. This approach is quick and interactive when data sets are of reasonable size. For very large data sets (many millions of rows) it is faster to build the filter manually in the Drilldown Manager.

Here is the final result:

We now save this result as a new version of the **Customers** table**.** Simply save as a table, and leave the name as the default of **Customers**. This automatically creates a new version of the Customers table. This will save in the background – it will show as a job running on the Jobs icon in the Quick Bar. Once it has saved it will be available in the project.

Completing the 3-way join

Using the Relationship Finder again, we can now see that the new version of Customers does indeed have a fairly strong relationship to the newly created **Valid Orders** table.

Once again, we can see there are customers who have never ordered (not really an issue) but there are orders that have no corresponding customer number! This is an obvious data quality issue requiring further exploration.

For our example, we are happy to just take the valid information for all orders made by customers who are telephone contactable. We simply click on the number in brackets in the yellow centre of the Venn Diagram to get the join between Customers v2 and Valid Orders.

You may recall that:

1.   **Valid Orders** is the join of **Purchase Order Header** and **Purchase Order Detail** (for all detail records with a matching header).

2.   **Customers v2** is the cleansed version of **Customers** where the key has been modified to match the format of that originally found in **Purchase Order Header,** the telephone number corrected and all non-contactable customers removed.

3.   The final result is the 3-way join of **Customers, Purchase Order Header** and **Purchase Order Detail** taking into account we have eliminated non-contactable customers and any stray orders, detail records and customers that have never ordered.

At this stage, it is usual to use grouping, counting, filtering and instant profiling (right-click on column headings) functionality to validate the results of all the rules which have been applied.

Before performing any further work on the results of the join, you need to save the relationship.
To do this, right-click on the relationship row in the Table Relationships and select the **Save** command.

You can now **Save As > Drilldown** to checkpoint your work before moving on to Mapping.

Source-to-Target Mapping

As part of the process of mapping source data to target data, it is often the case that the names and associated metadata of the source columns do not exactly match those of the target columns, despite the data content being correct.

You could manually set up the mapping, by right-clicking each individual column heading in the drilldown, and manually editing the column details to match the desired target.

But Experian Pandora provides a Mapping Editor that greatly simplifies and speeds up the task of matching columns in a drilldown to those of a target layout.

The Mapping Editor provides the ability to automatically map the data you have prepared to a fixed target structure. Where field names are not identical, fuzzy matching techniques are applied to give the best fit. The final details can be specified by the analyst.

Continuing our example, we now wish to map the final result from the previous section using a pre-determined layout.
The first thing to do is to click on **Mapping > Create,** on the relationship drilldown, for any of the discovered relationships.
Using the drilldown actions menu:
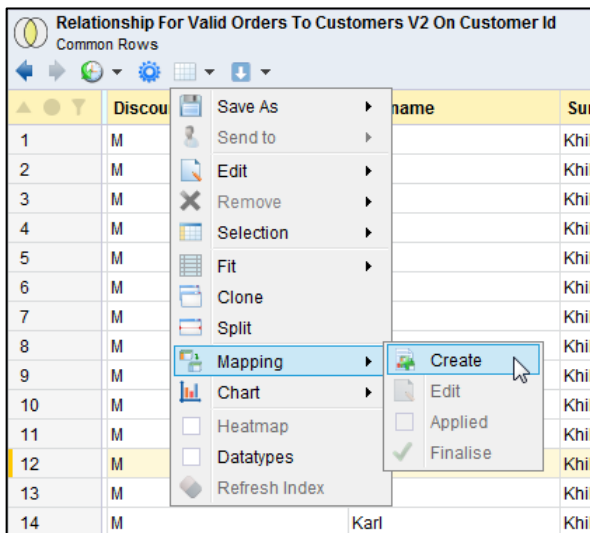
When the Mapping Editor opens, you need to choose a target to map this drilldown to. We have several choices at this point as to where we get the target layout information from:

Source	Description
Table	Use the layout and metadata of an existing Table in Experian Pandora
View	Use the layout and metadata of an existing View in Experian Pandora
Server Metadata	Use layouts specified in either xml configuration format or ANSI DDL format from the metadata area of the Server
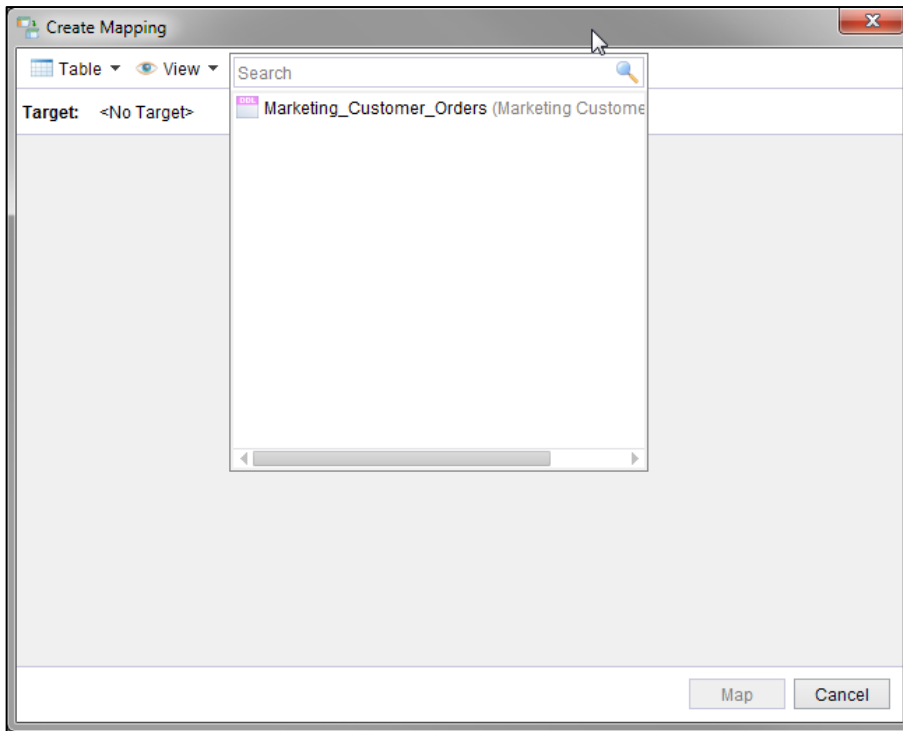
Here we have loaded the target as a DDL metadata file **Marketing Customer Orders.DDL** which has the target layout we want to use. Once you have selected the mapping target, you can instruct the Mapping Editor to automatically match source columns to target columns based on their names. You can try matching names exactly using **Auto-map > Exact name match**, and then try **Loose name match** which uses a number of fuzzy rules that allow similar names to be matched. In each case you can choose to retain existing mappings, or remove all of the mappings and start from a clean slate.
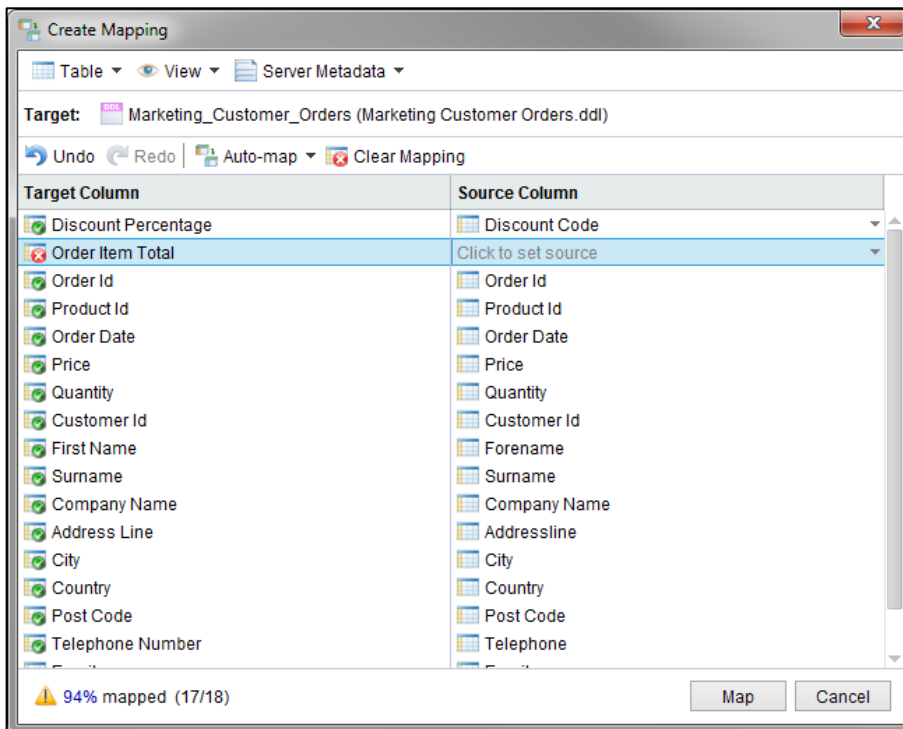
Map, review and refine the mapping

You can review the effect of the mapping after trying the auto-mapping by pressing the **Map** button. This will apply the target column names and types to the source columns and produce a new drilldown, named to correspond to the target layout.

You can now see if the mapping appears as you intended.

Notice that the drilldown title bar has changed its icon and subtitle to show that mapping is in progress, and is not yet finalised.

In this example, you can see that the **Discount Percentage** initially found by the Loose Name Match doesn't look right, and **Order Item Total** hasn't been mapped.

You can re-enter the Mapping Editor to change the mappings simply by pressing the Drilldown Actions button and selecting the **Mapping > Edit** option.

Choosing alternative sources

Where auto-matching has found multiple choices of source column for a potential match in a target column, you will see a drop down arrow to the right of the source column. Simply click on this down arrow to choose the correct source from the list of alternatives. In this example, we can see for the target **Discount Percentage** the source **Discount Code** was selected as a loose match, but we know that **Discount Rate** is the correct choice.
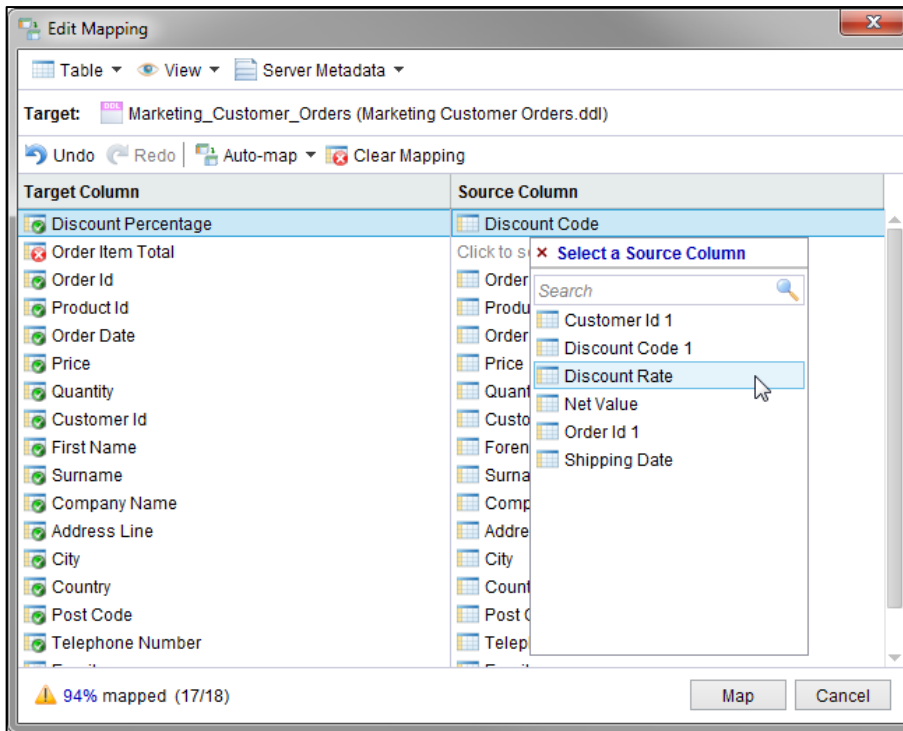
Choosing sources when no match found automatically

The **Order Item Total** column could not be auto-mapped. Clicking on the **Click to set source** prompt offers you all the remaining unmapped fields. If the field you wanted has been wrongly mapped elsewhere you need to first clear that mapping to make it available to be mapped elsewhere. Here we select the **Net Value** column as our source.
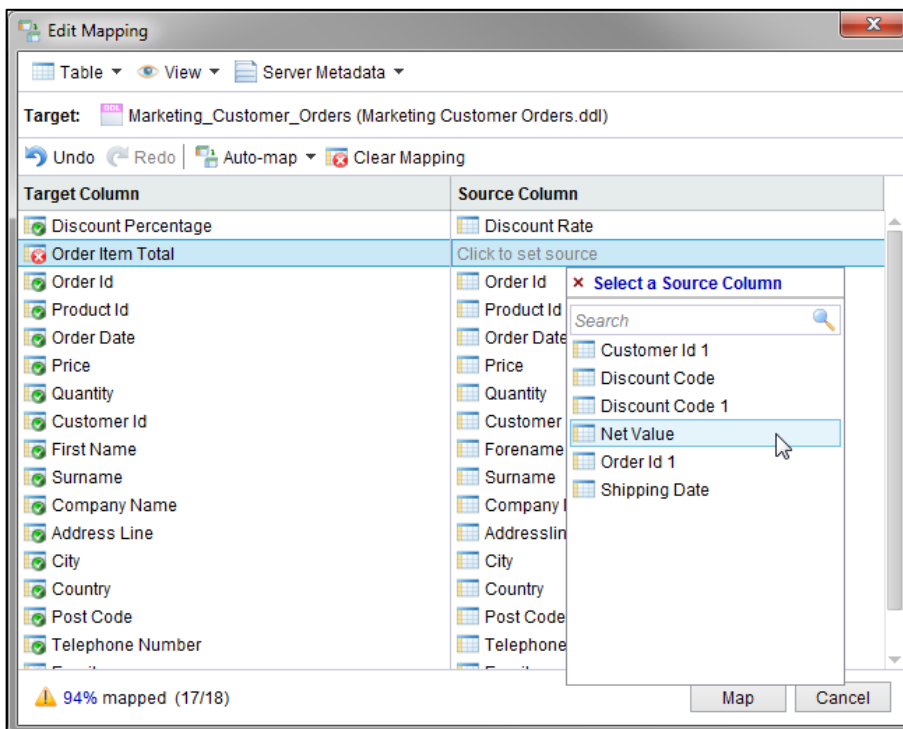
After refining the mapping and pressing the **Map** button again you will return to the drilldown with fresh results

	Discount Percentage	Order Item Total	Order Id	Product Id	Order Date	Price	Qua
1	11	242.08	1096742	SMLLBL716	17-Jul-2009	GBP 4	68
2	11	308.385	1096742	ST611	17-Jul-2009	GBP 3.5	99
3	11	44.5	1096742	KT794	17-Jul-2009	GBP 10	5
4	11	12460	1096742	BMPR381	17-Jul-2009	GBP 5	280
5	11	48.06	1096742	STRP480	17-Jul-2009	GBP 9	6
6	11	73.425	1096742	RBBRPL127	17-Jul-2009	GBP 1.25	66
7	11	77.875	1096742	BLBX001	17-Jul-2009	GBP 2.5	35
8	11	422.75	1108046	SCRW1113	31-Jul-2009	GBP 1	475

Additional transformations

Now the Discount Percentage looks correct. But the Order Item Total column includes numbers to 3 decimal places. For the sake of discussion, let's say that we only want to deal with whole pennies or cents. Fixing this requires an additional transformation.

Experian Pandora allows you to switch between viewing the drilldown with the mapping applied, or as it appeared before you started mapping. This allows you to adjust the source columns without losing the mappings you have already defined. You control this by pressing the Drilldown Actions button and checking or unchecking the **Mapping > Applied** option.

Now that the source columns are visible you can add a new column that calculates the **Net Value** rounded to 2 decimal places.

When you right click on the Net Value column heading, the options are slightly different while you are still mapping. Instead of adding a new column you are creating a **New Source Column** which will initially be hidden, but is available to the Mapping Editor.

After defining the custom rule and giving the new column the name **Rounded Net Value**, you can return to the Mapping Editor to complete the mapping.
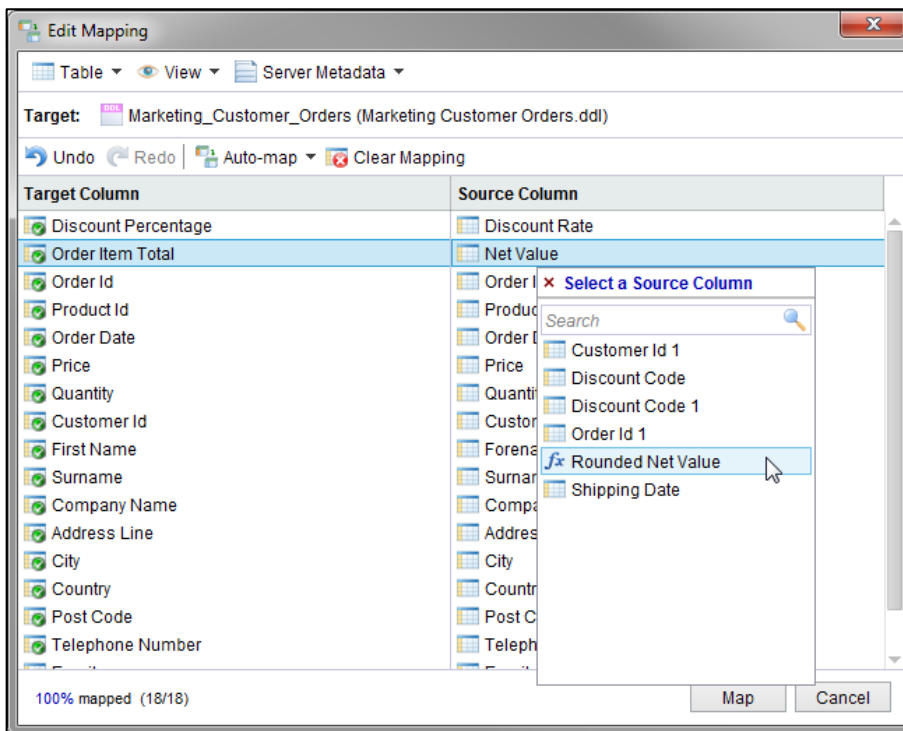
Here you can see that the **Rounded Net Value** column is now available to be used as the source for the **Order Item Total**.

	Marketing Customer Orders.ddl					18 of 26 Columns
	Currently being mapped (18 of 18 columns - 100%)					
						360,174 Rov
▲ ● ▼	Discount Percentage	Order Item Total *f*	Order Id	Product Id	Order Date	Price
1	11	242.08	1096742	SMLLBL716	17-Jul-2009	GBP 4
2	11	308.38	1096742	ST611	17-Jul-2009	GBP 3.5
3	11	44.5	1096742	KT794	17-Jul-2009	GBP 10
4	11	12460	1096742	BMPR381	17-Jul-2009	GBP 5
5	11	48.06	1096742	STRP480	17-Jul-2009	GBP 9
6	11	73.42	1096742	RBBRPL127	17-Jul-2009	GBP 1.25
7	11	77.87	1096742	BLBX001	17-Jul-2009	GBP 2.5
8	11	422.75	1108046	SCRW1113	31-Jul-2009	GBP 1
9	11	843.72	1108046	KT908	31-Jul-2009	GBP 12
10	11	9612	1108046	SCRW892	31-Jul-2009	GBP 2.4

After performing the mapping again you can review the results with the rounded **Order Item Total**.

Finalize the mapping

When you have finished tweaking the mapping, you need to select the **Mapping > Finalise** command. This completes the process of editing the mappings, and is important because it releases Experian Pandora from restricting the editing that you can perform while mapping.
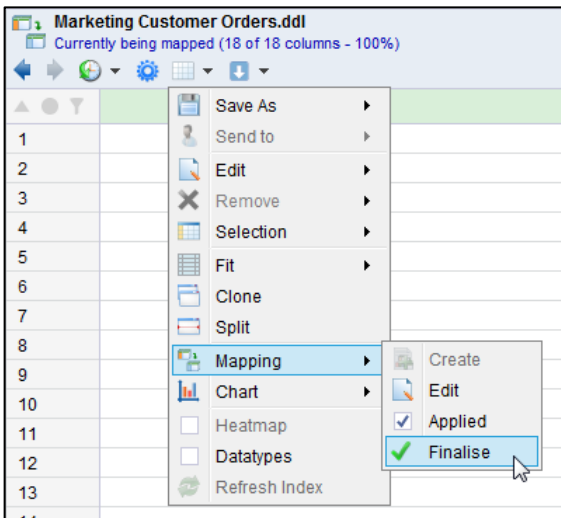
The resulting drilldown should then be saved as a dynamic view to allow the generation of the mapping report. **Save as > View** and naming it Marketing Customer Orders View. The description could be used to reference which tables and columns have been used to create the view.

Producing a Mapping Report

A Mapping Report can be produced from any View or Saved Table via the right-click pop-up menu on the corresponding explorer item and selecting **Mapping Report**:
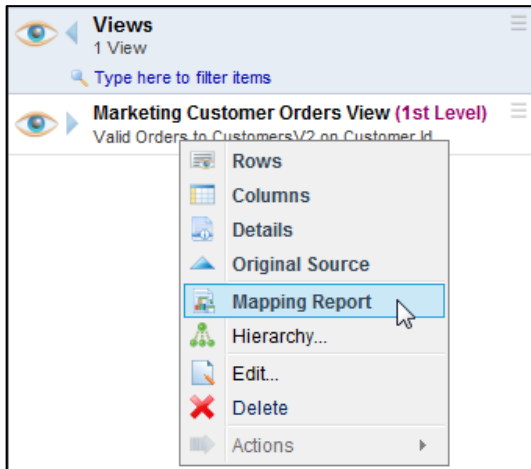
This is also available from the Drilldown menu on the drilldown of rows for a saved table or view:

You can also produce a quick mapping report at any time from a drilldown with transformations, regardless of whether you have saved it. Here we are viewing the set of relationships we were looking at earlier:

Mapping reports are dynamically created every time you request them.

In this example, we will create the mapping report by right-clicking on the saved view **Marketing Customer Orders View** and selecting **Mapping Report**.
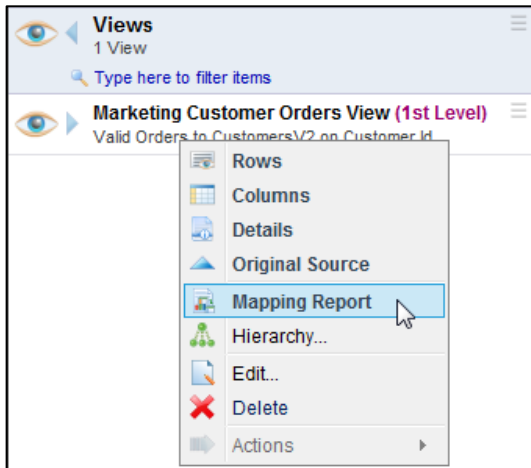
Mapping Report Example

The Mapping Report will typically contain the following information (which will be discussed further in a worked example in the next section):

- A list of **Source Input Tables** (the source tables involved in producing the target table, including lookup tables)
- A list of **Source Table Relationships**
- A diagram illustrating the source table relationships, including join type and cardinality
- A list of **Source Column Definitions** (the name, data type, description and mapping usage of the source columns for all the source input tables)
- A more detailed table for each source input column, which lists various statistics about each of the columns
- A list of **Target Column Definitions** for the target table (i.e. the derived or saved table that the Mapping Report was produced for). The target column definitions information includes the target column name, description, data type, mapping type and the source columns that were used to derive the target column.

The Mapping Report viewer provides facilities to selectively enable/disable sections of the report as well as exporting the report to HTML.

Mapping documents as a rule are constrained to a single set of inputs feeding an output. Experian Pandora will traverse multiple levels of transformation to work out the shortest path from the original inputs to the final target. It cannot, however, traverse certain steps

that fundamentally change the data, and will instead break up the mapping into multiple logical steps. This is analogous to the steps you would have to implement in real life and provide the maximum clarity and understanding in the specification. In complex examples this will result in a **Mapping Book** being created with multiple referenced Mapping Documents.

The following sections describe the key parts of the Mapping Report in further detail with reference to the previous worked example. The document is presented much like a hyper-linked web-page with a navigable index that is hidden by default to quickly navigate between sections. This report can be exported as a HTML web page and behaves identically in terms of look and navigation.

Source Input Tables

The first section of the mapping report summarises the source input tables. In the example we used the following source input tables to generate our target table:

- Purchase Order Detail
- Purchase Order Header
- Customers (v1)

These can be seen listed in the **Source Input Tables** section below:

Mapping Report - Martketing Customer Orders View (View)

Show Contents Page   ← →   Options   Save As   Refresh   Take Screenshot   + Expand All   — Collapse All

Mapping Report

Name	Martketing Customer Orders View
Creator	Administrator
Created	19-Jun-2015 11:18:36
Version	5.0.0
Notes	This is user defined document text

1 Schema

This section provides details on all of the source tables that are inputs to this specification. Details of the structure of the tables are provided along with quality profiles of each of the columns that are involved in the data mapping. Also included in this section are details of the relationships (joins) between the source tables and a diagram that clearly explains the relationship between the source tables and how they are combined to produce the target table.

1.1 Source Input

Name	Original Name	Description	Columns	Source Type	Percentage Mapped	Schema
Customers (v1)	Customers (v1)		14	File	100%	Default
Valid Orders	Valid Orders		14	File	100%	Default
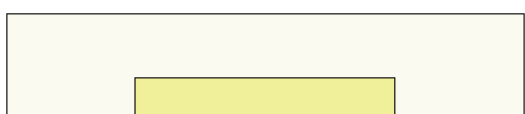
1.2 Source Table Relationships

Join Id	Left Table	Left Column	Right Table	Right Column	Join Type	Cardinality
1	Customers (v1)	Customer Id	Valid Orders	Customer Id	Inner Join	M:M

1.2.1 Definition for Join Expression Column *Customer Id*

Source Columns	Customers (v1).Customer Id
Source Tables	Customers (v1)
Transformation Type	Transformation

Transformation	**After(**
	    Input_Value = Customers (v1).Customer_Id,
	    Search_Value = '-',
	    Ignore_Case = 'FALSE'
	)

1.3 Source-To-Target Mapping Diagram

The lookup table is also mentioned - Discount Code and a second tab, provide a summary report about that table.
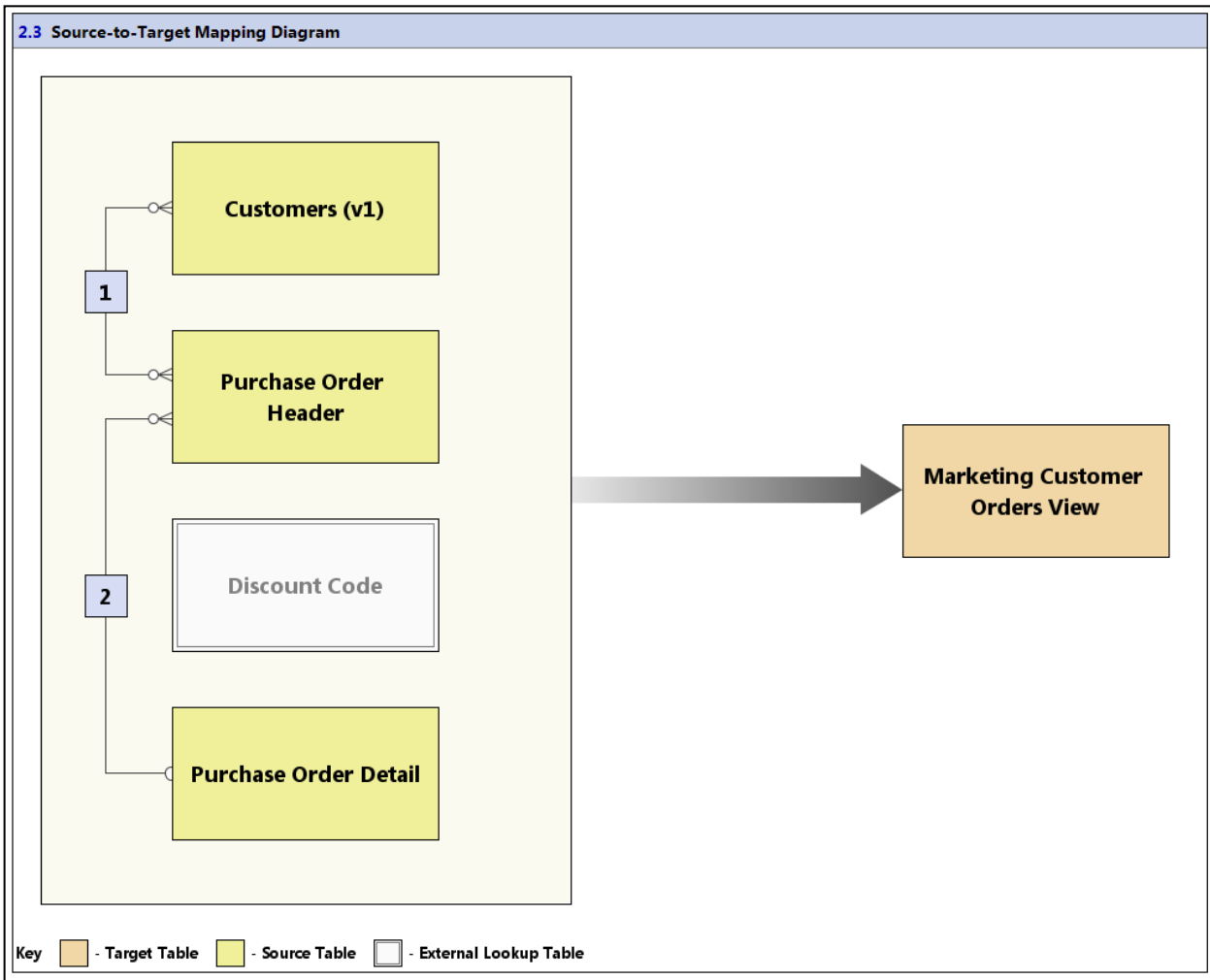
Source Table Relationships

The relationships between the source tables are summarised in the "Source Table Relationships" section of the mapping report. This includes the specification of the transformation of the join key.

2.2 Source Table Relationships						
Join Id	Left Table	Left Column	Right Table	Right Column	Join Type	Cardinality
1	Purchase Order Header	Customer Id	Customers (v1)	Customer Id	Inner Join	M:M
2	Purchase Order Detail	Order Id	Purchase Order Header	Order Id	Inner Join	N:0

2.2.1 Definition for Join Expression Column *Customer Id*	
Source Columns	Customers (v1).Customer Id
Source Tables	Customers (v1)
Transformation Type	Transformation

Transformation	**After(**
	    Input_Value = Customers (v1).Customer_Id,
	    Search_Value = '-',
	    Ignore_case = 'FALSE'
	)

Source-to-Target Mapping Diagram

The Source-to-Target Mapping diagram illustrates the relationships between the source input tables.

In the above screenshot extract, we can see that the diagram shows the two relationships visually along with the join cardinality and the inclusion of a lookup table. The relationships use traditional ER diagram notation to show 1:M cardinality and optionality.

Source Column Definitions

The next section in the mapping report summarises the layout of each source table and provides source column definitions. The **Usage** column in the table identifies how each source column has been used.

Columns that were not used in the final target table are indicated as having "No Mapping" in the **Usage** column.

2.4 Source Column Definitions for *Purchase Order Detail (83% mapped)*

Index	Name	Original Name	Datatype	Nullable	Usage	Description
1	Order Id	Order Id	Integer (7)		Mapping Joining	
2	Product Id	Product Id	Alpha Numeric (10)		Mapping	
3	Price	Price	Money (4,2)		Transformation Mapping	
4	Quantity	Quantity	Integer (4)		Transformation Mapping	
5	Discount Code	Discount Code	Alpha Numeric (1)		Transformation	
6	Shipping Date	Shipping Date	Date		No Mapping	

The report then details each line from the source that is used in the Mapping.
Example:

2.5 Profile for Source Column *Purchase Order Detail.Order Id*

Statistic	Value
Datatype	Integer
Values	452,095 duplicates out of 526,374 values (14.11% unique)
Minimum	1076579 (15 occurrences)
Maximum	1150876 (2 occurrences)
Least Frequent Value	1079210 (Only one occurrence)
Most Frequent Value	1110879 (26 occurrences)
Constant Format Pattern	9999999
Precision	7
Scale	0
Constant Length	7
Standardisation Options	Auto-parse Integers, Auto-parse Decimals, Leading Zero as Number, Auto-parse Scientific Numbers, Treat Blanks as Nulls, Remove Leading Blanks, Remove Trailing Blanks
Source Transformation	**Trim(** Order_Id **)**
Usage	Joining, Mapping

Target Section

The Target section contains the detailed transformation specification for every target column. This can be implemented either by hand coding or using any data integration product. Intermediate results used to derive final values but which are not themselves required in the target are included as *virtual columns*, for example if a transformation was used in a join but not visible in the output.

Target Column Definitions

The Target Column Definitions section lists all the columns of the target table layout we mapped to.

The **Sources** column in this section lists the source columns that were used to derive the target column. In the example above most of the target columns originate directly from a source column.

The **Telephone** target column (which we transformed with an expression) is shown as being derived from a transformation, as is the **Customer Number** and **Net Value** columns.

3 Target

This section provides details on the target table that is the output of this mapping specification. A target table is created inside the database server after any number of development steps. Details of the structure of the table are provided along with a quality profiles of each of the actual data in the target as modelled inside the database.

3.1 Target Column Definitions for *Marketing Customer Orders View (100% mapped)* ⌃

Index	Name	Original Name	Datatype	Nullable	Sources	Mapping	Description
1	Customer Id	Customer Id	Alpha Numeric (7)		Purchase Order Header.Customer Id	Direct Mapping	Valid Orders.Customer Id Purchase Order Header.Customer Id
2	First Name	Forename	Alpha Numeric (16)	Yes	Customers (v1).Forename	Direct Mapping	Customers (v2).Forename
3	Surname	Surname	Alpha Numeric (25)	Yes	Customers (v1).Surname	Direct Mapping	Customers (v2).Surname
4	Company Name	Company Name	Alpha Numeric (40)	Yes	Customers (v1).Company Name	Direct Mapping	Customers (v2).Company Name
5	Address Line	Addressline	Alpha Numeric (40)	Yes	Customers (v1).Addressline	Direct Mapping	Customers (v2).Addressline
6	City	City	Alpha Numeric (28)	Yes	Customers (v1).City	Direct Mapping	Customers (v2).City
7	Country	Country	Alpha Numeric (14)	Yes	Customers (v1).Country	Direct Mapping	Customers (v2).Country
8	Post Code	Post Code	Alpha Numeric (12)	Yes	Customers (v1).Post Code	Direct Mapping	Customers (v2).Post Code
9	Telephone	TELEPHONE	Alpha Numeric (18)		Customers (v1).Country Customers (v1).Telephone	Transformation	
10	Email	Email	Alpha Numeric (42)	Yes	Customers (v1).Email	Direct Mapping	Customers (v2).Email
11	Credit Limit	Credit Limit	Alpha Numeric (9)	Yes	Customers (v1).Credit Limit	Direct Mapping	Customers (v2).Credit Limit
12	Order Id	Order Id	Integer (7)	Yes	Purchase Order Detail.Order Id	Direct Mapping	Valid Orders.Order Id Purchase Order Detail.Order Id
13	Order Date	Order Date	Date	Yes	Purchase Order Header.Order Date	Direct Mapping	Valid Orders.Order Date Purchase Order Header.Order Date
14	Product Id	Product Id	Alpha Numeric (12)	Yes	Purchase Order Detail.Product Id	Direct Mapping	Valid Orders.Product Id Purchase Order Detail.Product Id
15	Price	Price	Alpha Numeric (14)	Yes	Purchase Order Detail.Price	Direct Mapping	Valid Orders.Pricey Purchase Order Detail.Price
16	Quantity	Quantity	Integer (4)	Yes	Purchase Order Detail.Quantity	Direct Mapping	Valid Orders.Quantity Purchase Order Detail.Quantity
17	Discount Rate	DISCOUNT_RATE	Integer (2)	Yes	Discount Code.Discount Code Discount Code.Rate Purchase Order Detail.Discount Code	Transformation	
18	Net Value	NET_VALUE	Alpha Numeric (14)	Yes	Discount Code.Discount Code Discount Code.Rate Purchase Order Detail.Discount Code Purchase Order Detail.Price Purchase Order Detail.Quantity	Transformation	

Row Transformation Criteria

This section is present if anything was done to transform the number of rows. In this case, we have a filter applied to remove rows where Telephone is not null.

3.3 Row Transformation Criteria

Filter	(Telephone **IS_NOT_NULL**)

Mapping Definitions for Target Columns

For each column in the Target Column Definitions section, there is a link to a pseudo-English transformation detailing what was done to the data to create the target column in question. If the target was composed of many different steps then all functions used would be amalgamated into a single expression at this point.

Here are two examples of the transformed columns and their Mapping Definitions:

3.3.1 Mapping Definition for Target Column *Telephone*	∧
Source Columns	Customers (v1).Country Customers (v1).Telephone
Source Tables	Customers (v1)
Transformation Type	Transformation

Transformation	(Case
	  When
	    Customers__V1.Telephone CONTAINS_FORMAT 'A' OR Telephone
	LENGTH_NOT_BETWEEN 11 AND 13
	    Then Null_Value
	    Else Concatenate(
	      First_Value = Convert_Iso(
	        Input_Value = Customers (v1).Country,
	        From = 'country',
	        To = 'dial'
	      ),
	      Additional_Value = Text_Transform(
	        Text_Value = Denoise(
	          Input_Value = Customers (v1).Telephone,
	          Remove_Vowels = 'FALSE',
	          Remove_Digits = 'FALSE',
	          Remove_Whitespace = 'TRUE',
	          Remove_Alphabetic = 'FALSE'
	        ),
	        Transformation = 'ZTRIM_LEADING'
	      )
	    )
	End)

3.3.2 Mapping Definition for Target Column *Discount Rate*	∧
Source Columns	Discount Code.Discount Code Discount Code.Rate Purchase Order Detail.Discount Code
Source Tables	Discount Code, Purchase Order Detail
Transformation Type	Transformation

Transformation	Table_Lookup_List(
	  Replace_Type = 'ALL',
	  Distinct = 'FALSE',
	  Ignore_Nulls = 'FALSE',
	  Return_Column = Discount Code.Rate,
	  Lookup_Column = Discount Code.Discount_Code,
	  Lookup_Value = Purchase Order Detail.Discount_Code,
	  Default_Value = '0'
	)

The source columns used in the transformations are also detailed.

Document Viewing Options

Clicking the **Options** button on the document viewer toolbar will display a dialog that allows you to specify exactly which sections of the Mapping Report to include, effectively removing them from the document if de-selected:
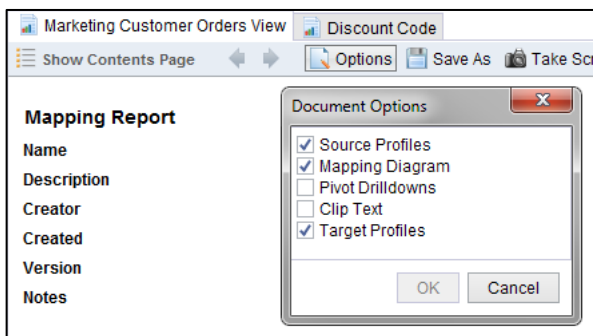
Mapping Annotations

Mapping annotations can be added to target tables or views by adding properties to them. Simply editing the table or view in question and clicking on the **Properties** button in the edit settings panels gives you access to adding properties. Here is an example of adding a Mapping Annotation to the Customers table. These can be used in transformation functions or simply used to provide additional documentation in the mapping report.
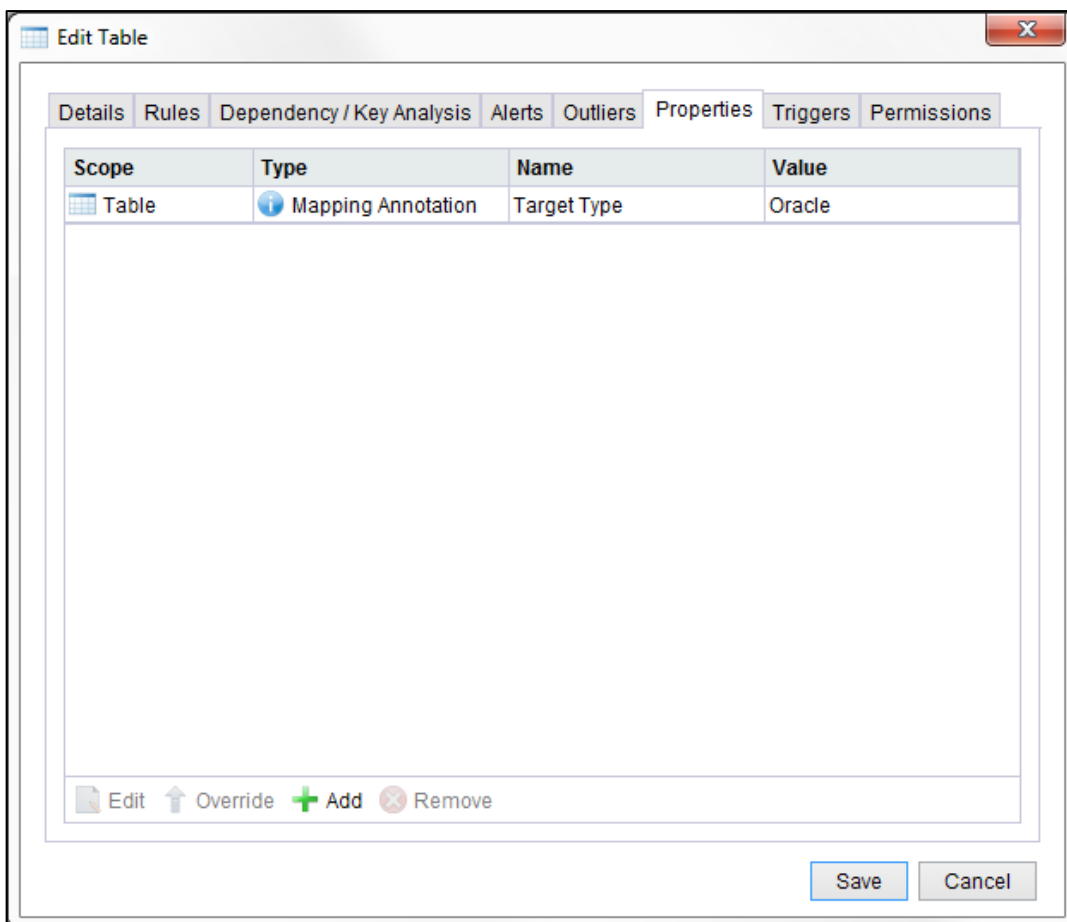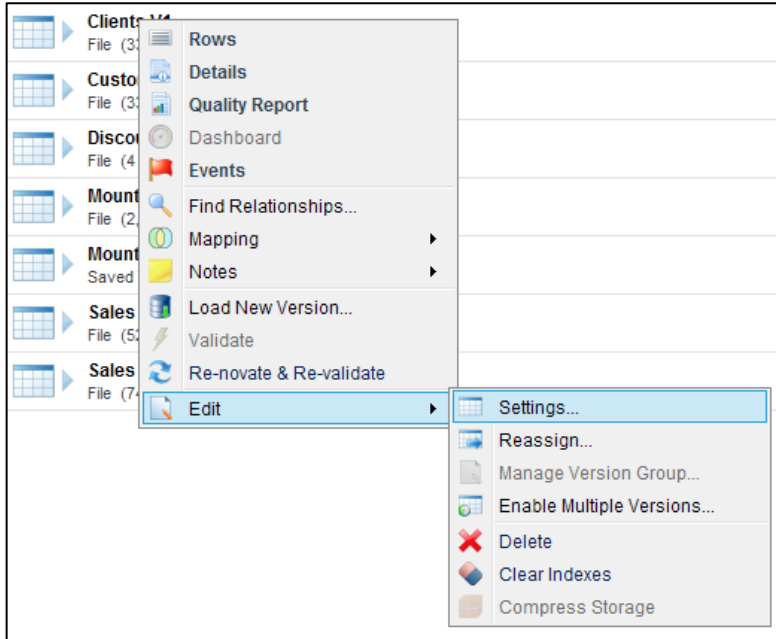
The target section includes a table of mapping annotations. If you are currently viewing the Mapping Report, when you add an annotation, hit Refresh and it will be included in the report.

2.6 Mapping Annotations for *Customers (v1)*		
Index	Name	Value
1	Target Type	Oracle

Navigating the Mapping Report

The Mapping Report, when viewed from the Experian Pandora client, can be navigated in a similar way to a web-page. Any navigable item is underlined and clicking on it will position the document at the appropriate position.

The back/forward buttons on the document viewer toolbar can be used to navigate the position history, again, in a similar way to navigating a web page.

Exporting the Mapping Report

Clicking the **Save As** button on the document viewer toolbar will show a standard file save dialog allowing you to specify the location of the export file. Mapping reports can currently be exported as web pages (HTML) or as a PDF document.

Mapping Report Contents Page

Clicking the **Show Contents Page** button on the document viewer toolbar will show a navigable contents page for the document being viewed, as shown in the screenshot extract below:

The button will now change its name to **Hide Contents Page**. Clicking this will hide the contents page again. Clicking on any item in the contents navigator will navigate the document to that point.

Dynamic Table Union

Experian Pandora provides the ability to dynamically union Tables held in a project. What this does is vertically stack the Tables within the project into a single drilldown, but the columns in that drilldown are the superset of all the columns across all the Tables. Where the column names are the same, then data from different Tables appears within the same column. To force data to appear in the same columns, simply rename the columns of any Table to match those of another.

In this example, we load 3 tables into the Project **Union Example**:
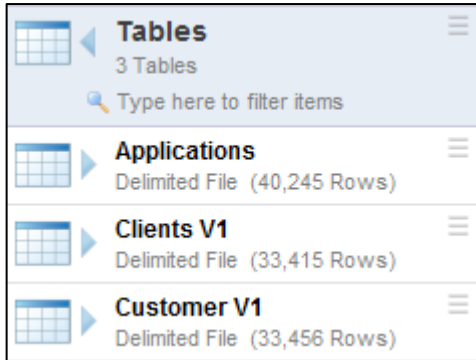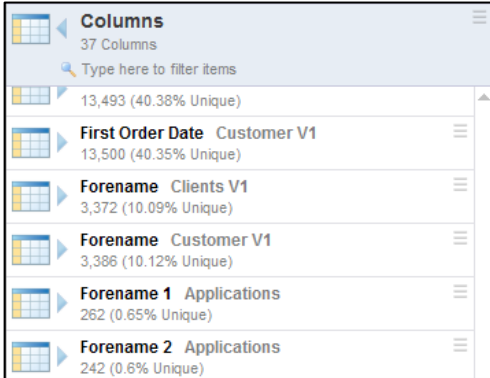
To view the union, simply right-click on the **Tables** and select **Table Union**:
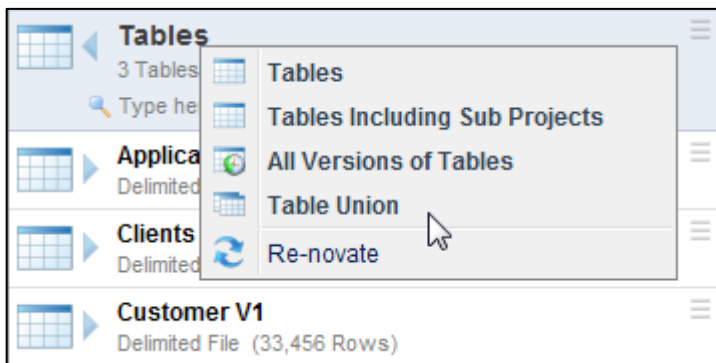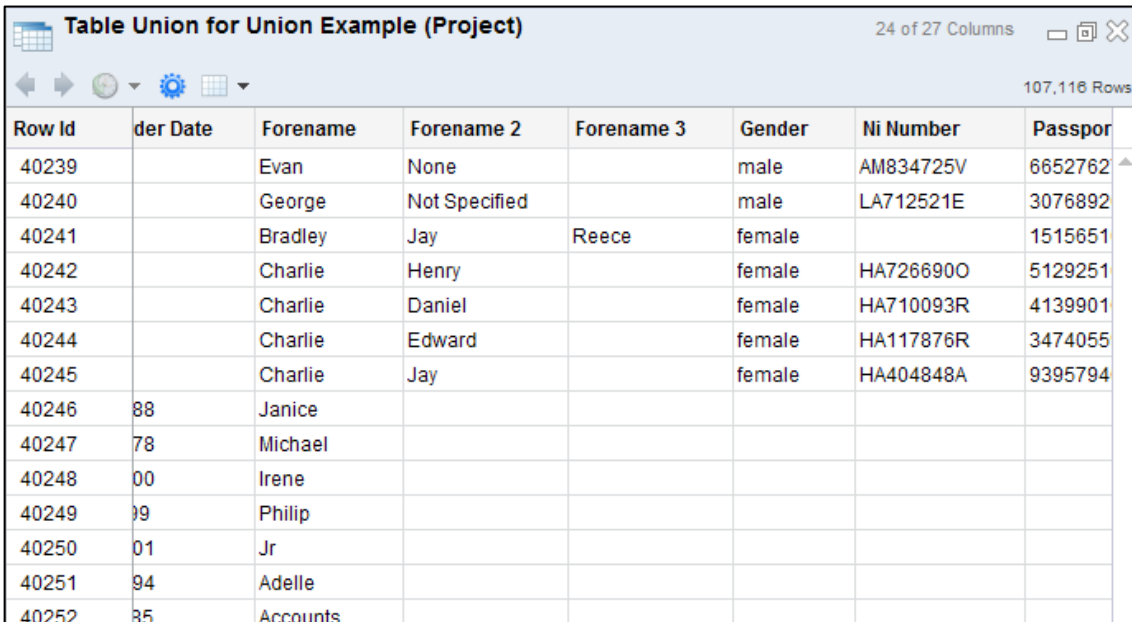
This gives us the Table Union drilldown:

Row Id	e	Forename	Forename 1	Forename 2	Forename 3	Gender	Ni Number
40241			Bradley	Jay	Reece	female	
40242			Charlie	Henry		female	HA726690O
40243			Charlie	Daniel		female	HA710093R
40244			Charlie	Edward		female	HA117876R
40245			Charlie	Jay		female	HA404848A
40246		Janice					
40247		Michael					
40248		Irene					
40249		Philip					
40250		Jr					
40251		Adelle					
40252		Accounts					
40253		Paul					
40254		Nigel					
40255		Reginald					
40256		Roger					

As you can see, we have strategically positioned the drilldown at the point where it transitions from one table to another. We can see clearly that the data in **Forename 1** really is identical to that in **Forename**. To coerce them together, simply rename Forename1 to Forename. We do not even have to think about which Table that comes from as we can do this from the list of columns across all tables.

Simply right-click on the **Forename1** column and select **Edit** to change its name. Depending on your user preferences this may produce a warning message – don't panic!   Re-issuing the union (which is always an instant operation), will show the result and we can now see the Forename1 column has vanished, with the data now in the correct position, namely the **Forename** column.

Table Union for Union Example (Project) — 24 of 27 Columns

107,116 Rows

Row Id	der Date	Forename	Forename 2	Forename 3	Gender	Ni Number	Passpor
40239		Evan	None		male	AM834725V	6652762
40240		George	Not Specified		male	LA712521E	3076892
40241		Bradley	Jay	Reece	female		1515651
40242		Charlie	Henry		female	HA726690O	5129251
40243		Charlie	Daniel		female	HA710093R	4139901
40244		Charlie	Edward		female	HA117876R	3474055
40245		Charlie	Jay		female	HA404848A	9395794
40246	88	Janice					
40247	78	Michael					
40248	00	Irene					
40249	09	Philip					
40250	01	Jr					
40251	94	Adelle					
40252	35	Accounts					

Note that columns in the union drilldown are presented in alphabetic order for convenience. You can now manipulate this data using all of the standard transformation and manipulation facilities.

Automatic Novation of a Table Union

If the Table Union has been derived from all of the **Tables** under a Project's folder, then saved as a drilldown, any future loads of the Tables in the Project will automatically update the drilldown to refer to the newest versions.

To generate a drilldown that updates in this way, right-click on **Tables** in the Explorer under a project and select **Table Union**.
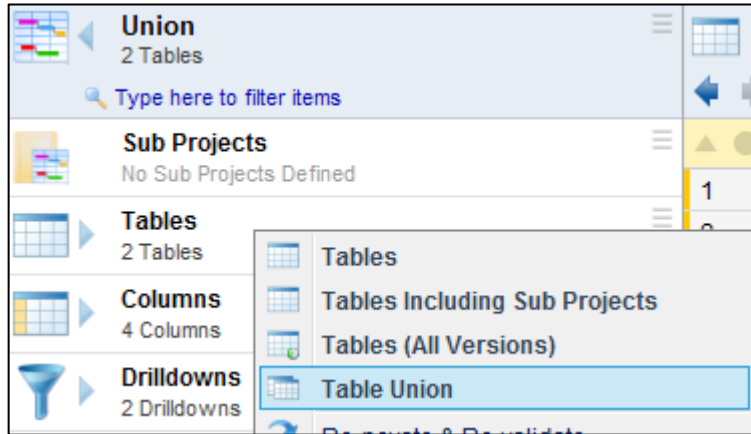
Limitations of Table Union

Currently the table union is not supported in the generated mapping specification.

Table Cell Transformations

Experian Pandora is unusual in providing functions that can operate on cells (row/column positions) within a drilldown. This is because the data is held in a grid format not unlike a spreadsheet.

The functions belong to the Drilldown function library and provide a wide diversity of capabilities including the ability to do running sums and averages over a range of records, or even reference cells at a specific location.

Care should be taken in using these capabilities. If you are using the output data directly then these are valid functions to use, but where you intend taking the mapping specification forward to a different implementation technology you may find that that technology does not have a comparable capability.

Using a Table as both a Source and Target

When a mapping is evaluating from a target table to produce a Mapping Document, it attempts to traverse all the way through to the original sources. In some cases, you do not want this to happen.

An example of this is modelling a data movement that goes via an intermediary step, such as a staging area. Initially you may have created the target staging tables, and produced the mapping specifications for those. You then move on to producing the actual targets, mapping from the staging tables. When you come to generate the mapping specification, Experian Pandora automatically traverses **through** the staging tables back to the original source tables, which does not result in the intended mapping specification.

To change this behaviour, you can disconnect the mapping hierarchy behind the staging tables. This removes all references to all objects used in the creation of the staging table, so that when you generate the mapping document for the final target, it stops its traversal at the staging table.

Hierarchy disconnection is temporary, providing the original objects that were related are not themselves deleted (i.e. the original source tables etc.) You can **reconnect** the hierarchy at any time. To disconnect the hierarchy for a saved table, right-click on it and select **Mapping > Disconnect Hierarchy.** To reconnect it at a later stage, right-click on the table and select **Mapping > Re-connect Hierarchy**.
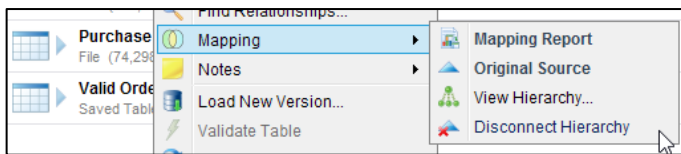
Other reasons for disconnecting hierarchies

The other reason for disconnecting a hierarchy is that Experian Pandora will not allow you delete the sources used to create the saved table if the hierarchy is still connected. You may have simply cleansed an original source table, and no longer wish to keep the source. Simply disconnect the hierarchy for the saved cleansed table, and you will be allowed to delete the original source table.

Data Integration

Having built the necessary data on-screen with Experian Pandora you may wonder if you can *just use it*. The answer is *sometimes*. Remember that Experian Pandora does not offer transactional or batch processing. Experian Pandora is used to prototype repetitive data integration tasks which are then implemented using, for example, an extraction and transformation tool (ETL).

In data integration projects however, especially in data migration, there are many tasks which are **one-off** data movements, such as non-volatile reference data. An example of this is the preparation of a consolidated list of suppliers from two different sources.
In these situations it is possible to simply export the data prepared by the analyst in Experian Pandora to a file and use it in the production data integration process.

Lookup tables, translation tables and reference tables are also good candidates for this.

Exporting Drilldowns

The majority of Drilldowns can be exported to a variety of different formats. For a user to be able to export, they must have the **Export** capability in their User Role. To Export to your PC use the **Save As > File** option or to Export to the Server or a JDBC Connection, use the **Save As > Target** option.

Save as File

In any Drilldown, click on the Drilldown Action menu and select **Save As > File** to save to your PC.

The export dialog has the following options: (NB: these will depend which file format you have chosen to export to)
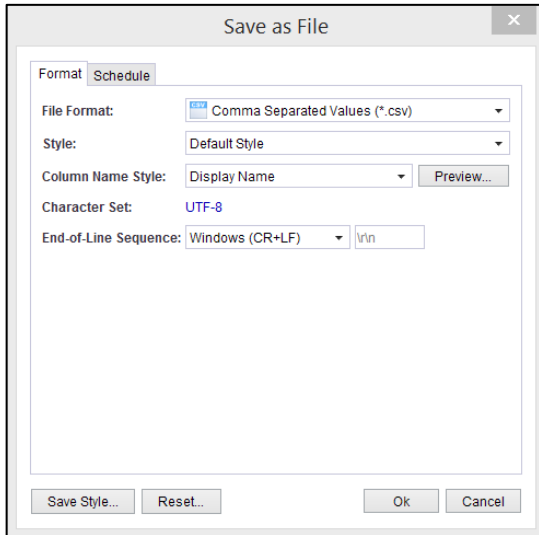
File Format
The export file format you wish to export to.

Style
If you are going to be exporting a number of tables with a particular style, once you have created the style select Save Style from the bottom left hand corner, name it and it is now available as an option for re-use.

Force Quotes?
If you chose Comma, Tab or Pipe delimited formats then you can choose whether you wish to force quoting of fields for output as opposed to the normal automatic quoting of fields that occurs when a field contains the chosen delimiter.

Quote Character
Allows you to choose the quote character to use when exporting to Comma, Tab or Pipe delimited outputs.

Pivot Table?
This is only relevant to HTML and PDF outputs. When set to Yes the output will be shown in the exported file in blocks of records with a three column output. Each block is equal to a single row in the original Drilldown. Each row within that block is equal to a column in that row. The first output column is the Row Id. The second output column (Column Name) contains the name of the Drilldown column and the third (Value) has the value of that Drilldown column on that row. This enables to output very long rows on a format that is fixed in width (for example, a PDF file). Set this to No to have it exported in the normal layout as seen in the Drilldown. Automatic makes a decision based on the width of the drilldown.

Column Name Style
When exporting delimited files, this is the style that will be used to generate column names on the first row. Selecting NONE will suppress the exporting of column header information in the output file. See Metatdata data export below.

Character Set
Specifies the character encoding to use when exporting delimited files.

End of Line Sequence

Specifies which character defines the end of line.

Metadata Export

When exporting a file without headings, having selected NONE in the column name style you can now select the Metadata style to export the heading in a DDL. , .xml or .json format. This will be saved with the exported file as a separate file.

Delimiter

The character to use to separate values in the exported file.

Changing Export File Path

For security reasons, you cannot chose server filenames that result in writing to unsolicited areas of the server. By default, all exports go to the EXPORT folder which can be set by your administrator. If you specify an absolute path (e.g. c:\somewhere\docs) in the filename it will strip the absolute part of the filename and make it relative to the export directory (e.g. somewhere\docs) and automatically create the sub directories in the path.

Clip Text?

If selected the text will be clipped/cut to fit into the cell. If not selected the full text will be exported.

Reset Button

If you have made changes to the option you can reset to revert to the default using this button.

Schedule
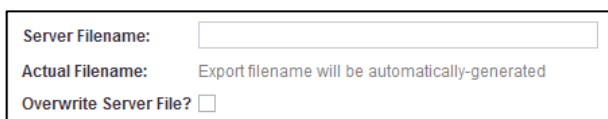
Exports are executed by the Scheduler as Jobs so may not happen immediately. This can happen in a different session if the export happens to be of a large data set which will take time to complete. A popup box will appear on the desktop to notify you that the export the job has completed if you are still logged in.

You can choose when you want to run the Export by opening the Schedule tab and selecting **Later**. This will give you a time option and a calendar.

Save as Target

In any Drilldown, click on the Drilldown Action menu and select **Save As > Target** to save to the Server. Files will be saved to the configured server exports location on the server. This location is defined in the **Server Exports** server setting, which can be found in **Home > Settings > Server Settings > Storage**. See the section below on saving to JDBC connections.

The dialog box will contain all the options available for Save as File along with the following :

Server Filename:	
Actual Filename:	Export filename will be automatically-generated
Overwrite Server File?	☐

Server File Name

This is optional. If you wish to save the file under a specified name enter the name in here. The file path will be displayed in the Actual File Name below.

Actual File Name

If you have not selected a Server File Name Experian Pandora will automatically generate one. It will be in the format of "export_USERNAME_Rows_for filename_20140512112013.csv". The numbers are the date and time that the file has been generated.

Overwrite Server File?

If you are saving a file with a specified name and wish to overwrite a previous version tick this option.

Changing Export File Path

For security reasons, you cannot chose server filenames that result in writing to unsolicited areas of the server. By default, all exports go to the EXPORT folder which can be set by your administrator. If you specify an absolute path (e.g. c:\somewhere\docs) in the filename it will strip the absolute part of the filename and make it relative to the export directory (e.g. \data\export\somewhere\docs)

and automatically create the sub directories in the path. To set a custom export location on the server, refer to the section "Custom Export Locations" below.

Special Export Locations

If you specify a relative path it will treat it as relative to the export directory unless the first path element is:

PUBLIC_IMPORT
The public directory for data loading on the server for all users to access.
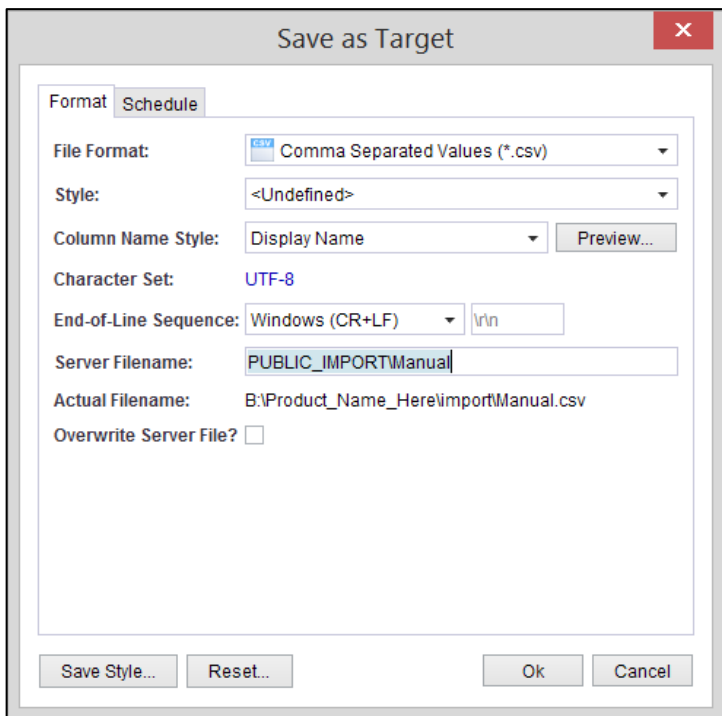
USER_IMPORT
Your private import directory for data loading on the server.

TEAM_IMPORT
The import directory for your Team for data loading on the server.

CONTENT
The CONTENT directory used for importing content for Domains. This allows you to export data and immediately re-import it as a Domain.

Custom Export Locations

Administrators can sanction exporting to specific directories on the server by creating additional keywords in the **paths.properties** file, which should be saved in the database root directory on the Server. The database root location is defined in the **Database Root** server setting, which can be found in **Home > Settings > Server Settings > Storage**. The paths.properties will need to be created manually before it is first used.

The file is a simple 'keyword=value' file, of the format:

> **myplace = c:/somewhere/over/therainbow**
> **myotherplace = c:/somewhere/else**

You can then use the key word in the Server Filename field of a Save as Target dialog: Specifying **myplace\Manual** will substitute keyword **myplace** with the file path defined in the **paths.properties** file. This access applies to all users so should be enabled with care.
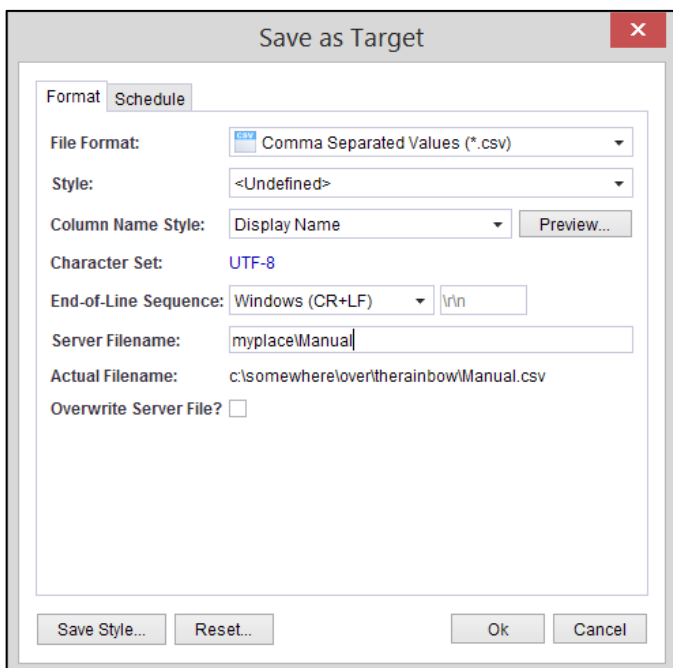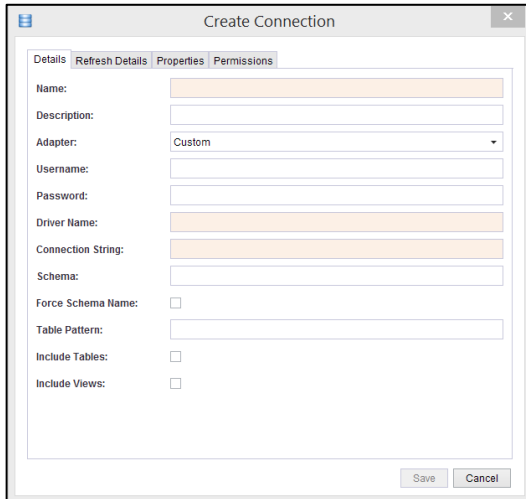
JDBC Exports

Virtually any tabular data can be exported via a JDBC connector to a remote database table. First you will need to create a connection to the remote database.   If you already have a connection in place please continue to Export to Target.

Creating a New Connection

Navigate to **Home > Connections**, right click on the required **Connections** and **New…**

When you specify the "**Connection String**" you should also ensure that the appropriate JDBC driver Jar file has been placed in Experian Pandora \drivers\jdbc\

Creating a Target

Navigate to **Home > Connections** then right click on one of the remote tables and select **Save as Target…**
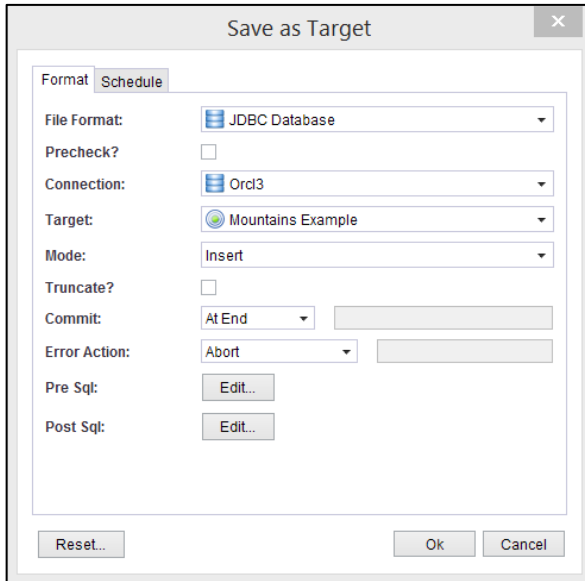
The **Create Target** dialog has two tabs, **Details** and **Permissions**. In **Details** enter the name and a description for the table.   Specify the **Permissions** for this table on the Permissions tab.

The remote table will now have a target icon:
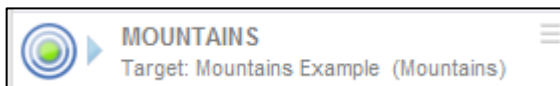
Export to Target

In the drilldown window of the data you wish to export, click on the Drilldown Action menu and select **Save As > Target.**

File Format
Select JDBC Database from the drop down selection list. The JDBC export dialog has the following options:

Precheck?
If you tick the **Precheck** box all the data will be checked prior to starting the export. This should be used in conjunction with the **Error Action** which defines the behaviour of the export when an error is encountered (see Error Action). The data checks for every cell are:

- If a Primary key(s) has been defined for the target then check that the data in the primary key column(s) is unique and not null
- If a target column is not nullable check that the data is not null
- Check that the source data type, i.e. Experian Pandora data type, is compatible with the target column's data type.
- For numeric datatypes, check that the remote column's scale and precision is at least as large as the remote columns scale and precision. For Alphanumeric data check that the column width is large enough.

Connection
This allows you to choose which data source you will be exporting the data to.

Target
The **Target** selection box is populated with Remote Tables associated with the selected **Connection.**

Mode
The **Mode** selection box is used to select how the data is inserted into the target:

> **Insert** (see also Truncate)
> Mapped columns will be exported to the database. It is not necessary to map all source or all target columns. However, if any target column is not nullable then it must be mapped since unmapped target fields will be set to null. Warnings are presented if source or target fields are not mapped but they do not prevent further execution.
>
> The equivalent SQL statement is:
>
> "INSERT INTO tablename (column1, …) VALUES (value1, …)"

Update

The mapping dialog will show which source columns are mapped to which target columns and which ones have been specified as keys.   You can change the mapping and keys at this stage by right clicking on the relevant target column name.

A key field (or fields) must be specified so that the target records can be identified. It is not necessary to map all fields and only mapped target fields will be updated.

The keys referred to are those columns that will appear in the equivalent SQL statement's WHERE clause:

 "UPDATE tablename SET (column1=value1, …) WHERE (columnName=value, …)"

Delete

It is only necessary to map the key field(s). Any further mapping has no effect.

The keys are the column names used in the WHERE clause:

"DELETE FROM tablename WHERE columnName=value, … "

Insert (or Update if Insert fails)

If a row insert fails then the statement will be retried as an update.

Update (or Insert if Update fails)

If the update fails then the statement will be retried as an insert.

Create Table

Selecting this option will automatically create a table in the remote database. If you select this mode you will need to enter the table name in the TARGET selection box.

Target:		▼
Mode:	Create Table	▼

The table name will be validated against the data source i.e. you may only type in valid identifiers for the database vendor which conform to DB identifiers. In addition there is no support for creating primary keys on the table; however you may use the pre sql edit panel to alter the table column characteristics to specify a primary key prior to inserting the data.

If the Mode is INSERT, UPDATE or DELETE then these statements will be batched and executed every commit interval. This is known as homogeneous batching. If no commit interval has been defined then the batch is executed every 1000 statements.

Truncate?

If the mode is **Insert** you may truncate the remote table before applying any Inserts.

Commit

You may commit the statements either at the end or in batches, also see Error Action.

Commit at End

Any errors encountered during the export will result in execution termination and no changes will have been applied to the target.   This is not recommended since you risk losing all the export that has taken place so far and will have to start again. If you also select Abort at End in the Error Action (see next section) the export will be continuously retried, skipping over failed statements) until it reaches the end.

Commit in batches

You may specify how many statements can execute prior to a Commit, i.e. the batch size. Any statements issued after the last commit will be lost if any errors are encountered. If you commit in batches you must specify the batch size in the text field to the right of the selection box.   This must be a valid positive integer value.

Restartable Commit

This is similar to Commit in batches, however all statements that will be executed have been written to a transaction log file first. If there are any errors during execution, typically these will be Network errors, power outages etc. rather than data errors, then you may restart the export and it will resume statement execution from the last commit onwards. See Restarting failed Exports for further details.

Error Action

An Error Action is typically used when you request data Prechecking, however it may also apply to errors encountered during the actual export of the data.

Abort

The first error encountered during data prechecks will cause job failure. This is the default mode during export as well.

Abort at End

The data precheck will run through all data, reporting all errors in the job's log file, then abort.

Abort After Percent

You may specify a percentage. If the (error count/rowCount)*100 > percentage then abort the export.

Abort After Count

You may specify an absolute error count. If the number of errors exceeds this count then abort the export

Ignore

All precheck errors and data export errors, i.e. statement execution error, will be reported in the Jobs log file and then ignored.

All errors are reported in the Export Jobs log file with a unique number. The client is notified of job failure via a popup.
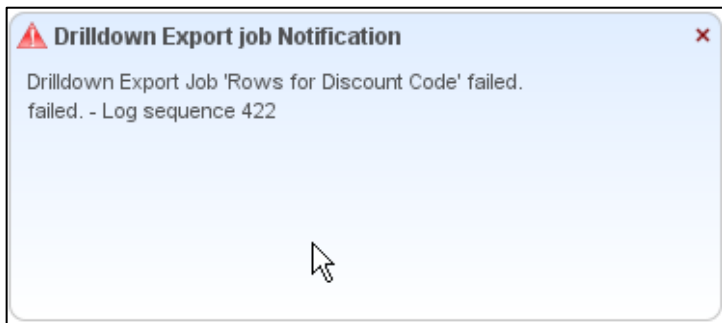
You may view the reason for the failure by navigating to **Home > System > Jobs** and right-clicking on the relevant job number.

Pre SQL

Users may specify a SQL script that will be executed prior to the Export. The Edit dialog box will allow you to import a script file or you may type the SQL statements in by hand. You may also validate the SQL, however validation failures may occur if the SQL uses Vendor specific extensions to standard ANSI SQL statements.

The SQL may contain any valid statements and will be executed as is. Users should ensure they have valid permissions on the database for any statements they intend to execute.

Note that there is no automatic Commit submitted after the script has executed, so if you want the statements committed then you must include a COMMIT statement at the end of your script.
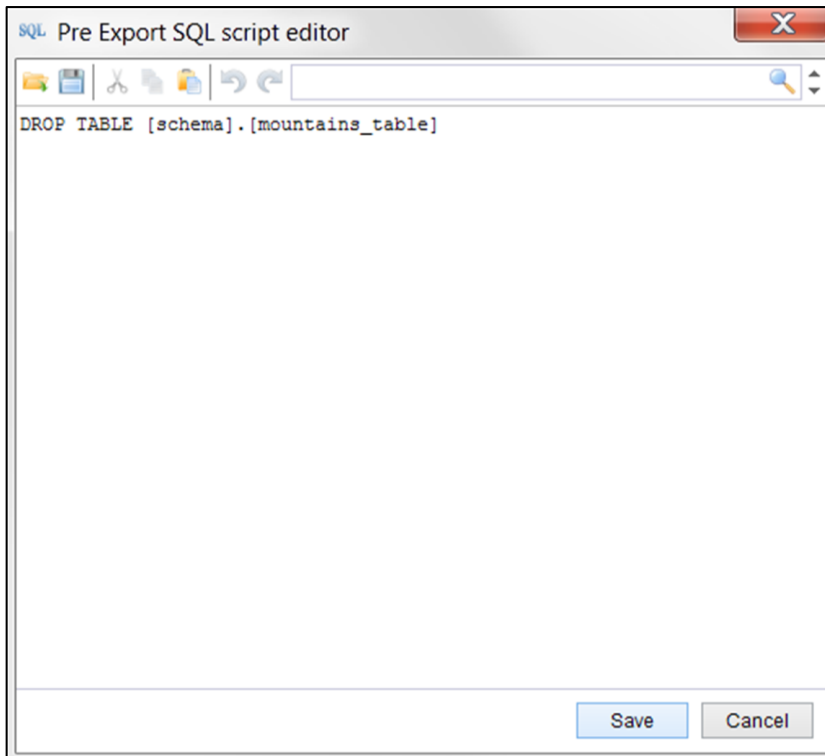
Post SQL

This dialog is identical to the "Pre export SQL" except it is executed after all other statements have been executed.

Once you have selected all the options and you click OK a dialog will open to show the mapping:
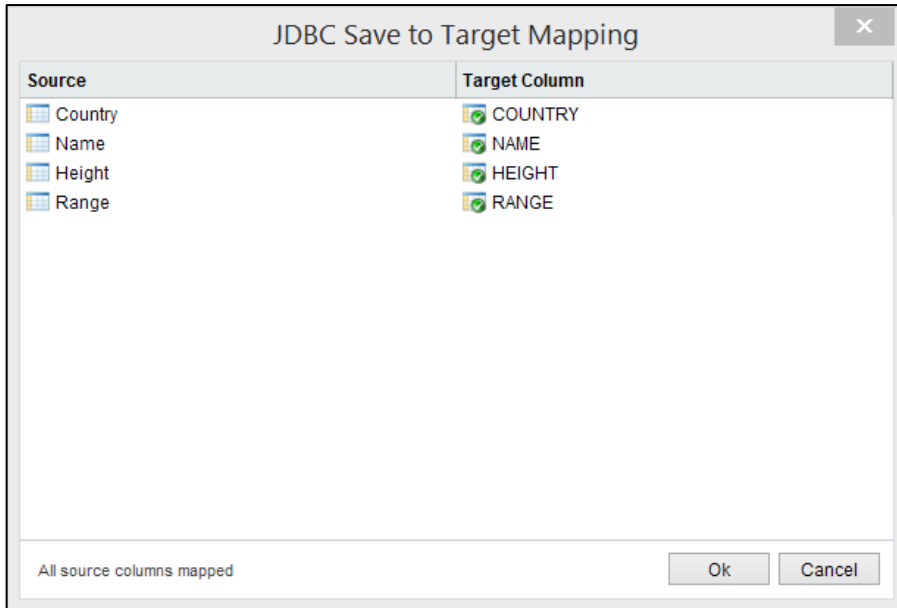
If there are unmapped columns these will be seen in the results:
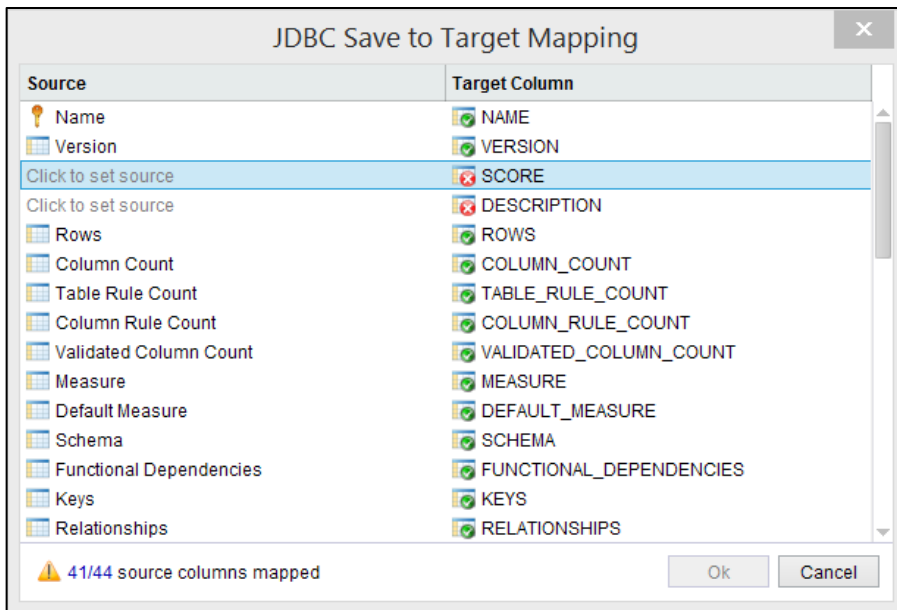
Where columns have not auto-mapped clicking on the **Click to set source** prompt offers you all the remaining unmapped fields. If the field you wanted has been wrongly mapped elsewhere you need to first clear that mapping to make it available to be mapped elsewhere.

Source	Target Column
🔑 Name	🔲 NAME
▦ Version	🔲 VERSION
Click to set source	❌ SCORE
Click to ... ✕ **Select a Source Column**	❌ DESCRIPTION
▦ Row ⟨Search 🔍⟩	🔲 ROWS
▦ Colu ▦ Byte Order	🔲 COLUMN_COUNT
▦ Tabl ▦ Expression Count	🔲 TABLE_RULE_COUNT
▦ Colu ▦ Filter Count	🔲 COLUMN_RULE_COUNT
▦ Valid	🔲 VALIDATED_COLUMN_COUNT
▦ Meas	🔲 MEASURE
▦ Defa	🔲 DEFAULT_MEASURE

The three columns that are available are not related to the two unmapped Target columns and we do not want to export these three columns so we need to go back into the drilldown and hide them before we can complete the Export. Once this has been done all source columns will be mapped and we can select OK to Export the data.

Export Jobs Log File

Every Export Job will have a log file. If the job failed then this log file will contain the reason for the failure and a report on the error types.

A successful export will contain statistics for the job

Timestamp	Message Type	User	State	Task Id	Name	Message
10-Jun-2014 13:46:49	Info	Administrator	Initialising	9137	Drilldown Export	The process is initialising
10-Jun-2014 13:46:49	Info	Administrator	Running	9137	Drilldown Export	Created new table "DCcj"
10-Jun-2014 13:46:49	Action	Administrator	Running	20	Monitor Orcl3	Reconnecting due to parameter changes
10-Jun-2014 13:46:50	Info	Administrator	Running	9137	Drilldown Export	Oracle JDBC driver v12.1 (JDBC 4.1) connected to Oracle
10-Jun-2014 13:46:50	Info	Administrator	Running	9137	Drilldown Export	Exporting view "Discount Code to table "DCcj" (4 rows)
10-Jun-2014 13:46:50	Info	Administrator	Running	9137	Drilldown Export	Starting export of 4 rows starting at row 0
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	The data export has completed
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	4 INSERT statements executed (0 failed)
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	3 COMMIT statements (0 failed)
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	Time preparing SQL statements: 00:00:00.000
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	Time executing SQL statements: 00:00:00.024
10-Jun-2014 13:46:51	Info	Administrator	Running	9137	Drilldown Export	Total time executing Export: 00:00:00.027
10-Jun-2014 13:46:52	Info	Administrator	Completed	9137	Drilldown Export	The process completed successfully

Log Messages for Job 18: Drilldown Export - Rows for Discount Code
Job Status - Completed (10 June 2014 13:46:52 BST)

Restarting a failed Export

This only applies to Exports where the Commit mode was Restartable.

Navigate to **Home > System > Jobs** then right click on the failed job and select **Requeue.**

All Scheduled Jobs
Last Updated on 09 April 2014 16:18:56 BST

Sequence	Name	State	Progress
420	Rows For Di...	...d	0%
419	Rows For D...	...eted	100%
418	Rows For D...	...eted	100%
417	Rows For Di...	...d	0%
416	Rows For D...	...eted	100%
415	Rows For D...	...eted	100%
414	Rows For D...	...eted	100%
413	Rows For Discount Code as "	Stopped	0%
412	Rows For Discount Code as "	Completed	100%
411	Rows For Discount Code as "	Completed	100%

View Log
Start Asap
Requeue
Stop
Delete
Filter

You may view the reason for the failure by either right-clicking on the failed job and selecting "View Log" or double-clicking on the job.

Programmatic exports

The mode of the export may be determined on a per row basis using the values of column data for that row. This option is only visible if you have created a new column on the Drilldown that uses the Expression Builder to create a USE_EXPORT_ACTION final expression.

On the relevant column right-click, select **Insert Column > Custom** to open the Expression Editor. In this example we are going to determine the export action based on the value in the "Discount Code" column. First we provide a Translation Expression that will translate the various values - H, M, L and the default action for any other value. These translations are selected by Export Action expression.
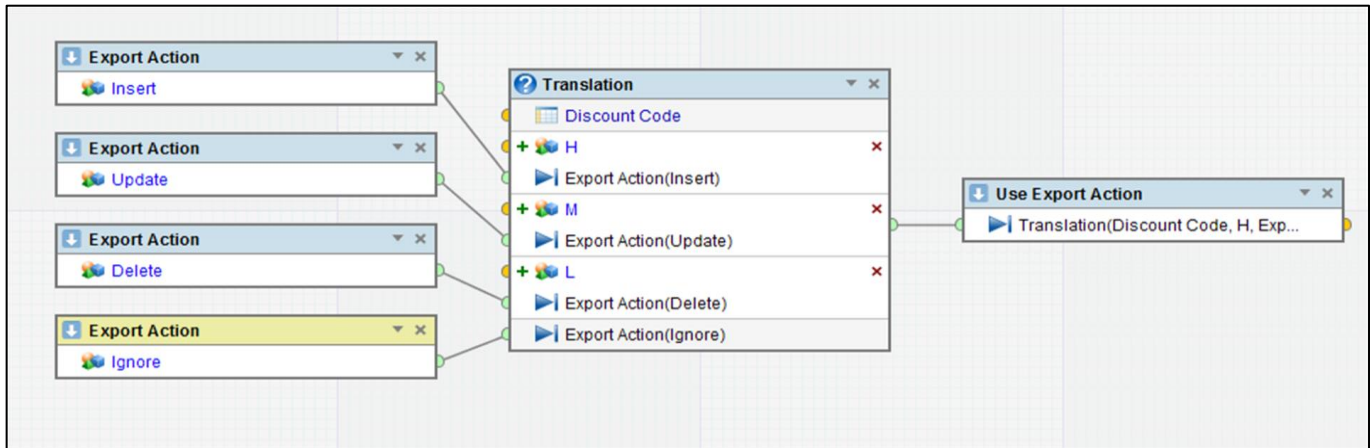
The following "Export Action" selections are available:

- Abort – abort the entire export up to the last good commit
- Delete – delete this row
- Ignore – ignore this row (i.e. no action)
- Insert – insert this row
- Update – update this row
- Insert or Update – Insert this row (or Update it if the insert fails)
- Update or Insert – Update this row (or Insert it if the update fails)

The output of the translation is fed to a **Use Export Action.** Name the Expression Export Programmatic Export and you have a new column containing the export action associated with each row.

Row Id	Customer Id	Discount Code	ƒProgrammatic Export
1	A10-E100750	M	Update
2	A10-E102508	H	Insert
3	A10-E102659	H	Insert
4	A10-E102982	M	Update
5	A10-E103408	M	Update
6	A10-E103489	H	Insert
7	A10-E105402	N	Ignore
8	A10-E105686	N	Ignore
9	A10-E105688	M	Update
10	A10-E106043	N	Ignore
11	A10-E108032	N	Ignore
12	A10-E108768	N	Ignore
13	A10-E108908	M	Update
14	A10-E109669	M	Update
15	A10-E111093	N	Ignore

The Business Glossary

While a common view of a Business Glossary is simple categorisation of data at a business level, Experian Pandora extends that concept to amalgamate all of the aspects of reference, categorisation and enforcement required to define all aspects of the business that require measurement and definition.

The Glossary can be accessed from the top level of the explorer.
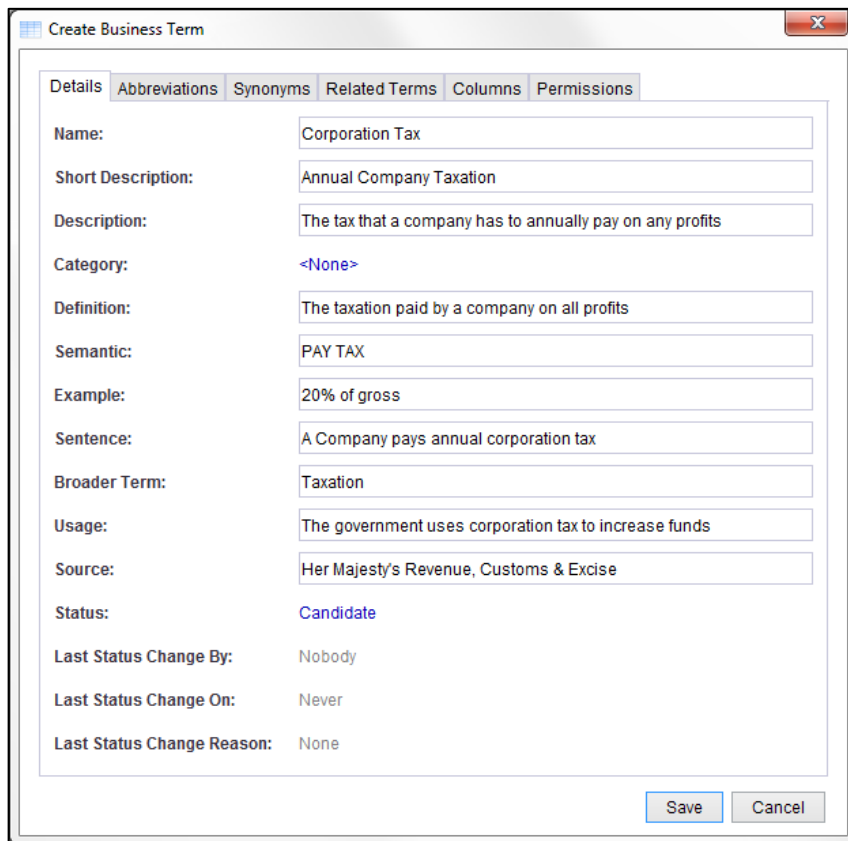
Business Terms

Business Terms are definitions of aspects of the business that provide clear understanding and business meaning to technical data items. They help bring together a myriad of technical column names together in a single business context. They can be optionally linked to Tables and/or Columns of actual data in the Experian Pandora database. Terms can exist outside a category or can be a member of any Business Category, which in themselves can be organised in a hierarchy. Business Terms are a method of expressing business understanding of data which can be abstract or even conflicting when working with a shared repository of information such as that provided by Experian Pandora. Facilities are provided to relate terms together, both to link exact definitions (**synonyms**) and to link closely related definitions (**related terms**).

To create a Business Term, navigate to **Glossary,** right-click on **Business Terms** and select the menu item **New**. The Create Business Dialog window is a multi-tab detailed definition of a Business Term.

Details

The **Details** tab contains basic information about the Business Term:

Attribute	Description
Name	The name of the Business Term
Short Description	A short description to help identify the Term in a list of Terms.
Description	A detailed description of the Term
Category	The Category the Term belongs to (optional)
Definition	A description of the business term, specifying a broad term classification, distinctive characteristics and/or function.
Semantic	The grammar form that this term represents. For example a noun, or a verb.
Example	An instance of the Business Term as it is seen in everyday life.
Sentence	A sample sentence that contains the Term.
Broader Term	General class to which the Term belongs
Usage	Description of how the item being defined is used.
Source	The source of the definition of the Term.
Status	The status of the Term, can be one of: **Candidate** **Authorised** **Replaced** **Archived**

Abbreviations

The **Abbreviations** tab is a list of alternate abbreviations for the Term to assist searching.

Synonyms

The **Synonyms** tab allows you to bi-directionally relate a term with others. Making one Term synonymous with another will result in the other term also being made synonymous. The Available Terms is a searchable list. Simply double click on the terms you wish to associate and they will appear under the Synonyms list. Alternatively select the terms and click **Add**. To remove a Synonym select it from the Synonyms list and click **Remove**.

Related Terms

Related Terms are a one-way relationship from one Term to one or more others. This is for relating Terms that have some association, but could not be considered Synonymous. The operation of this tab is the same as that of the **Synonyms** tab.

Columns

The **Columns** tab allows Columns to be associated with Business Terms. Select a Table, then the Columns from that table to relate and click **Add** to add them to the list of related columns.

When Columns are related to a Term or Category, additional items appear under the Table and Column information denoting the related Terms and/or Categories. Here are examples of those references appearing under both a table and column. Tables are not explicitly referenced; they are implicitly referenced by the related columns linkages.

You can then view the related Terms or Categories by drilling down onto the entries on the Table or Column information drilldowns or expanding the items in the explorer. This provides a subset list of related Terms and/or Categories.

You can also easily see which columns are related to a term. From the above drilldown, simply right-click on the business term in the drilldown (Customer Identifier) and select **Related Columns.** This information about your data landscape is often used to help project and task scoping.

Business Terms can be profiled, meaning that all the data from all the related columns is analysed as if the Term were a single, physical column. Select **Profile** on the right-click menu to view a drilldown which shows each Value, how many times it appears per table, column and schema. Use the right-click menus to further investigate the values across Experian Pandora. Users can determine the global list of values (domain) for a Term across multiple systems, identifying unusual values, finding local favourites etc. One way of describing this functionality is as "Profiling of Semantic Names" and is a massive productivity gain for MDM projects.

Business Categories

Business Categories are containers for Business Terms and enable a comprehensive hierarchy of terms to be created. There is no limit to the number of levels in the hierarchy that are allowed.

To create a top-level Category, go into **Glossary** right click on **Business Categories** and select **New**. The **Create Business Category** window appears. Here you can detail the Name, Short Description and Description for the Category. Click **Save** to create the new top level Category.

To create a sub-category, right click on any existing Business Category and select **New Sub-Category**. The same dialog as above appears with the **Parent** category automatically populated. Alternatively, you can create using the method above and simply select the parent category manually.

Field	Definition
Name	A short, meaningful name reflecting the business subject that the Category denotes.
Short Description	Several words designed to easily locate and differentiate one Category form another.
Description	A long description explaining in detail the purpose of the Category.
Parent	The optional parent category

Importing Business Terms and Categories Definitions

The Business Glossary supports an easy-to-use import facility that can import Business Terms and Categories from the same file, dynamically creating whichever structures are required, and validating the import process prior to final application. To begin the import process right-click on **Glossary** and select **Import**.
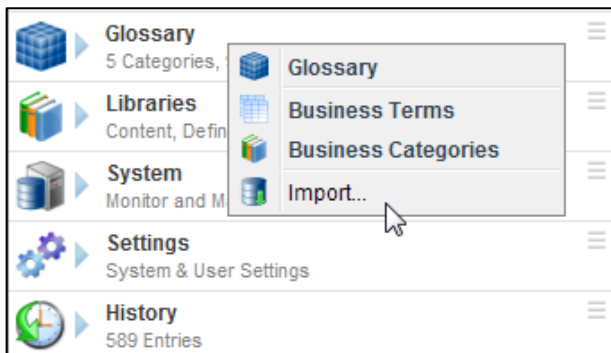
You will be presented with a standard file selector, with additional options. These are for specifying the character set (Code Page) of the import file, whether it ends in Carriage-Return/Line Feed, or just Line-Feed, whether there are headings on the first line and whether just categories should be imported, or all categories and terms.
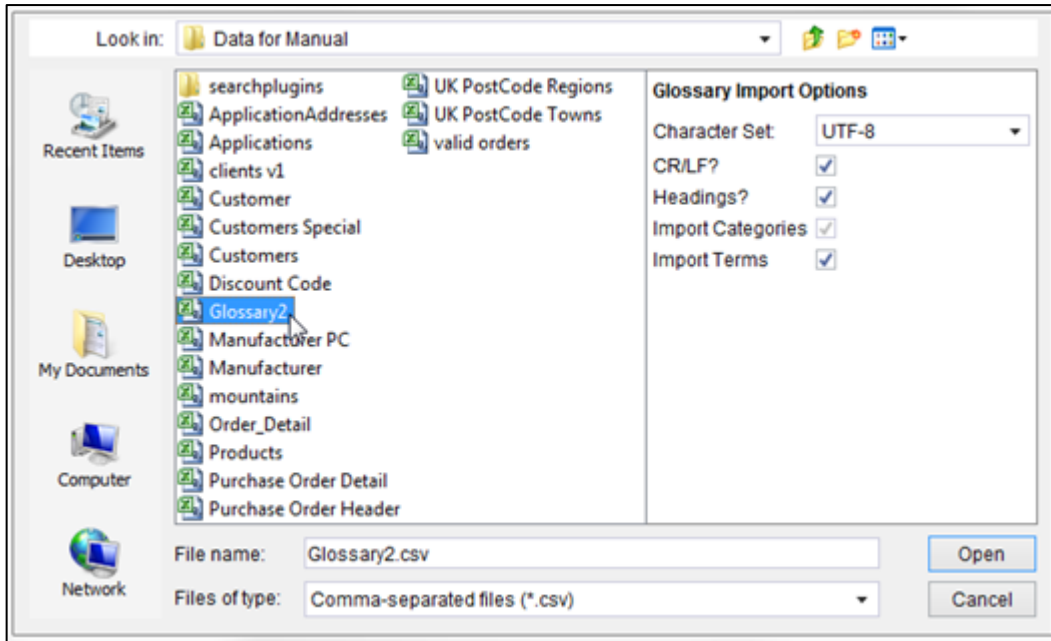
Typically, a Comma-Separated Values (CSV) file will be used as an import source, which can be easily created from Excel. The source file need not have headings, but if it does, Experian Pandora will attempt to match up the headings to standard expected headings. Experian Pandora also provides the ability to dynamically map columns to expected headings, being very flexible in operation.

Each row of the import file must specify a **category** and optionally a **term**.

- The first time a category or term is encountered it is created in the Glossary.
- Subsequent occurrences are regarded as new versions and therefore are potential updates to the existing occurrence.
- Values on new versions replace the values on previous versions, but optional fields which have no specified value are simply ignored (they do not wipe out existing values on existing Glossary definitions).

This example file has a combination of category and term definitions, and a file header line is present.

Category	Category Description	Category Parent	Term	Full Description
Specification	Orderable Automobile Specification			
Paint	Paint Attributes	Specification	Metallic	Metallic Paints
Paint	Paint Attributes	Specification	Simple	Solid Paint
Wheel Type	Wheel Type	Specification	Alloy	Cast or Forged Alloy
Wheel Type	Wheel Type	Specification	Pressed Steel	Simple Pressed Steel
Trim	Interior Trim Type	Specification	Leather	Leather Trim
Trim	Interior Trim Type	Specification	Half	Half Leather Trim
Trim	Interior Trim Type	Specification	Cloth	Full Cloth Trim
Glass	Glass Type	Specification	Tinted	Green or Grey Tinted Glass
Glass	Glass Type	Specification	Plain	Plain Glass

Selecting this file to import presents the import dialog with the understanding Experian Pandora has of what to interpret the fields as. You can simply click on the headings to reassign any of the import fields accordingly, or assign them where there is no header in the import file.
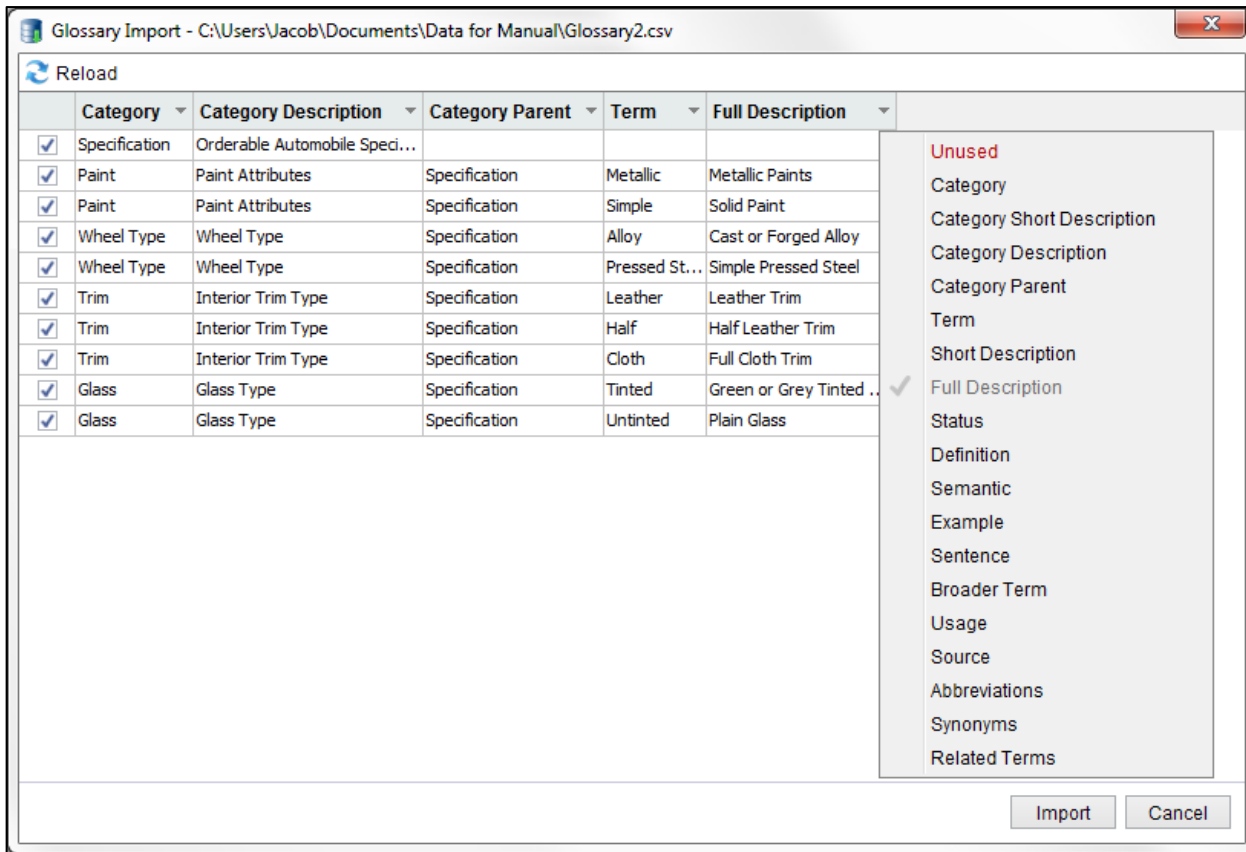
The valid column headings are shown in the drop-down above and are self-explanatory.

Some of the parameters can take comma-separated lists (**Abbreviations**, **Synonyms** and **Related Terms**).

Simple error checking is carried out by Experian Pandora principally to avoid cyclic definitions and errors are signalled. It is possible to de-select rows on the import window by simply unchecking the boxes to the far left.

You should see this pop-up window after a successful Glossary import, providing a summary of how many terms and categories were created:
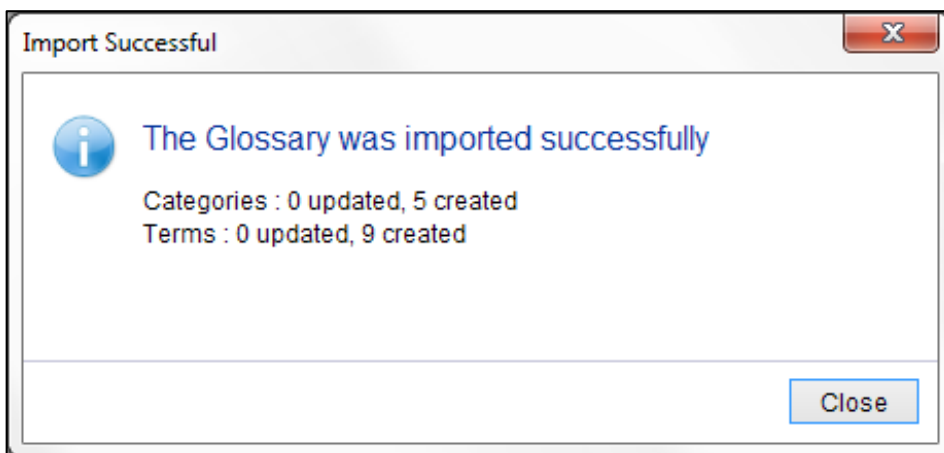
We can now view the imported terms in a drilldown by right-clicking on **Glossary** and selecting **Business Terms** from the menu:

Business Terms

		Name	Category	Short Description	Description
1		Alloy	Wheel Type		Cast or Forged Alloy
2		Cloth	Trim		Full Cloth Trim
3		Corporation Tax	<None>	Annual Company Taxation	The tax that a company has to annual...
4		Half	Trim		Half Leather Trim
5		Leather	Trim		Leather Trim
6		Metallic	Paint		Metallic Paints
7		Pressed Steel	Wheel Type		Simple Pressed Steel
8		Simple	Paint		Solid Paint
9		Tinted	Glass		Green or Grey Tinted Glass
10		Untinted	Glass		Plain Glass

We can also inspect the imported Business Categories by right-clicking on **Glossary** and selecting the **Business Categories** option from the menu:

Business Categories

		Name	Short Description	Description	Parent	Owner
1		Glass		Glass Type	Specification	Administrator
2		Paint		Paint Attributes	Specification	Administrator
3		Specification		Orderable Automobile Specification		Administrator
4		Trim		Interior Trim Type	Specification	Administrator
5		Wheel Type		Wheel Type	Specification	Administrator

Note the top-level category **Specification** is shown as a **Master** category as it did not have any parent.

Libraries

Domains

Domains are a way of making reference data available to Experian Pandora dynamically. Domains are held outside of the Experian Pandora database structure as external CSV files and are dynamically loaded on demand and held in memory and indexed by the Experian Pandora server.

There are four types of domain files:

Value Lists

This is a simple list of values. They are referenced, case-insensitive, by all functions.

Format Lists

This is a simple list of Experian Pandora format patterns.

Regular Expression Lists

This is a list of regular expressions.

Name/Alias Translation Lists

This is a table of names and corresponding aliases. Each line of the domain file can have one or more aliases, or there can be multiple lines with the same name and different aliases, or any combination of both. Experian Pandora can even handle single fields containing multiple aliases.

Example Domains

Experian Pandora ships with several example domain files which are not intended to be exhaustive implementations:

English Dictionary

This is a 313,627 word dictionary of English words provided as a simple value domain.

First Names

A list of commonly used first names as a simple value domain.

Surnames

A list of over 9000 western surnames as a simple value domain.

Offensive Words

A list of rude and offensive words as a simple value domain.

UK Counties

Counties in the United Kingdom as a simple value domain.

UK Islands

A list of Islands in the UK as a simple value domain.

UK Towns and Cities

A list Towns and Cities in the United Kingdom as a simple value domain.

US States

A list of the 50 States in the United States as a simple value domain.

US Towns and Cities

A list of the Towns and Cities in the United States as a simple value domain.

Using Domains via Expressions

Experian Pandora provides a series of functions that can utilise domains, as detailed in the section on functions.

These provide functionality to:

- Check whether a domain has a matching value/format/regular expression
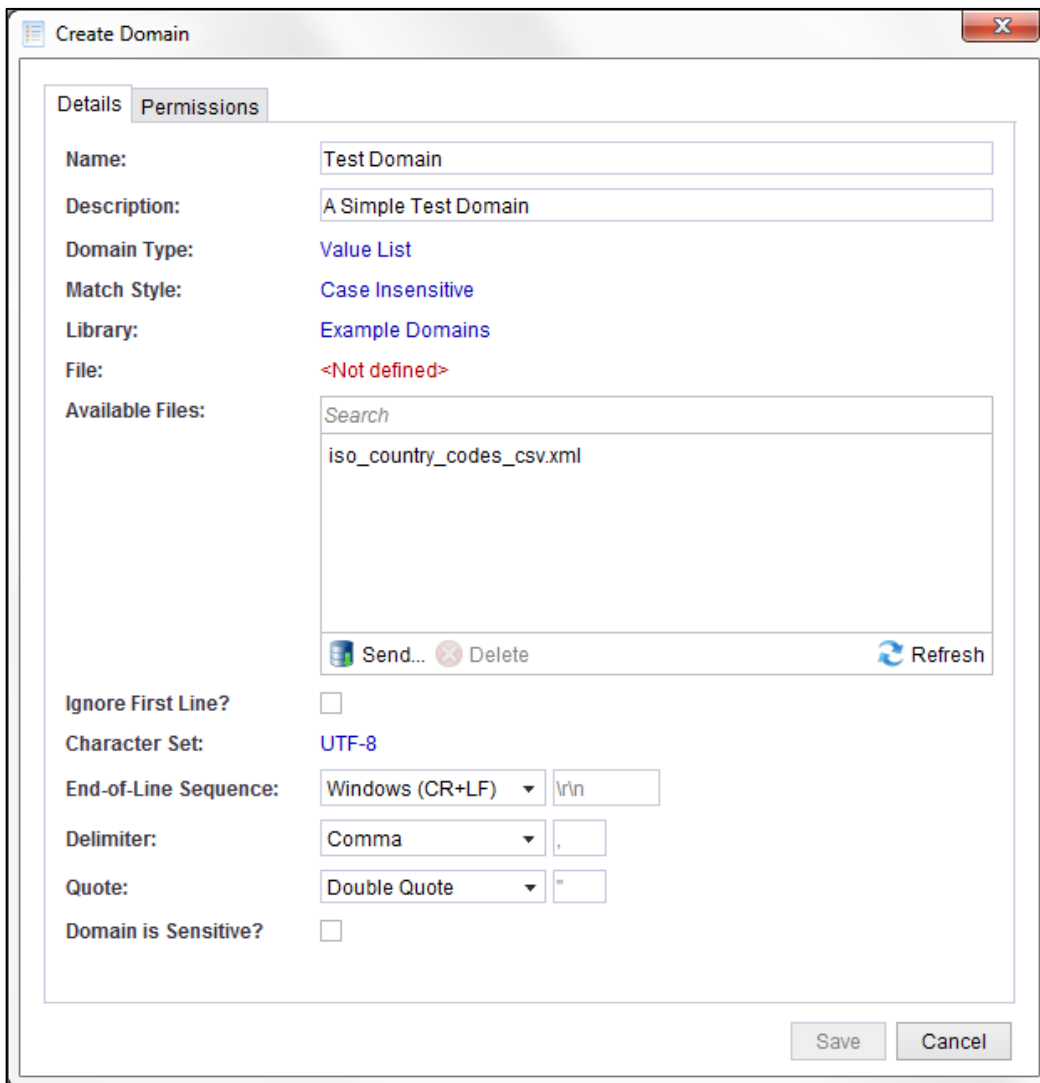
- Check whether a value exists in a domain
- Obscure values from free text that match a given domain (e.g. for replacing rude words in sentences)
- Extract values from free-text that match a given domain (e.g. for extracting key words)
- Remove values from free-text that match a given domain (e.g. for removing noise words)
- Translate an alias to its corresponding parent value using a value/alias domain
- Translate a value to all its potential aliases using a value/alias domain

Domain Libraries

A Domain Library is simply a container for domains to help in sub-setting and categorising them.

Domain Creation

To create a Domain, you must first create a Domain Library to contain it. From the previous section, we have a default library called **Example Domains**. Right-click on that library from the explorer and select **New Domain**: This displays the domain creation dialog:

Field	Description
Name	The name of the domain
Domain Type	The type of domain which can be one of Value List, Value List with Aliases, Format List or Regular Expression List
Match Style	The level of looseness of the match: • Case Sensitive – Matching performed in strict case. • Case Insensitive – Matching performed ignoring case. • Normalised – Matching performed by normalising (ignoring case and foreign characters). • Standardised – Matching performed by standardised (ignoring case, foreign characters and resolving punctuation). • Standardised: No Digits – Same as Standardised but with digits removed. Web Friendly - Matching performed by standardised (ignoring case, foreign characters and resolving punctuation), but retaining – for domain names and allowing '.' for a break character.
Library	The library to place the domain in, here pre-selected to **Example** as we right-clicked on that library to create this domain
File	The server data file to use.
Available Files	The list of available files on the server, held in the **CONTENT** directory of the database.
Ignore First Line?	Denotes the first line should be skipped as it contains headings
Character Set	The character set encoding (code-page) of the domain content file
End-of-line-Sequence	The end-of-line (EOL) sequence of the domain content file
Delimiter	The field delimiter of the domain content file
Quote	The field quote character of the domain content file
Domain is sensitive?	The domain is sensitive in that the values may need to be supressed on display. A good example of this is the offensive words example domain. If the user has the preference option **Hide Sensitive Domains** set to **on** then the domain values will be suppressed when drilling down to the content of the domain. The values will still be visible on hover-over.

Sending domain files to the server

Note that the Available Files are those that are currently in the **CONTENT** directory of the Experian Pandora server. To place additional files in this directory from your local client, simply press the **Send** button and select the required files.
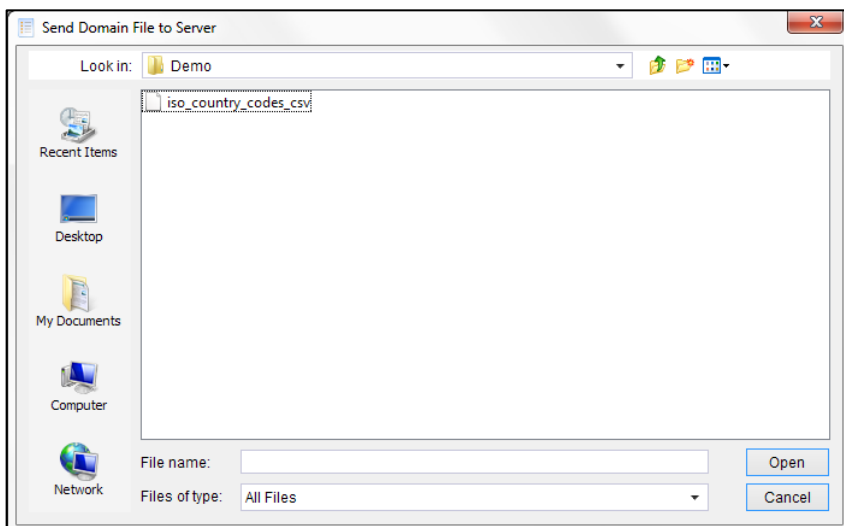
Removing domain files from the server

Clicking on the **Delete** button having first selected a domain file or edit dialog will remove the selected domain file from the **CONTENT** directory of the Experian Pandora server.

Refreshing the list of server domain files

Clicking on the **Refresh** button will refresh the list of files to the currently available list on the Experian Pandora server.

Updating an existing Domain

To change details of a domain you can simply edit it by right-clicking on it in the explorer and selecting **Edit**. The dialog is the same as the creation one except that you cannot change the domain library or the type of domain. You can however change which file is used. To update the content of a domain simply change the file. Additionally you can replace the file in the **CONTENT** directory on the Experian Pandora server as this is constantly monitored by the server, and the domain content will change dynamically. Care should be taken if the domain is in use as this will change any rules, drilldowns, views etc. that may be using the domain.

Drilling down to domain content

If we right-click on the domain, we can view the content of the domain by selecting **Content** from the menu.

Some of the values in the drilldown are highlighted in blue – this is because they are actually in existence in columns somewhere in the tables loaded in Experian Pandora. This is a real-time highlighting and part of the benefits of the Experian Pandora database system.

Viewing domain value usage

From the **Drilldowns** section of the Edit drilldown menu on the Content drilldown for a Domain, there are two additional drilldowns available to filter this display to show domain values that are in use (or not). Using standard Experian Pandora functionality, any Domain values can now be examined further to see exactly **where** they exist by selecting the menu option **Find Columns > Using this Value**.

Dealing with Domain sensitivity

Some domains may contain values that are sensitive or may offend, and it may be preferable that, when viewing the content, the values are not shown by default. An example of this is the **Offensive Words** domain shipped as one of the example domains. Hovering over any of the invisible values will show what they are in a popup. This is the case only if you have in selected in your preferences **Hide Sensitive Domains**.

Value List with Aliases Domains

Value/Alias domains are designed to provide one-to-many values to potential aliases. This can be used for a myriad of tasks, but a common one is standardising the names of people.

The format of value/alias files is a comma-separated values file, with optional header. The format is very flexible in that it allows all combinations of name/aliases as follows:-

- One value/alias per line
- One value field with multiple aliases fields
- Repeated values with aliases, one pair per line (duplicates ignored)

The content of a Value List with Alias Domain will have multiple columns for the aliases, which may appear as variable length given that not all values will have the same number of aliases. An **Alias Count** column is provided to give you an instant visual clue as to how many aliases exist for a value.

References to Domains

Whenever a domain is referenced in a permanently created object in Experian Pandora, a reference is made to prevent that Domain being deleted and accidentally breaking the object in question. For example, a view that uses a function that refers to a domain would need to ensure that the domain is not deleted.

The types of references that can occur are as follows:

Type	Description
Column	When a Column has an attached Validation Rule that uses a Domain
Table	When a Table has a Measure expression or attached Validation Rule that uses a Domain, or a drilldown has been saved as a Table, and an expression in the drilldown referenced a Domain
Function	When a user-defined Transformation or Validation Function references a Domain
View	When a saved View references a Domain in an expression Column

Drilldown	When a saved Drilldown references a domain in an expression Column

These references can be viewed by either expanding the Domain on the explorer, or by right-clicking on the Domain in the explorer and invoking the drilldowns under the **References From** sub menu.

Dealing with change

Domains by their nature are designed to be dynamic, yet are a powerful ally in both transformation and validation of data. Domains can be rapidly changed simply by changing their content on disk, without any load processes. The negative aspect of this is that without sensible controls this flexibility can easily invalidate transformation or validation results previously obtained.

We recommend that if a domain is not designed to be dynamic that it is protected from change, and a naming convention is used to either effectively timestamp or version the domains, allowing multiple variants of the same domain to co-exist at the same time. Experian Pandora will **not** prevent you from overwriting a Domain that is currently in use. For more help and guidance on this subject please contact your support representative.

Creating Domains from Drilldowns

It is easy to create a Domain from any drilldown. Simply manipulate the data as required and then export the drilldown to a comma-separated values file. Then use that file in the Domain creation process in the normal way.

Deleting Domains

Providing a domain has no references to it, it can be deleted by right-clicking on it and selecting **Delete**. Note that deleting a Domain does **not** delete the underlying Domain file.

Business Constants

Business constants provide the ability for commonly used values to be defined in a central place and referenced in both Transformation and Validation rules, as well as in conditional expressions such as filters and condition statements in Decision functions.

Types of Business Constants

A constant has a value and description and can be, for example:

- Literal Text
- Format Pattern
- Phonetic Value
- Regular Expression Pattern

Business Constant Libraries

Business Constants are organised in containers called Business Constant Libraries. These simply help group up similar constants into an identifiable category.

Creating a Business Constant Library

To create a Business Constant Library, on the explorer go into **Libraries** right click on **Business Constant Libraries**, then select **New** from the menu. You must create a library before you create a constant.

Editing a Business Constant Library
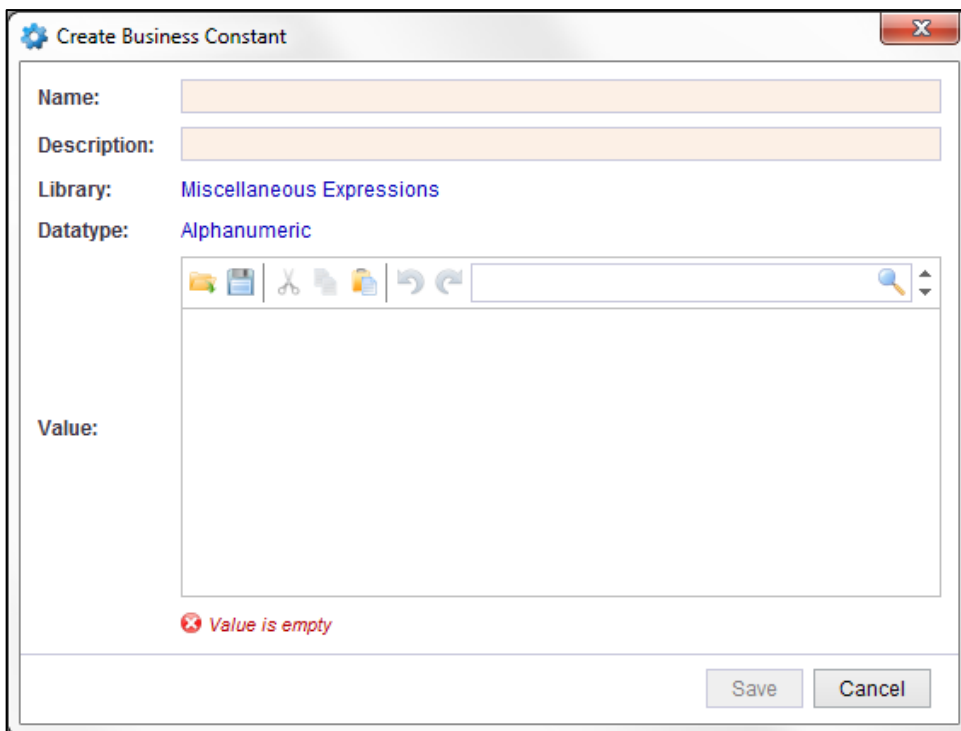
To edit a Business Constant Library, simple right-click on it from the explorer or a drilldown and select **Edit**.

Deleting a Business Constant Library

To delete a Business Constant Library, simple right-click on it from the explorer or a drilldown and select **Delete**. Note that the library must not contain any constants for the Delete option to be available.

Creating a Business Constant

Having first created a Business Constant Library to contain the constant, right-click on the library from either the explorer or a drilldown of Business Constant Libraries and select **New Constant**.

Field	Description
Name	The mandatory name of the Business Constant
Description	The mandatory description for the Business Constant
Library	The Business Constant Library this constant belongs to
Datatype	The type of the value in the constant. This is used as a hint to the server and used to validate entry in the Value field.
Value	The actual value of the Business Constant

Note the editor for the value is a multi-line editor as values could potentially be large, and supports copy/cut/paste/undo/redo, as well as the ability to load values in from external text files, and to save them to text files also.

Using Business Constants in Expressions

It is very easy to use a Business Constant in place of a value in an expression. Here is an example function element for multiplying price:
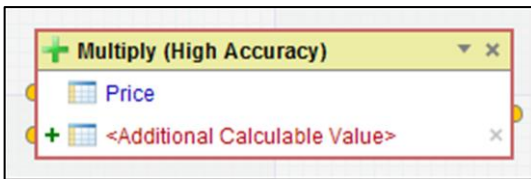
Normally, you would just type in the amount in the *Additional Calculable Value* field. Looking at the left, the symbol shows the type of input the field is expecting. In this case the symbol signifies a **column** – which you can choose from the drilldown. If you click on the icon on the left, it will present you with other input options:
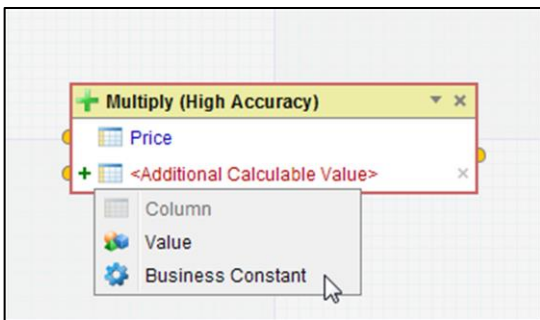
As you can see, one of the options is for a Business Constant. Select this and the behaviour of the *Additional Calculable Value* field changes. Clicking on the red highlighted text now presents a list of Business Constants, which can be subset by library (or chosen from an entire sorted list):
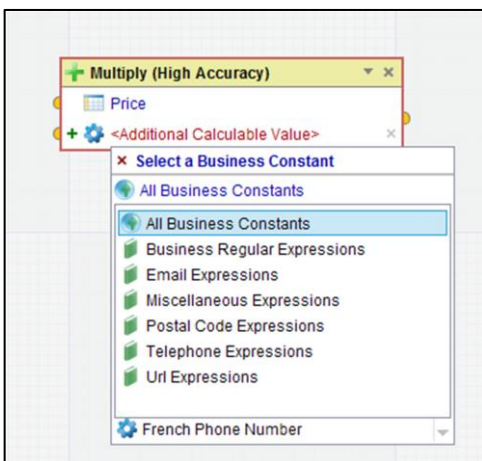
Select, for example, the **Low** discount Business Constant:
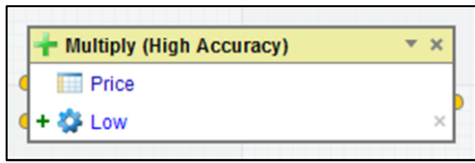
It is obvious from the blue icon that the value is actually a Business Constant "Low". These constants are referenced by id, not name, so it is safe to change the name at a later stage (which would be reflected in the expression if it was edited at a later stage).

Using Business Constants in Filters

It is very easy to use a Business Constant in place of a value in a filter. The filter is created in the normal way, but when it comes to selecting the value to match, we simply click on the **Constant** box on the value selection panel on the right, where we can now pick any available Business Constant, either from a complete list or subset them with the drop-down selection list of Business Constant Libraries.
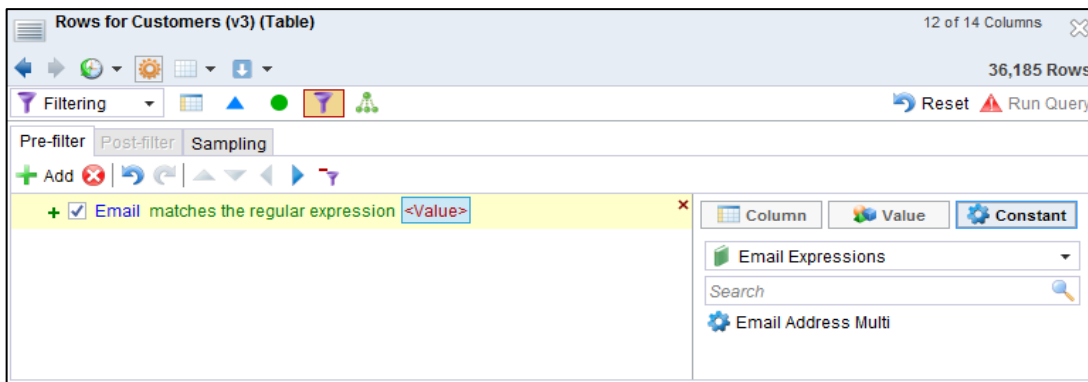
References to Business Constants

Whenever a Business Constant is referenced in a permanently created object in Experian Pandora, a reference is made to prevent that constant being deleted and accidentally breaking the object in question. For example, a view that uses a function that refers to a constant would need to ensure that the constant is not deleted.

The types of references that can occur are as follows:

Type	Description
Column	When a Column has an attached Validation Rule that uses a Business Constant
Table	When a Table has a Measure expression or attached Validation Rule that uses a Business Constant, or a drilldown has been saved as a Table, and an expression in the drilldown referenced a Business Constant
Function	When a user-defined Transformation or Validation Function references a Business Constant
View	When a saved View references a Business Constant in an expression Column
Drilldown	When a saved Drilldown references a Business Constant in an expression Column

These references can be viewed by either expanding the Business Constant on the explorer, or by right-clicking on the constant in the explorer and invoking the drilldowns under the **References From** sub menu.

Dealing with change

Business Constants by their nature are designed to be dynamic, yet are a powerful ally in both transformation and validation of data. Constants can be rapidly changed simply by changing their values. The negative aspect of this is that without sensible controls this flexibility can easily invalidate transformation or validation results previously obtained.

We recommend that if a constant is not designed to be dynamic that it is protected from change, and a naming convention to be used to either effectively timestamp or version the constants, allowing multiple variants of the same constant to co-exist at the same time. Experian Pandora will **not** prevent you from overwriting the value of a Business Constant that is currently in use.

For more help and guidance on this subject please contact your support representative.

Standard Function Libraries

A wide range of standard functions is provided by Experian Pandora for use by the user. Information on these functions can be found in the Standard Function Libraries section, located in **Libraries**. From here, you can view all the functions that have been provided as well as information about how they can be used and from what category they reside in.

User-Defined Function Libraries

Experian Pandora provides the ability to centralise common rules for transformation and validation. This essentially extends the palette of functions available in the expression editor or selectable as validation rules when applying validations directly to columns and tables. These saved functions are independent of any table or column and can be reapplied hundreds of times over by reference.

User-Defined Function Libraries are simple containers for grouping together user defined functions. These are identical to the standard libraries supplied with the product (which cannot be changed)

Creating a User-Defined Function Library

From the explorer, go into **Libraries** right click on **User-Defined Function Libraries** and select **New**.

A simple dialog appears allowing to enter the name and description of the library, as well as the standard security and ownership options.

Editing a User-Defined Function Library

From the explorer, right click on the function library you want to edit and select **Edit** from the menu.

Deleting a User-Defined Function Library

Simply right-click on the library on the explorer or from a drilldown of libraries and select **Delete**. You will only be able to delete libraries if they do not contain any functions, otherwise the Delete option will be disabled.

User-Defined Transformation Functions

Experian Pandora provides the ability to centralise common rules for transformation and validation. This essentially extends the palette of functions available in the expression editor or selectable as validation rules when applying validations directly to columns and tables. These saved functions are independent of any table or column and can be reapplied hundreds of times over by reference. They are implemented as new functions in the expression editor.

Creating Transformation Functions from an existing Drilldown

Referring to **A Simple Calculation Expression**, we demonstrated an expression column that converts Height in metres to Height in feet, and ensures the result is set to 2 decimal places.

Whilst it may be potentially a good conversion function, it could easily be made more universally useful by being further generalised. For example, it is not so much a height conversion – it is in-fact a conversion from metres to feet. It is also restrictive to set the scale of the result to a fixed value of 2, and instead this should be a parameter. In order to parameterise this, the expression must be first edited so that it does not have a constant, instead just put any column in its place so that is not already referenced in the expression. Experian Pandora will not parameterise constants, only variable arguments such as column names.

Here was the original expression:
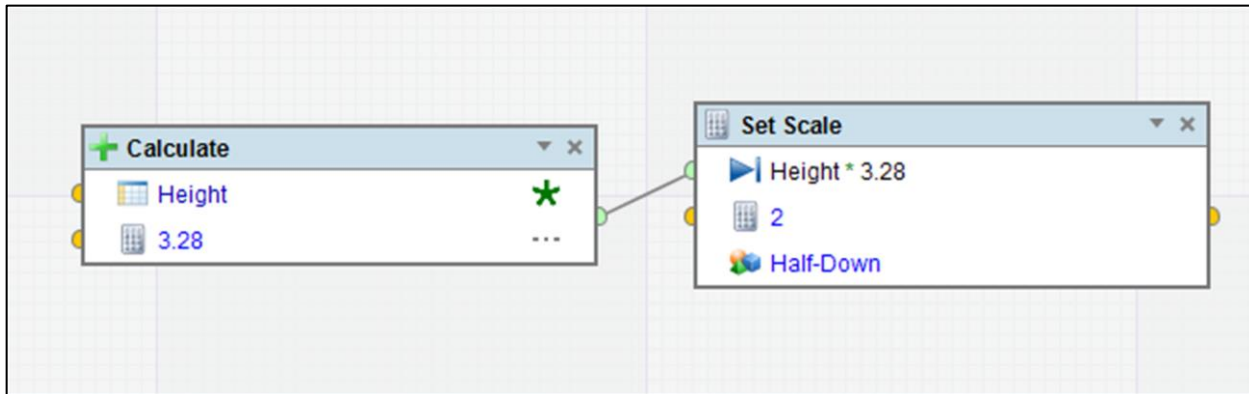
We can change the scale of **2** to just be any column in the drilldown. I will use **Row Id** just because it is convenient:
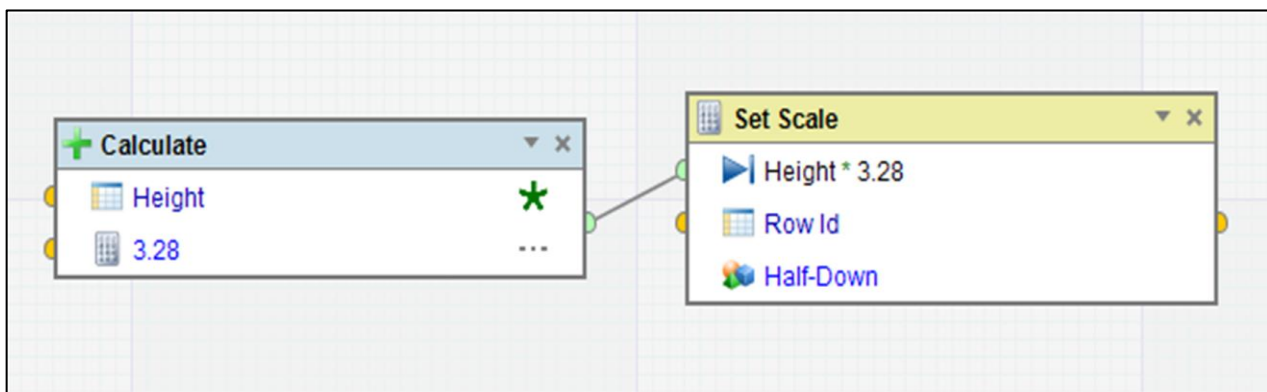
From the drilldown, we can simply right-click on the heading for the expression Column **Height in Feet** and select **Transformation > Save As > Transformation Function**

		Height	Height In Feet *f*		Range
1		1344	440	Edit...	Fort William to Loch Treig and Loch Leven
2		1309	429	Delete	The Cairngorms
3		1296	425	Hide	The Cairngorms
4		1295	424	Fit ▶	The Cairngorms
5		1291	423	Sort ▶	The Cairngorms
6		1265	414	Group ▶	The Cairngorms
7		1258	412	Insert Column ▶	The Cairngorms
8	earg	1249	409	Transformation ▶	ngorms
9		1245	408	Profile ▶	
10		1235	405	Validate and Enrich ▶	
11		1234	404		
12		1221	4004.88		
13		1221	4004.8799999999997		iam to Loch Treig and Loch Leven
14		1220	4001.6		iam to Loch Treig and Loch Leven
15		1215	3985.2		ngorms
16		1214	3981.9199999999996		on to Glen Dochart and Loch Tay
17	deir	1213	3978.64		ngorms
18		1197	3926.16		ngorms
19		1184	3883.52		ngorms

Transformation submenu:
- Edit...
- Save As ▶
  - *fx* Global Transformation Function
  - *fx* Global Validation Function
  - *fx* Table Expression...
- Remove
- Collections ▶
- Conversion ▶
- Data Cleanse ▶
- Date ▶
- Encoding ▶
- Number ▶
- Text ▶
- Validation ▶

We can now give the parameters proper names and descriptions:

Edit Arguments

Position	Parameter Name	Description
1	Height	
2	Row Id	

Edit Arguments

Position	Parameter Name	Description
1	Height in Meters	The height to convert in meters
2	Decimal Places	The number of decimal places to set the result to

You can now name the function and supply a description:

Create Function

Name: Meters to Feet

Description: Convert meters to feet with fixed decimal places

Library: Default

Versioned? ☐

[Save] [Cancel]

Click **Save**. This is now saved in Libraries.

Changing the Parameter Names

We can edit the custom function to change the names of the input parameters. Following on from the previous section, we right-click on the newly created function **Metres to Feet** and select **Edit**.
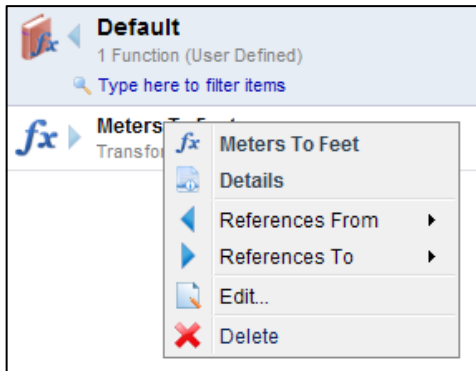
This will bring up the following screen to enable us to make changes. We have changed the location in the example below:
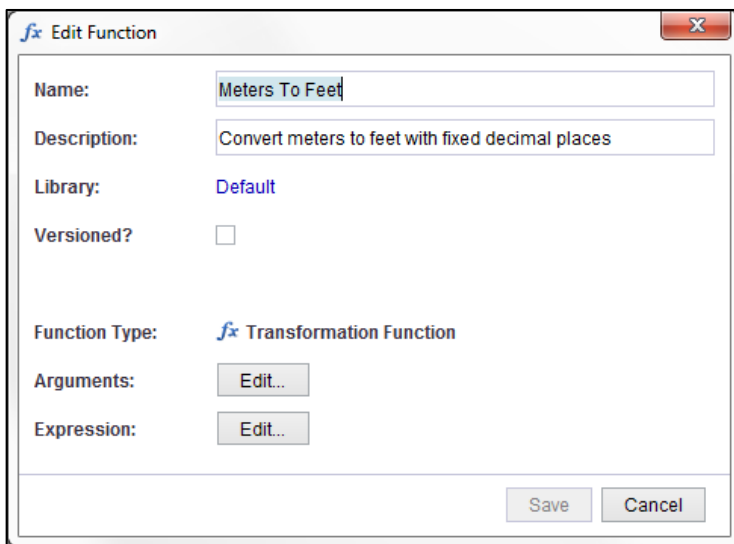
Expanding the Function Library displays the newly created function, and right-clicking on the library and selecting **Functions** will display the same information in a tabular drilldown:

We can now change the original drilldown to use our new standard transformation function. We once again edit the transformation. We simply delete the existing expression and now type in the search box on the left-hand side of the expression editor the name of the function we want, in this case our new function **Meters to Feet**.
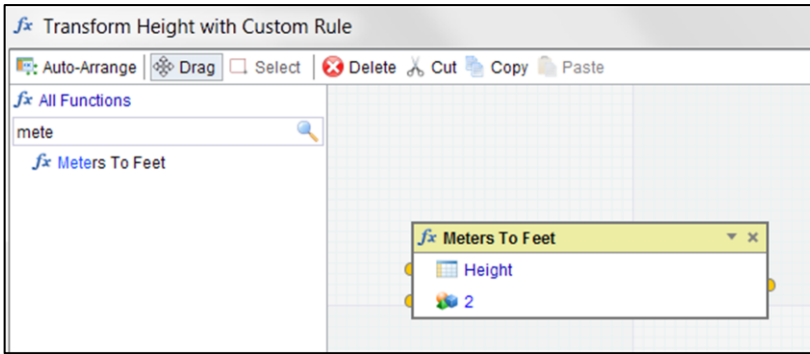
It only takes a few letters to find the new function (alternatively I could have subset the list of functions by picking my function library). I drag the new function on to the canvas and set the parameters simply to **Height** and **2**. I have exactly the same functionality, but the implementation is now hidden from this drilldown and can be easily changed in future.

Editing User Functions

You can also edit the actual expression for a Custom Function from its Edit window. This would, if saved, cause a **new version** of the function to be created. Function versions are individually addressable, relatable and can be deleted. You cannot change the operation of an existing, assigned version of a function. All links to functions are at a specific version.

To edit a function, right-click on it and select **Edit**.

Click on the **Versioned** checkbox to make this function versioned (disabled by default).
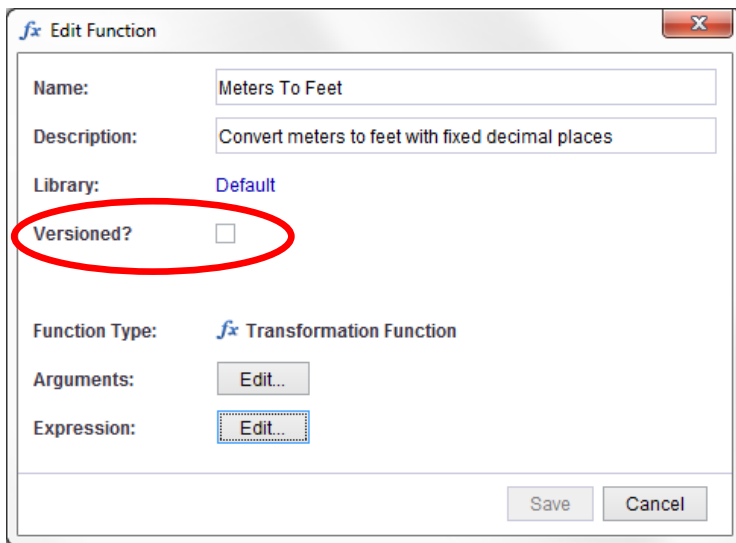
Next click on the **Edit** button for the **Expression**. The Expression editor appears with the expression in. You can now make any changes required. The only limitation is that you cannot change the number of parameters used by the expression. When you have finished your changes, select **Save**. This returns you to the function edit dialog above. Click **Save** on this dialog to create your new version. For this example, we change the rounding on the Set Scale part of the expression to **Half Even**.

The function has now changed on the explorer denoting that it has 2 versions:
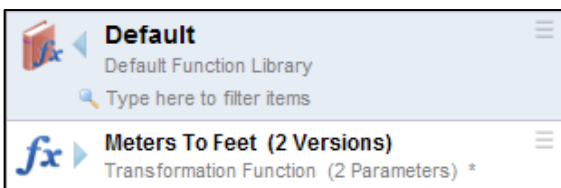
If we expand the function we can see more information:

Information is available including metadata about the function, and now a folder of versions. At the top level, the various references to and from other objects in Experian Pandora are an amalgamated set of references across all versions.

Expanding the **Versions** folder:

We can see the two versions of the function.

Expanding the latest version (2) we see information specific to that version, namely the specific references to and from other objects related to that version of the function.
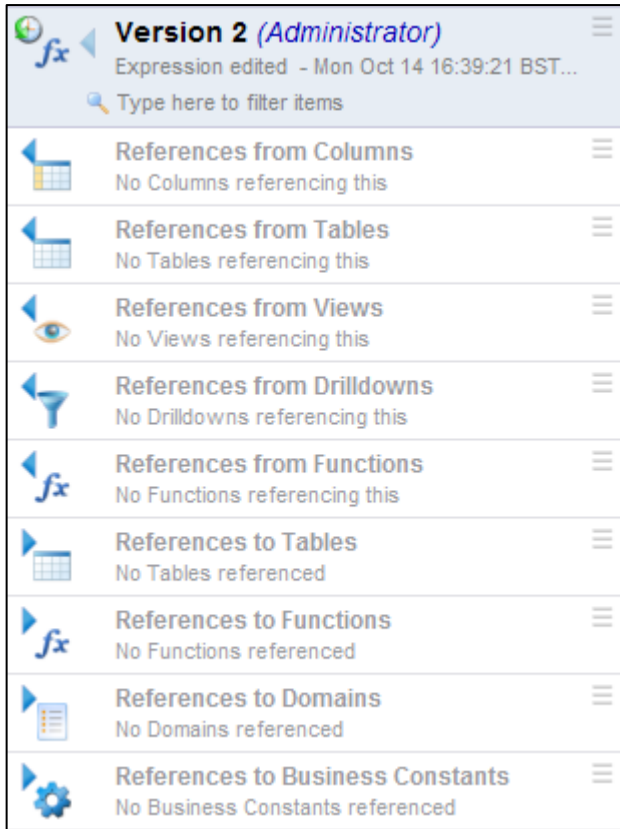
Deleting User Functions

Any version of a function can be deleted by right-clicking on it and selecting **Delete**.
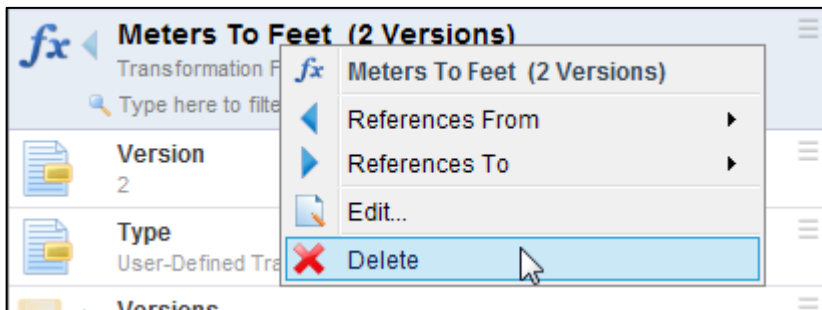
This will delete **all** versions of it, provided they are not referenced by other objects, i.e. not in use.

Deleting specific versions of User Functions

Individual versions can be deleted by expanding the versions folder of the function and right-clicking on an individual version and selecting **Delete**. Once again this is only allowed if that particular version is not in use by any other objects in Experian Pandora.

References to User Functions

Experian Pandora automatically manages usage links between a function and all of the objects that it depends on or that depend on it. This is to ensure that you can neither delete a function (or a version of it) that another object is using or delete something that the function is using, for example a look-up table.

The following links are tracked for things used by a User Function. Experian Pandora will prevent the deletion of any of these if they are used in a User Function:

Type	Description
Column	When a column is referenced in a Table lookup or other column related function within the expression for the User Function
Table	When a Table is referenced in a Table lookup or other Table related function within the expression for the User Function
Function	When the expression in the User Function makes use of another User Transformation or Validation Function
Domain	When a the expression in the User Function makes use of a Domain related function and hence uses a specific Domain
Business Constant	When a the expression in the User Function makes use of a Business Constant as a function parameter and hence uses a specific Business Constant

- Lookup Tables/Columns used
- Domains used
- Business Constants used
- Other Custom Functions used

These appear as **References To** folder items underneath the function (or a specific version of it) and also on the **References To** right-click sub menu of either the function or the function version.

The following links are tracked for things that **use** the Custom Function. Experian Pandora will prevent the deletion of the function if any of these links are in place. These are the usages of the function:

Type	Description
Column	When a Column has an attached Validation Rule that is a User Validation Function
Table	When a Table has a Measure expression that uses a User Validation or Transformation function, or an attached Validation Rule that is a User Validation Function.
Function	When another User Function uses this User Function within its expression
View	When a saved View has an expression Column that uses this User Function
Drilldown	When a saved Drilldown has an expression Column that uses this User Function

These appear as **References From** folder items underneath the function (or a specific version of it) and also on the **References From** right-click sub menu of either the function or the function version.

Different versions of the function have their own links which can be different from version to version, as the expression might be entirely different.

- **You can only delete a version of the function if it is no longer referenced**

- **You can only delete a function if all versions are no longer referenced**

- **You cannot change the number of parameters in a function with new versions**

- **Parameter names are shared from version to version**

User Validation Functions

User defined validation functions can be created in exactly the same way as user defined transformation functions, except for the following aspects:

- The expression of a User Validation Function must terminate with an existing validation function. The built-in function **Is Valid** is available to convert the output of other functions (e.g. Decision functions) to convert text or numeric responses to true/false. The functions **True** and **False** allow explicit response to be output, but these are most likely to be fed into a **Decision** function, which in itself is not a Validation Function. This is an example of where you would place the **Is Valid** function at the end of the expression.

  Experian Pandora understands all of the following values in converting values to validation results:

Value	Validation Value	Case Sensitive
TRUE	True	No
FALSE	False	No
YES	True	No
NO	False	No
1	True	No
0	False	No
Y	True	No
N	True	No

- A Validation Function can be attached to Columns or Tables for validation purposes, a Transformation Function cannot.
- Both Validation and Transformation functions can be used in another User Validation or Transformation function.

- A User Defined Validation function can be used wherever a user defined transformation rule can be used.

Saving User Validation Functions from a Drilldown

If an expression column in a drilldown is a Validation Function, then Experian Pandora will offer you an additional option on the **Save As** menu under the **Transformation** heading, which is **Transformation > Save As > Validation Function.** The rest of the process is identical to **Creating Transformation Functions from an existing** Drilldown.

Function Versioning

The examples in the manual have all shown the functions as being versioned. The default state of new user-defined functions is that they are **not** versioned.

This means that any changes to the expression will always be applied to (and replace) the saved expression and any references to this function will be immediately affected by the changed expression.

Enabling Versioning

The function can, at any time after creation, have versioning enabled. Once a function has had versioning enabled, changes will result in a new version of the function and references to the function can be made to any of the available versions of the function.

Disabling Versioning

If a function has previously been versioned and has been modified so that it has more than one version, then existing references to any version of the function will pick up the expression saved in the latest version if versions are disabled at a later date. The older versions are not actually deleted and the repository links are maintained in exactly the same way, just dynamically pointed to the latest version. This means you can actually re-enable versioning and get back the multiple versions and links as they were before.

Example

1. Create a function **EXAMPLE**.

2. Reference this function in two views, **VIEW1** and **VIEW2**.

3. Edit **EXAMPLE** a couple of times. The expression stored in version 1 will change and both Views, when executed, will pick up the latest change to the expression.

4. Now make **EXAMPLE** versioned and change the expression so that it now has 2 versions.

5. **VIEW1** and **VIEW2** will continue to pick up the expression stored in version 1.

6. Change **VIEW2** so that it now references **EXAMPLE** version 2. As expected, **VIEW2** will pick up the latest version of the **EXAMPLE**. **VIEW1** will continue to pick up version 1 of **EXAMPLE**.

7. Now make **EXAMPLE** un-versioned again. **VIEW1** and **VIEW2** will now both pick up version 2 of the expression. Change the expression on **EXAMPLE** again. This will result in changing version 2 and **VIEW1** and **VIEW2** will change accordingly when re-executed or refreshed.

8. Now re-enable versioning on **EXAMPLE**. **VIEW1** will now pick up version 1 of the expression and **VIEW2** will pick up the recently-edited version 2.

ISO Countries
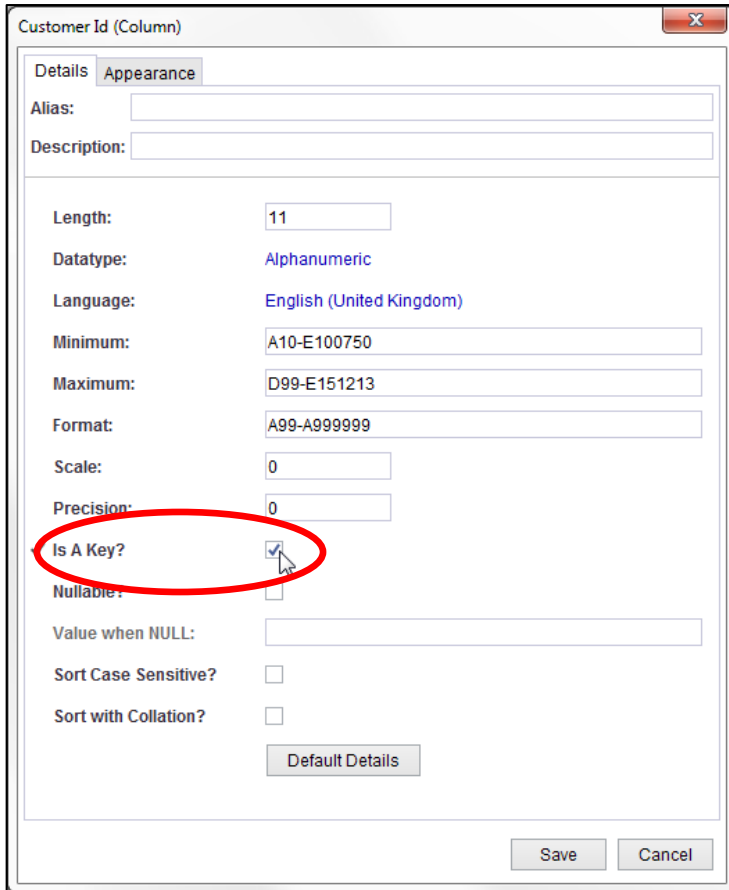
Experian Pandora provides a drilldown of all the ISO countries along with associated information and lots of pretty flags. ☺

Reference Tables Vendors

A **Reference Table** is a compact and portable table that can be used for the purposes of storing reference data which can be used by the **Reference Table Lookup** function. Reference Tables are grouped by **Vendor** (which is defined at the point of either creating the Reference Table, or importing it from a package provided by a specific reference data vendor).

Creating a Reference Table

Selecting the menu option **Save As > Reference Table** from the Drilldown Action Menu button on any drilldown, will save the data in the drilldown as a **Reference Table**. It is important that one or more columns in the drilldown is defined as a **Key**. A drilldown column can be defined as such by right-clicking on the column header, choosing **Edit…** and selecting the **Is a Key?** checkbox.

In the **Create Reference Table** dialog, as well as specifying the name of a Reference Table, you will also need to specify a Vendor Name, which is used to group similarly-named Reference Tables, as well as providing the future means of packaging for use in other instances of Experian Pandora.

The reference table is saved in **Libraries > Reference Table Vendors** under the vendor you have chosen or created.

Packaging the Data

To package the data to send to a recipient in **Libraries > Reference Table Vendors** select the required vendor then right click on the relevant table name either in the Explorer or the Drilldown and select **Package.**
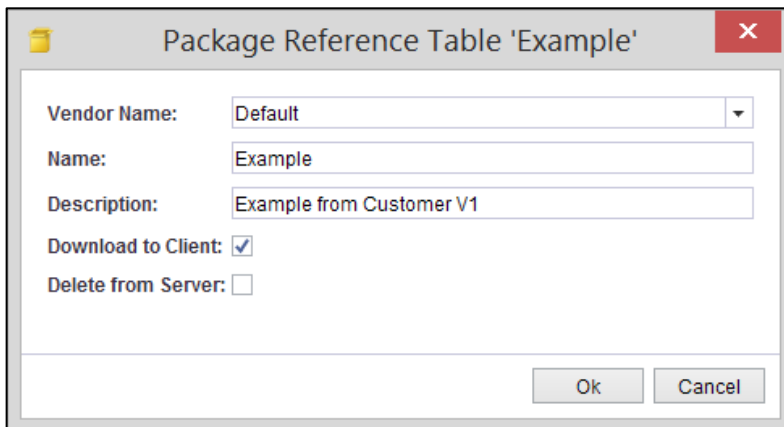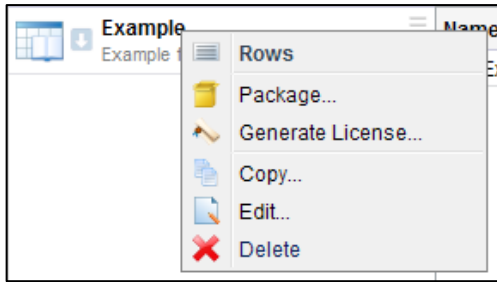
If the information in the dialog box is correct click ok or alter as required. Clicking OK will open up the Export Reference Table dialog box for you to select where you wish to save the data to. If the file already exists a warning will pop up asking if you wish to overwrite the original file. If you select No you can rename the file and continue to package it.

Creating a License

To create a license right click on the relevant table name either in the Explorer or the Drilldown for the required vendor and select **Generate License**.
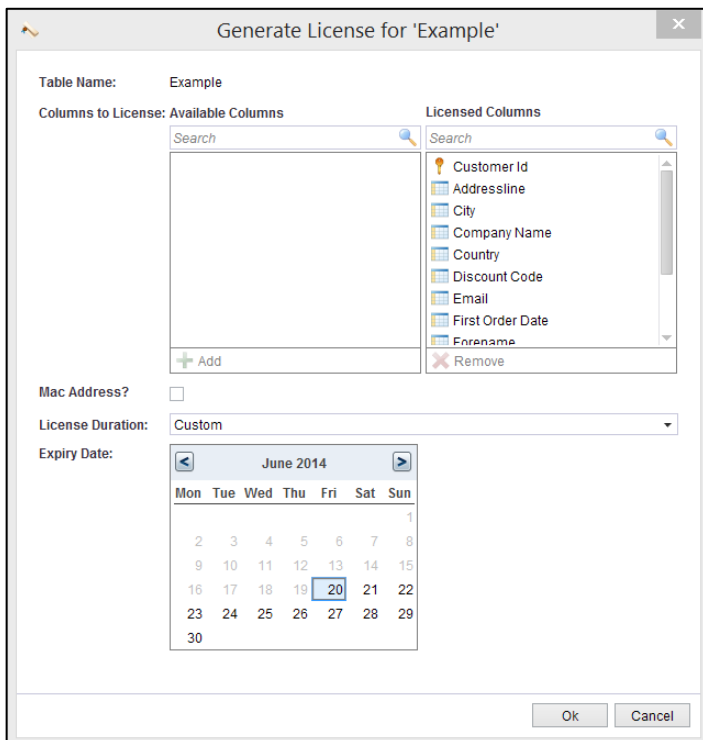
All the columns are currently in the licensed columns side on the right. Move the columns you don't want to share to the left hand side either by double clicking on a column name or hold down the control key, make your selection and hit Remove. These can be added back in using the Add button under Available Columns.

To tie the Reference Table to a server tick the Mac Address box and insert the MAC address of the server or, if you want more people to be able to view the information, leave this blank.

The license can be created for a set period of time by selecting one of the options from the License Duration drop down menu or a custom period of time can be set by selecting an end date.

If you have chosen to create a Perpetual license a warning dialog will open when you click OK to confirm that you do want a Perpetual license. NB: you will need to re-enter all the details from scratch if you select no.

The license is now created and will need to be named and saved on your Computer.

Once the Table has been packaged and the license created these can be sent to the recipient to enable them to view the data through Experian Pandora

Importing and Licensing Reference Tables

To import reference tables right click on **Reference Table Vendors** and select **Import Reference Tables**. This will open a dialog box for you to locate the required Reference Table file(s) and Import.

Once the Reference Table has been imported right click **Reference Table Vendors** and select **Apply Licenses**. Choose the relevant license and hit Apply.   A popup will confirm that the license has been applied.   You can now open and view the data. Multiple Reference Tables and multiple Licenses can be imported together.

If you have already loaded a table with the same name a warning will popup asking you to delete the old one before installing the new one. This is also a good way of checking whether you have already installed the table.   If any licenses do not match any of the installed Reference Tables a warning will be shown in the popup.
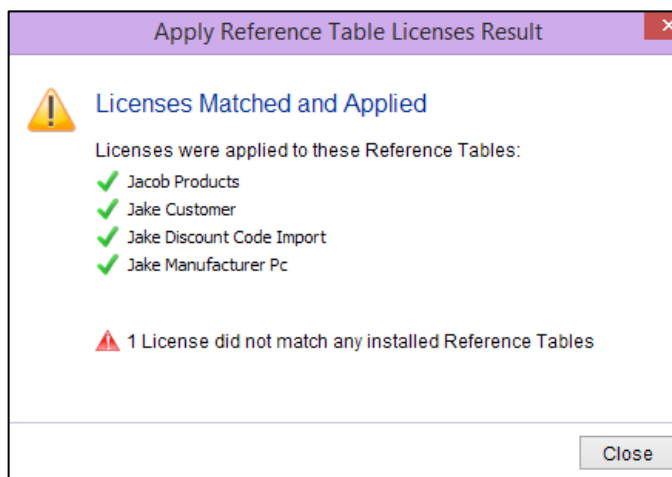
Reference Table Constraints

There are some constraints on data that may be stored in a Reference Table:

- The data must adhere strictly to the defined datatype of the containing Column (e.g. if a Column is defined as Integer, it must not contain any Alphanumeric values)
- The precision of decimal values must not exceed 19

If these constraints are not respected, the creation of the Reference Table will fail.

Extending the Repository with Properties

Experian Pandora allows most repository objects to be arbitrarily extended with additional information, known as **Properties**.

A Property is a key/value pair where the keys and values can be used for any purpose you desire, for example to provide additional documentation or metadata about the objects.

Properties can be added to most repository objects, and are inherited by child objects. For example, a property placed on a Table is automatically inherited by all of its Columns. A property may be overridden at a lower level – e.g. a Column can override a property set at the Table level.

Properties are categorised, and some have special meaning. For example, the **Mapping Annotation** properties automatically appear in the Mapping Documents that can be generated from the application.

Editing Properties

To edit the properties, select an object and edit it. For example, to edit a Table, select it on the explorer, right-click on it and select **Edit > Settings.** Click on the **Properties** button to reveal the **Property Editor.**

Property List

Select **+ Add** to show the property list at the top of the editor. This is the list of properties attached to this Table.

Scope

This is the scope of the property, which is the object type it was applied to.

Type

This is the type of property which can be one of:-

- General
- External Reference
- Mapping Annotation
- Mapping Parameter
- User Annotation
- User Parameter
- JDBC Property
- Script Discriminator

Name

The name of the property

Value

The value of the property

Edit

Select a property and click **Edit**. The **Edit Property** panel appears:
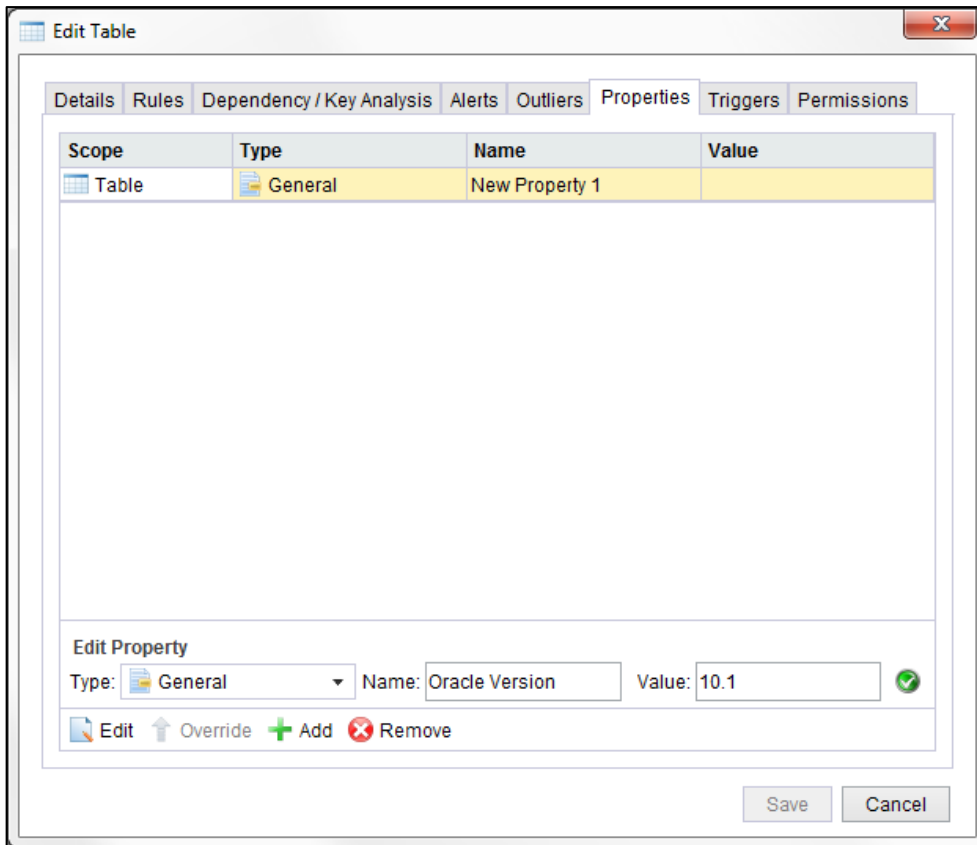
Here you can change the type of property, name and value as displayed in the property list. Click the green tick to accept the changes.

Override

Here we can see the properties of the Name column which is one of the columns in the Mountains table, and we can see it has inherited the **Oracle Version** property. If we select it, and click **Override**, we will override the inherited value with whatever value we choose.

New

Create a new property at this level.

Remove

Removes a property at this level, or removes an override of an inherited property.

Viewing Properties

These appear as static metadata for objects. Here we have added a property **Oracle Version** to the Mountains table. By looking in the **Info** folder for Mountains, we can see the property displayed:
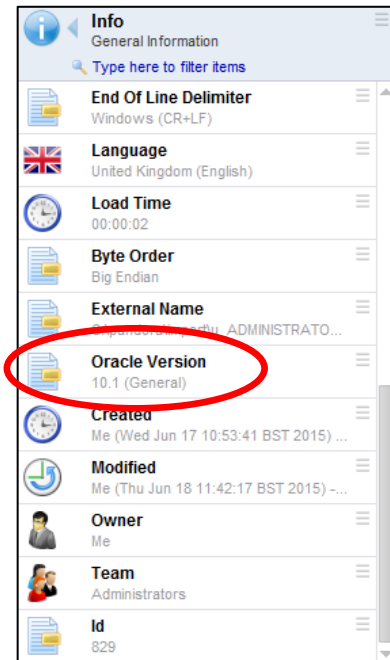
If we look at the **Name** column for Mountains, we can also see that it inherits the property from Mountains:
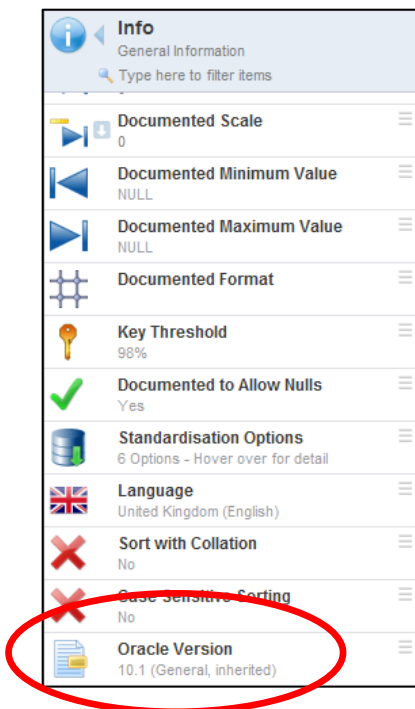
We can also view a drilldown of all properties, across all objects in the repository by right-clicking on **Data > Properties**.

Right-clicking on a property gives you access to the right-click menu for its parent object and thus access to the editing panel for the object, to allow you to change or remove the property.

JDBC Properties

JDBC Properties are used to add JDBC connection properties to a Connection. JDBC properties are simple key\value pairs and are added to the Connection's JDBC connection string. Refer to the specific JDBC driver's documentation for available connection properties, what they do, and what their valid values are.
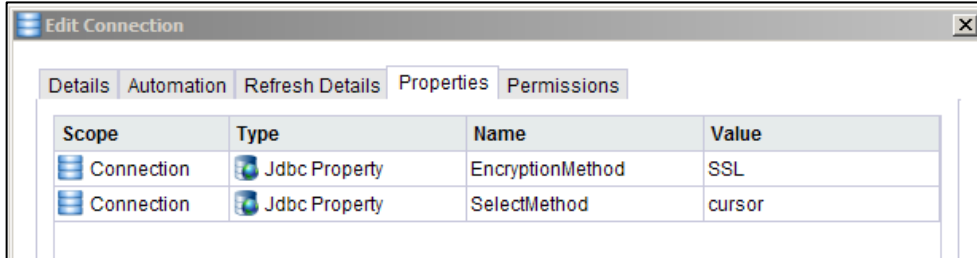
Table Expressions and Filters

Table Expressions

Column expressions, created within a drilldown of Table Rows, can be saved as Table Expressions. These Table Expressions persist as part of the table structure and are automatically recreated when loading a new version of the table.

Create the new column in the usual way using either the right-click on the column or Custom Expressions. For this example, we have transformed the Company Name Column in the Customers table, using a Text -> To Proper Case transformation:
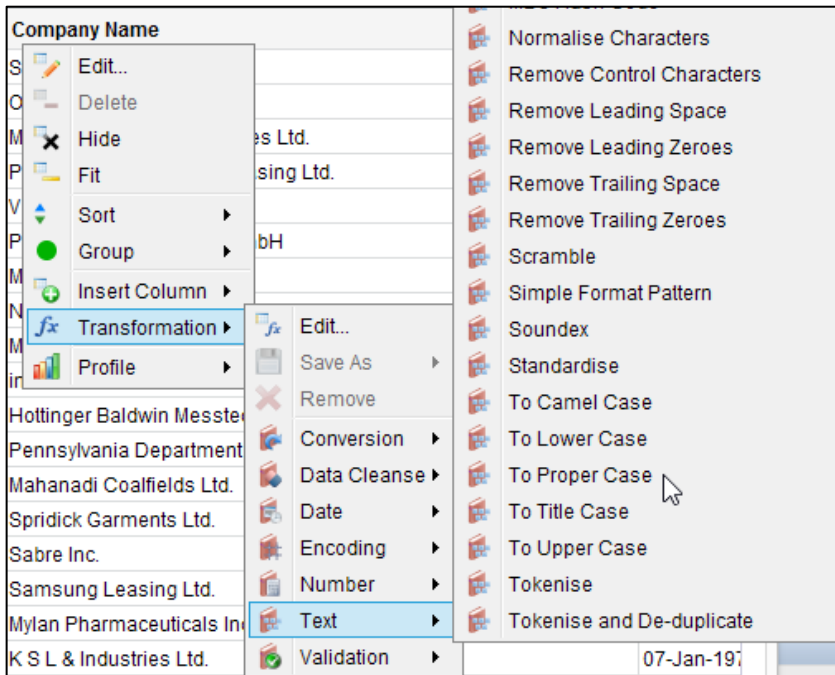
Once the transformation is completed, right-click on the new column and select **Transformation > Save As > Table Expression.**
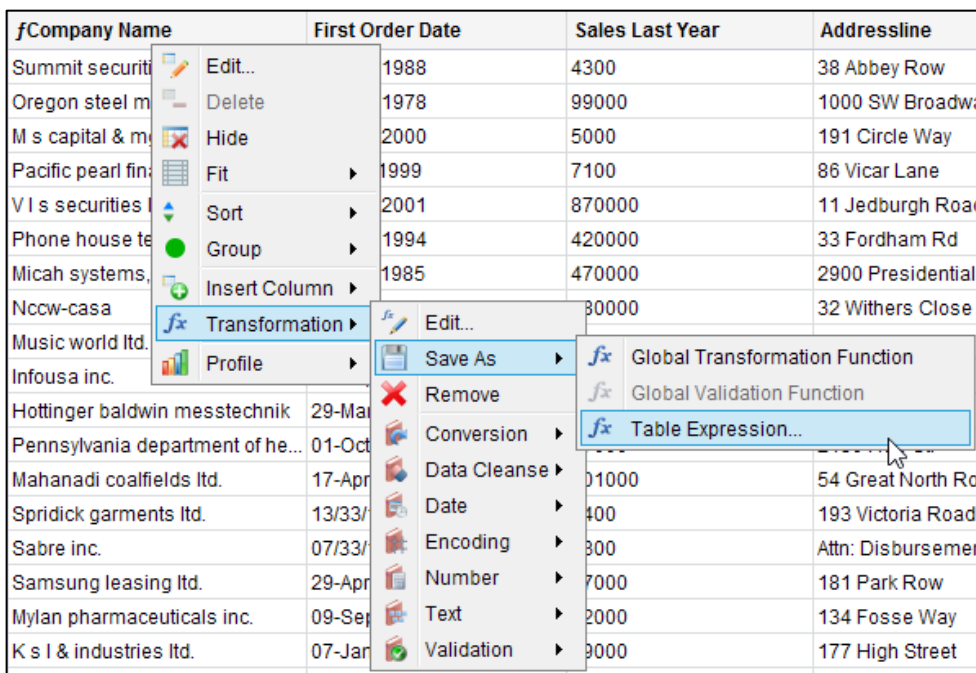
Note that Table Expressions can also be created from scratch by right-clicking on the **Expressions** folder in the explorer beneath a table, and selecting **New.**
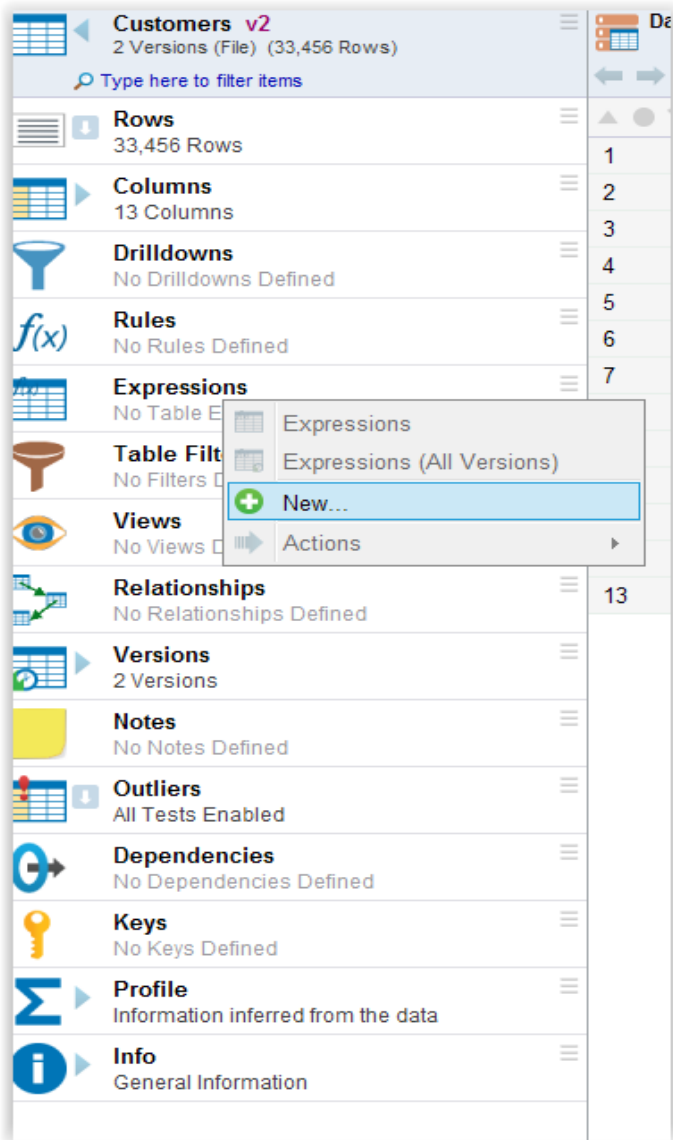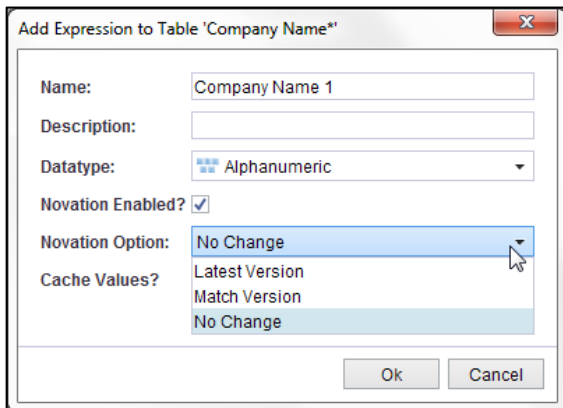
Datatype

Experian Pandora will select the datatype from the content of the column in the expression but you can change this if you so wish.
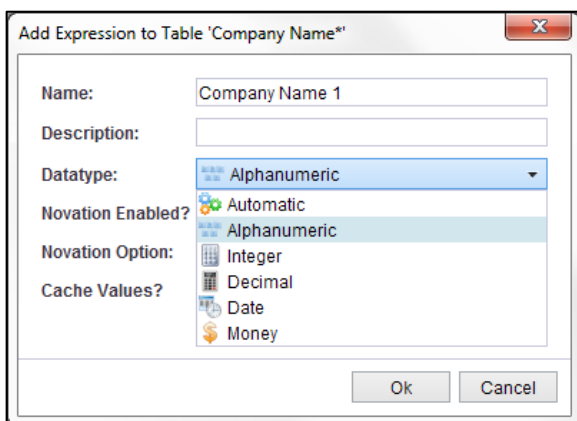
Novation Option

Novation options can be changed at any time by right-clicking the Table Expression in question and editing it.

Latest Version

This option will change the reference to the foreign table to be the highest current version of that table.

Match Version

This option will attempt to match the version for the foreign table to be that of the new version of the loaded table that the Table Expression is attached to. If no match is found it first searches up for the next highest version. Failing that it searches down for the next lowest.

No Change

This leaves the reference to the foreign table to be at the same version that it was created at.

Disabled

This option will prevent the selected drilldown from appearing in future versions of the parent table. This option is only available on saved Drilldowns.

Cache Values

To improve the performance of complex Table Expressions, you can choose to cache the values to improve performance when revisiting the Table Expression. It is recommended that this is only used where the Table Expression will not change.

Editing Table Expressions

We now realise that we have used **Proper Case** and should be using **Title Case** and need to edit the Transformation. We need to edit the transformation that we have created.

In the Explorer select **Expressions** then **r**ight-click on the Table Expression and select **Edit Expression.** This displays the Expression Editor with the original expression which can be changed. Simply change the **Proper Case** function to a **Title Case** one.

To see our changes, open a new drilldown to view rows with Table Expressions.
Note that all Table Rows drilldowns implicitly contain additional columns for any created Table Expressions, but these are hidden by default in the normal Table Rows drilldowns.
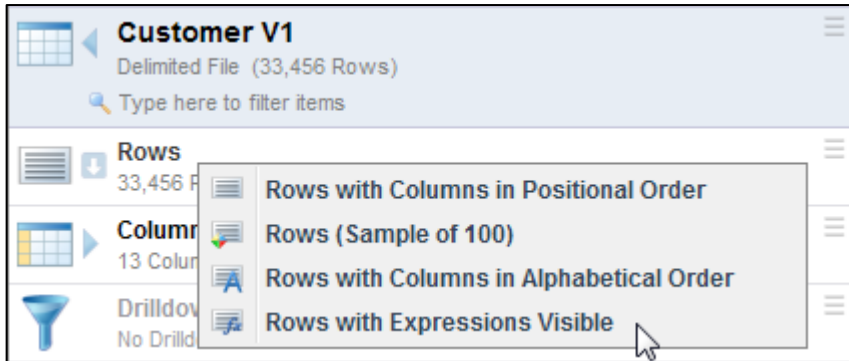
The Table Expressions will appear at the far right of the Table Rows drilldown.

When you edit a Table Expression on the rows drilldown, you actually edit the definition of the Table Expression which results on all uses of the Table Expression changing.
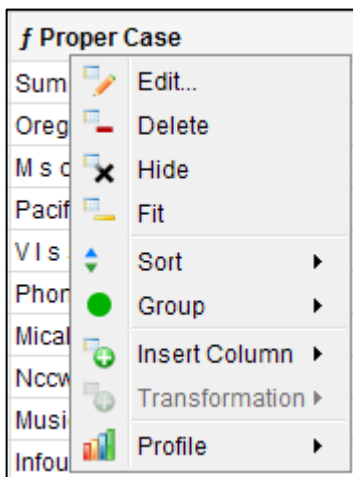
Note that if you add a new Table Expressions with a Table Rows drilldown for the table visible, it will NOT automatically add the new Table Expression to the drilldown. You must explicitly re-issue the rows drilldown.

Profiling a Table Expression

Simply right-click on a Table Expression and select the name of your table expression to view a profile of the values that are the result of the Table Expression being applied on every row of the table.
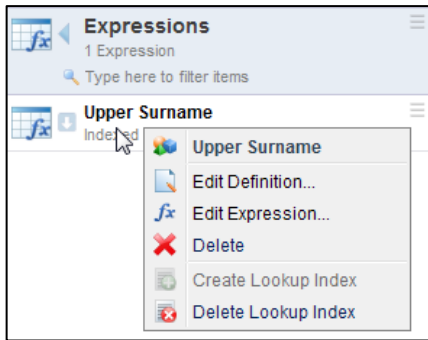
Indexing a Table Expression

Table expressions can be used for table lookups to compare values from one table to transformed values from another. This is done by creating a table expression, right clicking on its name, either in the explorer or on the **Expressions** drilldown, and selecting **Create Lookup Index**.
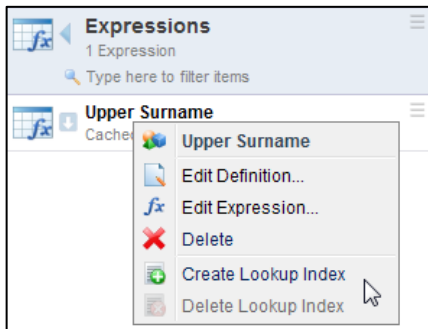
Having created the lookup index you can use a table lookup function to compare one table to the other containing the indexed expression. The option of looking up the column created by the table expression now appears in the Lookup Column section of the function.
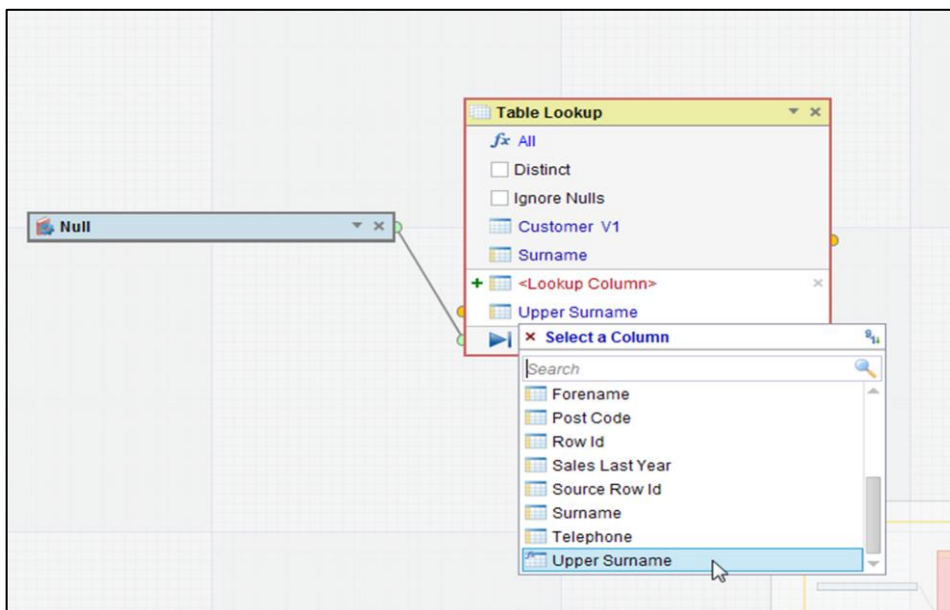
Table Filters

A Table Filter is exactly that – a saved filter of Table Rows of a table. The Table Filter may reference expression columns. The advantage of Table Filters over, say, Saved Drilldowns is that the Table Filter can be used in the evaluation of Validation and Quality rules.

Table Filters are automatically novated to new versions of a Table at load time.

To create a Table Filter, from a Table Rows drilldown, simply filter the drilldown as required, then click on the Edit dropdown menu button on the drilldown toolbar and select **Save As >Table Filter**.
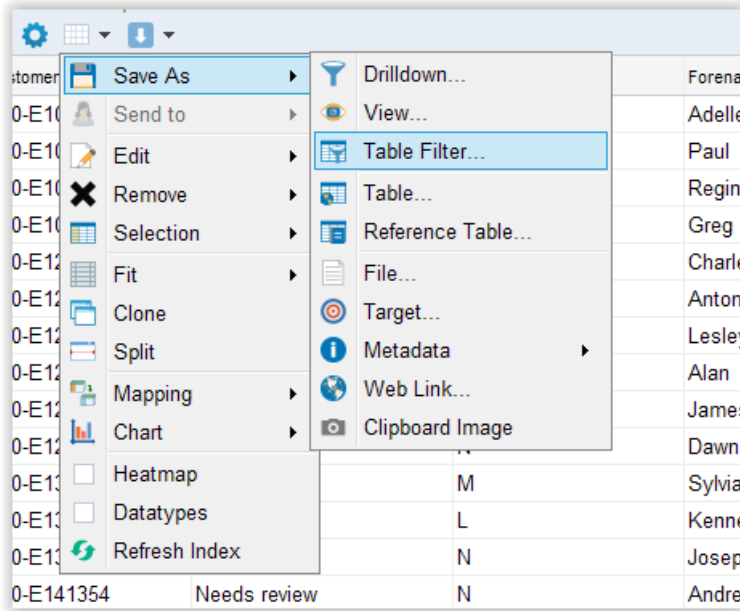
From the **Filters** folder of the Table, you can see the newly created Table Filter. This can be drilled-down on to reveal rows matching that filter. Simply drilldown to the filter in question to see the rows matching that filter.
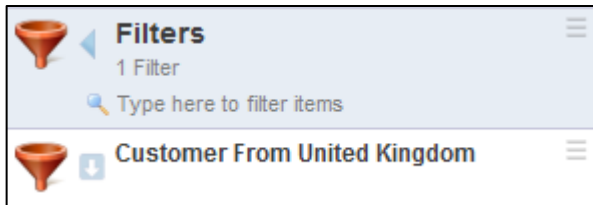
Note that you cannot save a Table Filter if the drilldown, including any of the table expressions it uses, references a dynamic function.

Handling Multiple Versions of Tables

Table Versions

Experian Pandora provides the ability to load multiple versions of the same table. When loading one or more tables whose name matches that of an existing table, you will be asked whether or not to load the table(s) as a new version of the existing one.

You can also explicitly load a new version of a table by right-clicking on it and selecting **Load New Version**.

This will open the Load File dialog. You can configure the table in the same way as when loading for the first time. New versions of tables can also be created from drilldowns by saving them as a new table with the name of an existing table.

New versions do not have to have the same schema names. Experian Pandora matches columns between versions using their name, but allows the introduction or removal of columns and provides information about the changes over time of columns added, removed and moved, as well as tracking changes in profiling statistics for the table and columns over time.

Projects, Schemas and Tables only ever show the latest version of a Table in all Explorer and Drilldown views.

Explorer differences for versioned tables

When a table is versioned, it still appears in the explorer under the relevant Table's folders as a single table, namely the highest version of the table.

Additional explorer folders

Versions

Previously loaded versions of the table are found under **Versions** in the table to which they relate.
This is a list of **all** versions of the table, including the highest one:
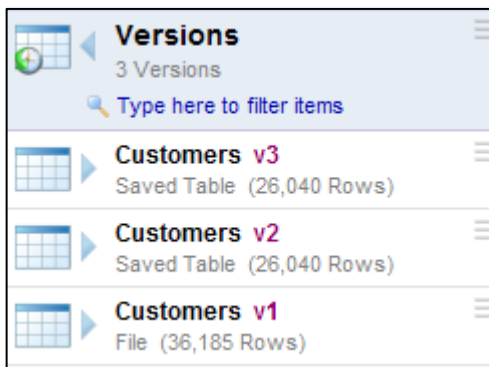
Column changes between versions

From the explorer menu above, you can go into a table's folder, right-click on **Columns** and select **Added, Moved** or **Removed Since Last Version** to bring up a drilldown showing the changes to the columns over each version.
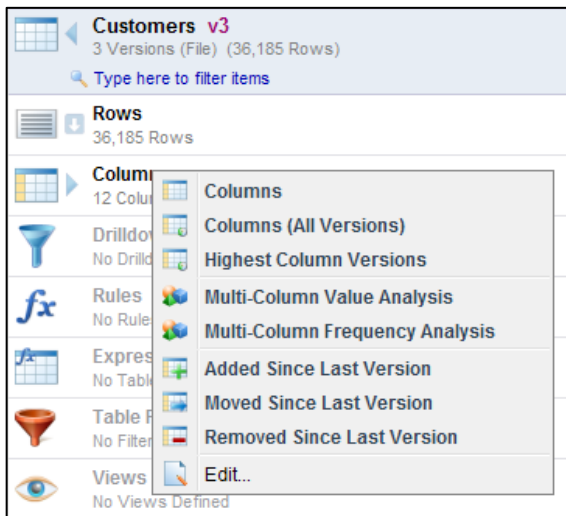
Table changes over time

Another feature used to track changes between versions is the **Change Over Time** selection from right clicking on **Versions** when in a table's folder. Select **Graph** to view a drilldown of the change in number of rows for each version of the table. More details later in this chapter.

Manually enabling versions

You can manually make a table multi-versioned by navigating to the table in the Explorer, right-clicking on the table name and selecting **Advanced > Enable Multiple Versions.**

Disabling multiple versions

Right-click on the table name and select **Advanced > Disable Multiple Versions**. All of the versions are reverted back to single tables. If they originally had different names, their original names will be restored. Any name clashes will be dealt with by Experian Pandora, generating unique variants of the table names to ensure that there are no name clashes.

Reassigning a table

To make a version of a table stand alone as a new table, click on the table in the explorer go into the **Versions** folder, then right-click on the version you want to make a table from. Select **Advanced > Reassign** and the Reassign Table dialog will appear.
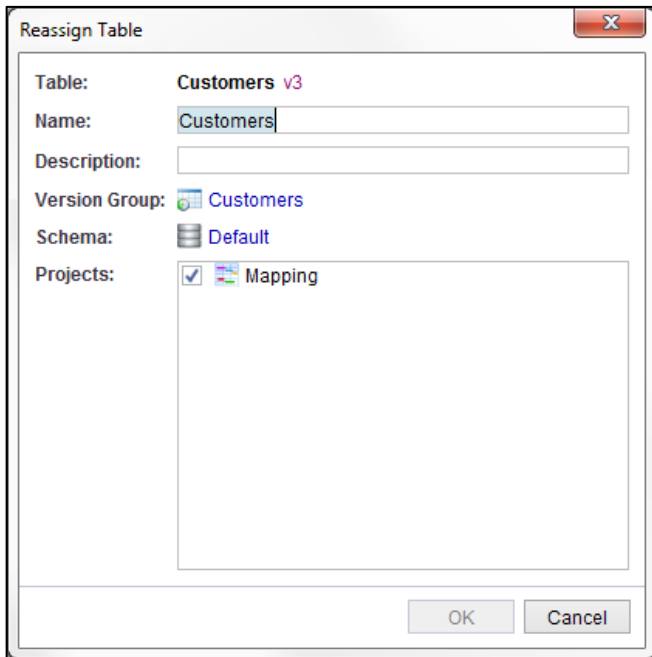
Name

Rename the table you want to create. This cannot be the same as the table versions you are extracting from

Description

Here you can write in a description of the table, this is optional.

Version Group

This option is used to move your selected table into another group of versioned tables. If you want the table to be on its own, or the first version of a new group of tables, select **<None>**.

Schema
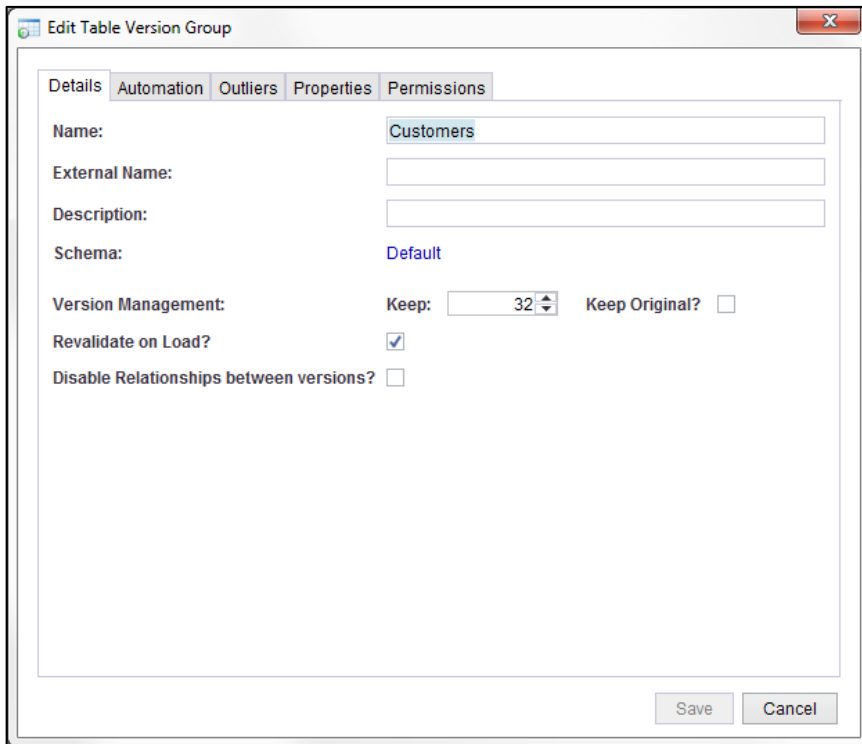
This will move the selected table into an existing Schema.

Projects

If you want the table to appear in an existing project, select the target project using the tick boxes.

Managing the version group

The version group provides a central point for managing aspects of all versions of the table. To edit the version group information, right-click on **Versions** and select **Manage Version Group:**

This dialog can also be found by right clicking on a table name and selecting **Advanced > Manage Version Group**.

Details Panel

Name

This is the name of the version group. This is the name that will be displayed on all versions of the table. The original table names are preserved and displayed as part of the information line on the explorer.

External Name

The name this group is known by outside of Experian Pandora

Description

A description about the table group

Schema

The schema that the table group (hence all versions of the table) belongs to

Version Management

The number of versions to keep before automatically pruning them

Revalidate on load
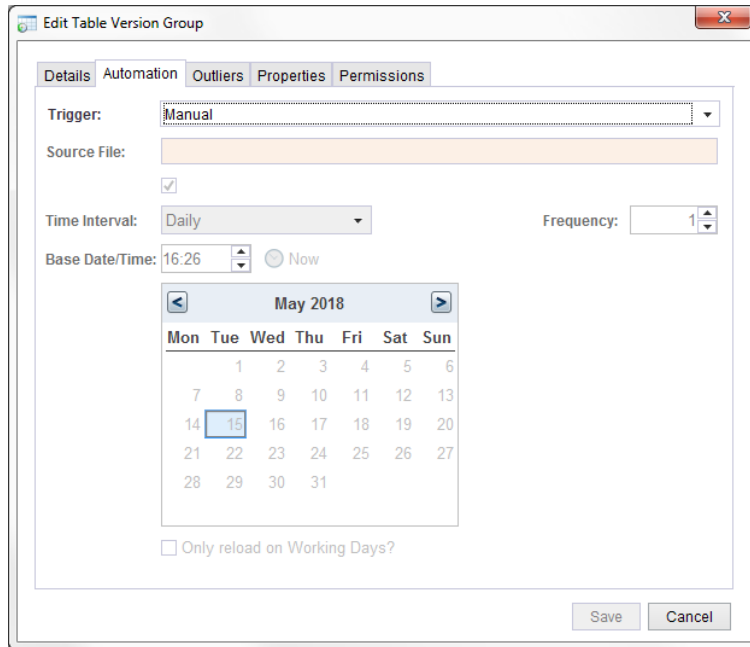
If the current version of a table has validation rules attached then this will ensure that when a new version is loaded, the validation rules are automatically applied to the new version, and the validation is actually run to get the results.

Automation Panel

Experian Pandora provides a number of methods for deciding how to load a new version. Whatever method chosen, you can always load a new version manually from the explorer.
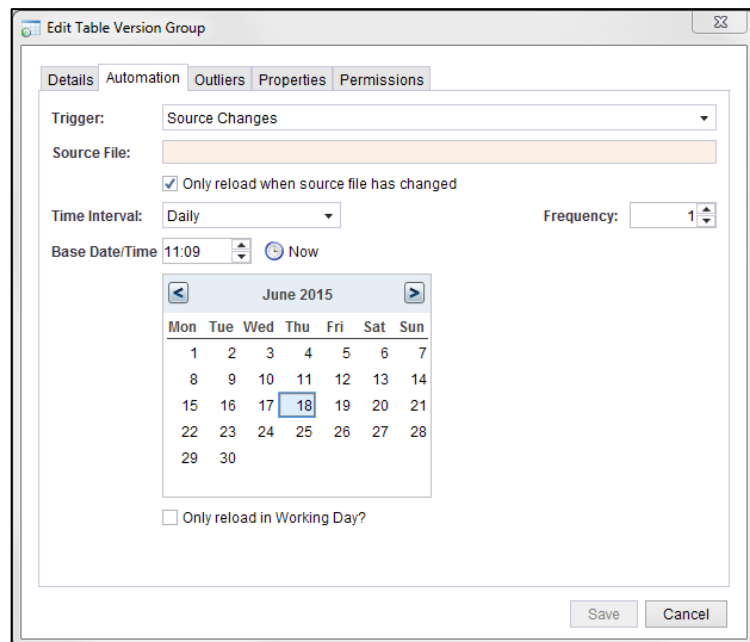
Automation Panel – Manual

The default is set to Manual where only manual versions can be created.

Automation Panel – Source Changes

Experian Pandora monitors the source for changes at a given time interval. Provided that interval arrives, the file is accessible **and** it has changed, then Experian Pandora will issue a new load process to load the new file. Note that this option should be used with care as it is up to the user to ensure that the file is actually available for loading, has actually changed, as opposed to being subject to some timestamp difference for other reasons, and that it is not being written to when Experian Pandora begins its load process. There are a number of strategies that can be used to get around this, most notably ensuring the check time is far later than the possible end time for creating/writing the input data file. Contact your support representative for more assistance for your particular use case.

Any table that is subject to source change monitoring will have an associated **Watchdog** task.

Source change monitoring is only possible for file based data sources currently. There is no reliable method for detecting change via a JDBC connector.

Source Name

The name of the input file to load from. If the source filename is not absolute (e.g. not preceded by a drive letter or file separator) the file is assumed to exist in the import directory associated with the Connection.

Time Interval

The interval to attempt to load the source which can be one of the following:

Time Interval	Description
By Intervals in Minutes	Every n minutes
Hourly	Every n hours
Daily	Every n days
Weekly	Every n weeks
Monthly	Every n months

Frequency

The number of Time Intervals between checking for the existence of the input source.
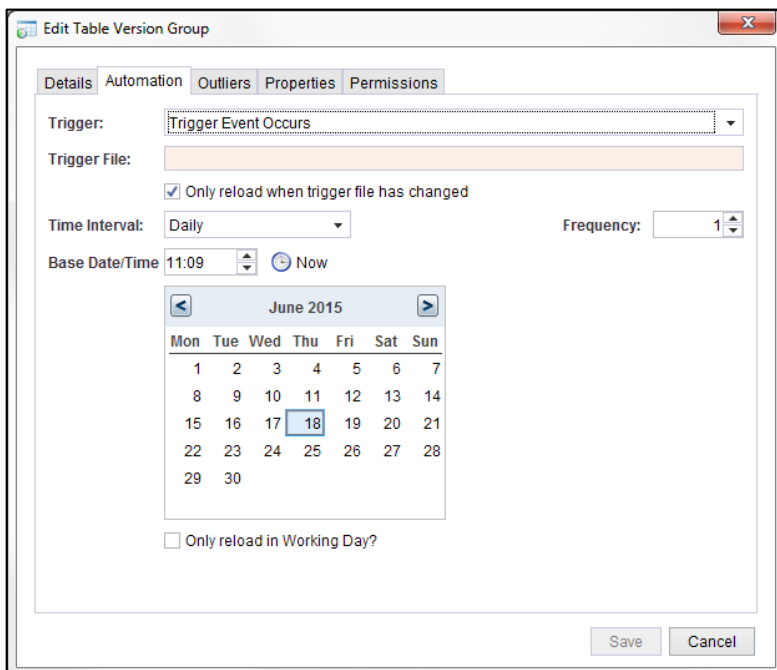
Base Date & Time

The starting date and time from when to start checking.

Working Days Only

Ignores any time intervals that arise on non-working days to ensure activity only occurs on working days.

Automation Panel – Trigger Event Occurs

An alternative to monitoring the actual source file arrival is to monitor the arrival of a trigger event. Currently this is the arrival of a file whose contents are simply the name of the file (or database table) to actually load. This trigger file will be placed in **\triggers** of the root directory. Experian Pandora monitors the trigger file for changes at a given time interval. Provided that interval arrives, the trigger file is accessible **and** it has a changed timestamp, then Experian Pandora will issue a new load process to load the new file or table referred to by the contents of the trigger file.

Source Name

The name of the input file (or database table) to load from. Defaults to the last known file/table loaded. This will need to be in the same Connection directory that the original file was placed in. If the trigger filename is not absolute (e.g. not preceded by a drive letter or file separator or the correct path) the file is assumed to exist in the import directory associated with the Connection.

Time Interval

The interval to attempt to load the source. Can be one of:

Time Interval	Description
In intervals of minutes	Every n minutes
Hourly	Every n hours
Daily	Every n days
Weekly	Every n weeks
Monthly	Every n months

Frequency

The number of Time Intervals between checking for the existence of the input source.

Base Date & Time

The starting date and time from when to start checking.

Working Days Only

Ignores any time intervals that arise on non-working days to ensure activity only occurs on working days.

Choosing a table group or a table which is part of a table group allows the **Graph Over Time** drilldown option. Experian Pandora automatically compares all statistics for all versions of the selected table and presents its finding in a graph.

Viewing Table Versions

Right-click on the **Versions** folder of the table and select **Versions** to view all the information about the Versions loaded. A Tables drilldown appears with all the different versions of the selected table.

Table metadata changes over time

As multiple versions of the table are loaded, Experian Pandora provides the ability to look at profile changes over time. Right click on **Versions** of the table and select **Changes over Time > Graph**

The graph appears on the Experian Pandora desktop. The dropdown on the toolbar contains **all** the statistics over time with those that have changed between versions marked in **bold**. Flipping this changes what is shown on the Y axis of the graph. Hovering over any bar gives the actual details of the statistic, and all of the bars can be clicked on to drilldown to the relevant supporting data or metadata.

The two toggle buttons on the far right of the graph provide an optional trend line and alternate logarithmic scale.

Changes can also be viewed in a tabular format. This allows all changes to be displayed at once for the versions of a table. Click on the table and select **Changes over Time > List**

Value	Clients v1	Clients v2	Clients v3
Row Count	33,415	▼ 29,966	▼ 27,632
Column Count	13	13	13
Columns Added	0	0	0
Columns Moved	0	0	0
Columns Removed	0	0	0
Unique Values	209,360	▼ 190,959	▼ 178,276
Unique Formats	18,314	▼ 17,169	▼ 16,431
Alphanumeric Values	206,382	▼ 188,189	▼ 175,645
Integer Values	2,978	▼ 2,770	▼ 2,631
Decimal Values	0	0	0
Number Values	2,978	▼ 2,770	▼ 2,631
Money Values	0	0	0
Date Values	0	0	0
Alphanumeric Formats	18,305	▼ 17,160	▼ 16,422
Integer Formats	9	9	9
Decimal Formats	0	0	0
Number Formats	9	9	9
Money Formats	0	0	0
Score	0	0	0
Rules	0	0	0
Rows Passed	0	0	0
Rows Failed	0	0	0
Rows Considered	0	0	0
Rows Ignored	0	0	0

The above drilldown presents a column for each version with a row for each of the profiling statistics that have changed over time. Red arrows and highlighting show the statistic is decreasing, whereas green shows that it is increasing. Black indicates no change.

All columns for all versions

Right-click on the **Columns** folder of the table and select **All Column Versions**. This provides a Columns drilldown of every column in every version of the table.

Highest column versions

Right-click on the **Columns** folder of the table and select **Highest Column Versions**. This produces a Columns drilldown of all uniquely named columns and the highest version they occur at.

Column changes over time

As multiple versions of the table are loaded, Experian Pandora provides the ability to look at profile changes for the columns of the table over time. Right click on any version of the column and select **Changes over Time > Graph**

The graph appears on the Experian Pandora desktop. The dropdown on the toolbar contains **only** the column profile statistics that have changed between versions. Flipping this changes what is shown on the Y axis of the graph. Hovering over any bar gives the actual details of the statistic, and all of the bars can be clicked on to drilldown to the relevant supporting data or metadata. The two toggle buttons on the far right of the graph provide an optional trend line and alternate logarithmic scale.
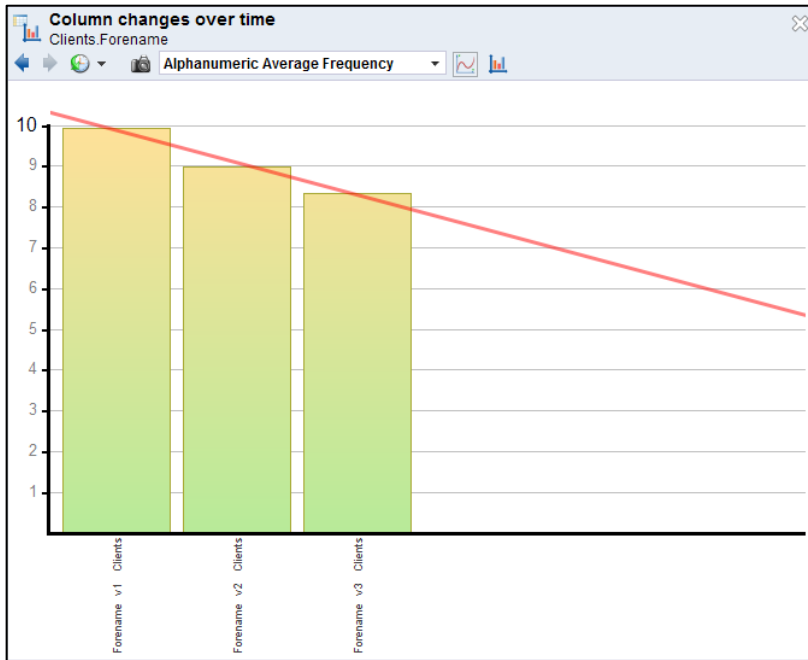
Changes can also be viewed in a tabular format. Click on the column and select **Changes over Time > List** which provides comprehensive information.

Changes to Forename over time

Value	Forename v1 Clients	Forename v2 Clients	Forename v3 Clients
Value Count	33,415	▼ 29,966	▼ 27,632
Distribution	10.09	▲ 11.17	▲ 12.03
Unique Value Count	3,372	▼ 3,348	▼ 3,325
Unique Format Count	52	52	52
Null Count	0	0	0
Blank Count	0	0	0
Zero Count	0	0	0
Negative Count	0	0	0
Shortest Length	1	1	1
Longest Length	16	16	16
Precision	2	2	2
Scale	0	0	0
Sum	41	41	41
Average	0	▲ 0	▲ 0
Average Length	5	5	5
Average Frequency	9.91	▼ 8.95	▼ 8.31
Alphanumeric Value Count	33,414	▼ 29,965	▼ 27,631
Alphanumeric Unique Value...	3,371	▼ 3,347	▼ 3,324
Alphanumeric Distribution	10.09	▲ 11.17	▲ 12.03
Integer Distribution	100	100	100

Relationships between versions

Using discovered relationships between different versions of a Table can help understand exactly what values have changed between versions and what records have changed. To find relationships just between versions of the same Table, click on the **Versions** folder of the Table to bring the drilldown, select which tables you want to find relationships from, right click on your selection and select **Find Relationships**.

Notes for all versions

From the **Notes** folder of a Table, right-click and select **Notes (All Versions)**. The Note Search dialog appears scoped for the versions of the Table only.

Automatic Novation

Experian Pandora provides automatic novation on metadata and rules when a new version of a table is loaded.

Table Validation Rules

Any validation rules applied at a table level are automatically applied from the previously highest version to the new version being loaded including references to Validation measures.

Table Expressions

Any Table Expressions are automatically applied from the previously highest version to the new version being loaded. If any Table Expression refers to a lookup table (i.e. by using the "Table Lookup" function) the version of the lookup table will also be novated according to the Novation Option set for the Table Expression.

Lookup table references can be re-novated at any time after the table has been loaded by accessing the **Re-novate** right-click option from an individual table, Schema, Project or globally (from the **Tables** item on the Explorer).

Column Quality Rules

Any quality rules applied at a column level are automatically applied from the previously highest version to the new version being loaded provided the column existed at the previous version.

Table Documented Metadata

Documented metadata on the table is automatically applied from the previously highest version to the new version being loaded.

Column Documented Metadata

Documented metadata on every column is automatically applied from the corresponding column on previously highest version of the table to the new version being loaded provided the column existed at the previous version.

Saved Relationships

Relationships are automatically applied from the previously highest version to the new version being loaded. Any saved relationship is examined on the previous version and a corresponding discovered relationship is searched for, replacing the previous version of the table with the latest one. Provided a discovered relationship still exists then it is novated and saved.

Business Terms & Categories

Any terms or categories applied to columns at the previous version are carried through to the newly loaded version.

Drilldowns

Any drilldowns attached to the table are automatically novated to create a new version-named copy of the original with all references changed to point to new elements of the new table version. This includes data and metadata drilldowns, and even joins of rows or values of Saved Relationships for the Table. For drilldowns parented by relationships, additional controls to decide when novation should take place are available. For example, when a particular table in the relationship changes, this decides whether the Drilldown should re-novate itself.

Keys and Dependencies

Once a Key or Dependency has been evaluated, it is possible to manually set up Novation, by right-clicking the table your Keys/Dependencies are on and selecting **Edit**. In the edit window, there is a check box which makes it possible to have Keys/Dependencies novated.

Automatic Re-novation

There is an optional ability to automatically re-novate or revalidate when a new version of a reference table is loaded.

Novation Options

You can choose novation options at the point of saving any of the following:

- Drilldowns
- Relationships
- Table Expressions
- Views

The only novation option that is specific to an object is **Disabled**, which is currently only available for **Drilldowns**.

Novation options can be changed at any time by right-clicking the object in question and editing it.

Latest Version

This option will change the reference to the foreign table to be the highest current version of that table.

Match Version

This option will attempt to match the version for the foreign table to be that of the new version of the loaded table that the Table Expression is attached to. If no match is found it first searches up for the next highest version. Failing that it searches down for the next lowest.

No Change

This leaves the reference to the foreign table to be at the same version that it was created at.

Disabled

This option will prevent the selected drilldown from appearing in future versions of the parent table. This option is only available on saved Drilldowns.

Data Validation

Validation and Quality Rules

Experian Pandora provides the ability to define sets of rules that can be applied to a given Table or Column to measure or validate characteristics about the data. Experian Pandora provides a wide range of built-in functions that can be used for this purpose or user-defined functions can be created in the Libraries node for multiple reuse.

What is a Rule?

In terms of Experian Pandora, a rule is the application of a built-in or user-defined validation function that returns a simple true (pass) or false (fail) result. The function itself could perform any kind of transformation, standardisation, if-then-else logic, etc.; whatever is required to distil the data passed to the function down to a pass or fail.

Types of Rules

Experian Pandora models two types of rules which are identical in the way they are defined. The difference is in where they are applied:

Validation Rules

The purpose of a Validation Rule is to validate records (rows) in a Table in order to ascertain the overall measure of validity for that Table. This would typically be done as part of an on-going validation process, for example, monitoring the results of transformation rules applied in an external Data Integration process.

Validation rules can use the values of one or more columns on each record and the rules are applied to a Table.

Validation rules tend to be used to check the syntax of data and its conformity to either metadata or known rules. This is often from the viewpoint of the content within a column, but can also extend to:

- **Dependency checks** – A column should have a certain value when another column has a certain value);

- **Referential checks** – A row should have a corresponding row in some other table.

Quality Rules

The purpose of a Quality Rule is to check data in a Column in terms of its *suitability* to the business or a given business process. For example:

- **Validity** – The country code is valid; the telephone number is properly formatted; the zip code is known; etc.

- **Conformance** – The ISIN is valid and known; the product code is valid and exists; etc.

- **Accuracy** – The Total Amount does equal the Price x Quantity; the discount code is valid and correctly calculated.

- **Completeness** – All required fields in a record have appropriate values.

The functions used for Quality Rules are identical to those used for Validation Rules. You are even allowed to use values from multiple Columns when deciding a Quality Rule. The principle difference is *where* the results are recorded. In the case of a Quality Rule, regardless of the inputs used, the result is recorded against the Column the rule is attached to. In the case of Validation Rules, it is always recorded against the Table.

Separation of Rules

Validation Rules and Quality Rules are aggregated entirely separately with the Column Quality rules having no impact to the results for a Table or the results of other columns.

Filtering data used in the evaluation of a Rule

You can reference a Table Filter when defining a Validation or Quality rule. This pre-filters data passed to the Validation Engine using the referenced Table Filter to ensure only required rows are considered for measurement. Filters can be applied at a Table, Column

or Rule level. Rules inherit Filters from their parents (Tables or Columns) as an *additional* filter, e.g. the parent and Rule filters are *ANDed* together.

Applying a business measurement to Rules

The results of a Rule naturally calculate the volume of records that pass or fail. The problem is that, in business terms, this is potentially misleading. Take for example a validation of a series of transaction records. 90% may pass a certain rule, giving the belief that most of the transaction data is fine. However, the 10% of records that failed may actually account for 90% of the financial value of the entire set of transactions. You can define a **Measure** to be used that alters the way that a Rule is scored to give it a business weighting. In this example, the measure could sensibly have been the **amount** of the transaction, which would have led to a 10% pass rate.

Measures can be applied to a Table and overridden at the Column level. Measures must produce a decimal or integer result to be valid. For instances of measures where the result is not a calculable value, you can define a default measure to use instead. This can be simply **1**, the biggest measure in the table, the smallest measure in the table or the average of all measures in the table for the rows that were considered.

Scoring

The score for each Rule is by default simply the number of rows that passed the rule, relative to the number of rows in the table represented as a percentage of the rows considered. For unfiltered rules this is the entire set of rows in the table. When a Measure is used, the score is the sum of the measures for rows that passed divided by the sum of measures for all rows considered represented as a percentage.

If you do not want null values to be calculated in the scoring, select the **Ignore Nulls?** checkbox when you are adding or editing a rule. This means that nulls are ignored when the rule is used.

Rule Categories

Validation and Quality rules can be further clustered by attributing them to a **Rule Category**. These are defined in the Settings and simply have a name and description. Once defined, they can be referenced by the definition of a rule. This then provides a further level of rule aggregation within a Table or Column. For example, you could have the Rule Categories for *Appropriateness* and *Completeness* and attribute different rules to each group, giving you topical sub-scores for the Table or Column.

How Rule Scores Are Combined

Each Rule Categories and containing Table or Column has an overall score. Each score is calculated by combining the results of the contained rules. For example, if you applied two rules to a Column, the overall score for that Column would be calculated by counting up the number of considered rows that pass **both** of the rules.

It is possible to exclude a rule from any such total. This is useful where you want to measure a particular characteristic of a Table or Column but not unduly influence its overall score. To do this unselect the **Include in Totals?** checkbox when you are adding or editing a rule.

Thresholds

Every Column or Table has two quality thresholds, expressed as a percentage. The Low Validation Threshold and High Validation Threshold columns are added to your table by default.

Lower Threshold

The lower threshold determines whether the quality of the data is acceptable.

Upper Threshold

The upper threshold is the target objective for the quality of the data. By default the target quality is 100%.

Rules inherit the threshold of the object they are attached to. These thresholds are for visualisation purposes only. By manipulating these two thresholds, the display of the quality of data will become red, amber or green.

1. Results that match, or exceed, the upper threshold will be displayed in green.
2. Results that are below the lower threshold will be displayed in red.
3. Results that equal to, or higher than, the lower threshold, but are less than the upper threshold are displayed in amber.

Rule Thresholds

If there is a specific low and high threshold that you use repeatedly, a Rule Threshold can be set up to store those values and select from a list when editing the Table or Column's validation settings.

These Rule Thresholds can be found in the 'Settings' section of the explorer. You can create these by right clicking on the option and selecting 'New…'.
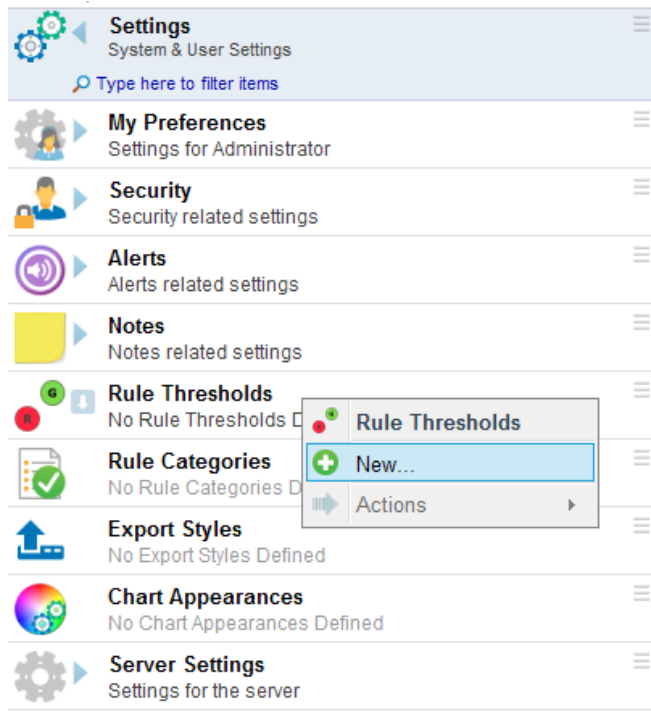
On the Create Rule Threshold dialog, give your Rule Threshold a name and an optional description, and set the required minimum and maximum thresholds:
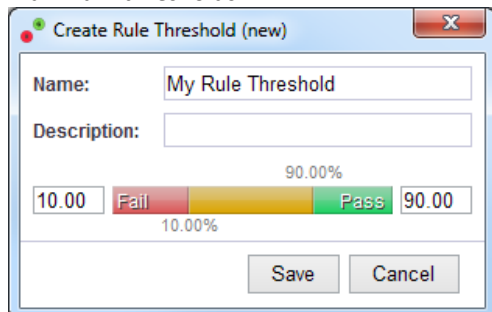
Once you have saved the Rule Threshold you can apply it to any Table or Column.

You can review all your Rule Thresholds in a drilldown by clicking on 'Rule Thresholds' on the right click menu in the explorer.

Editing a Rule Threshold will update any Table or Column referencing it including its validation status of 'Pass', 'OK' or 'Fail'.

Managing Table Validation Rules

Right-click on a Table, select **Edit** and then click on the **Rules** button:
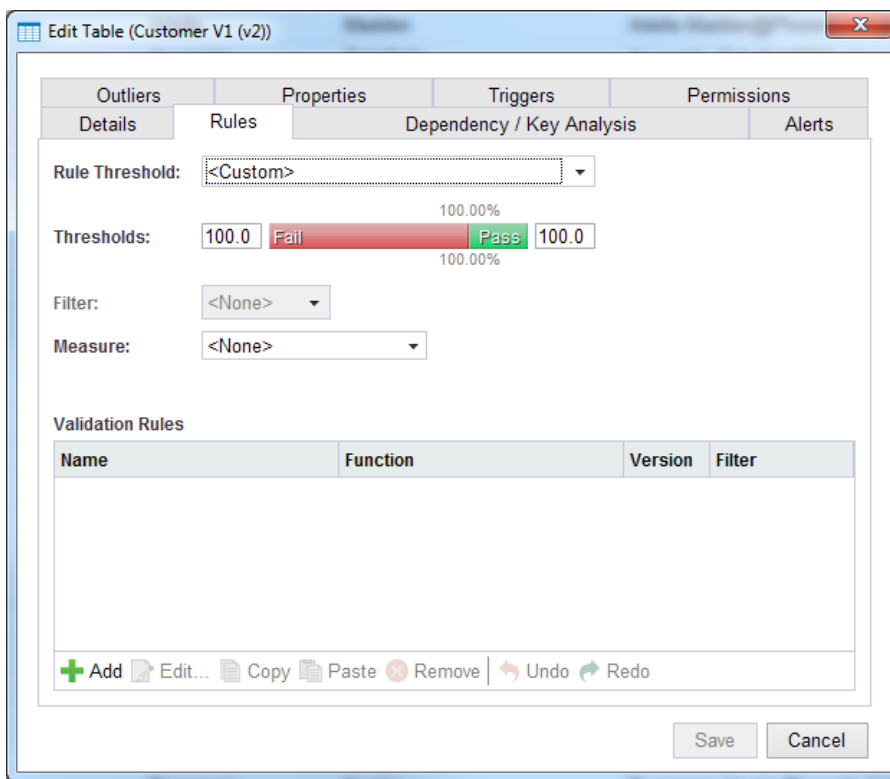
From this dialog you can Add, Edit or Remove Validation Rules as well as set or change the Measure and Filter definitions for the Table as a whole. You can also apply a pre-defined Rule Threshold.

Rule Threshold

This allows you to select a pre-defined Rule Threshold.

Threshold Control

This control allows you to set the lower and upper validation thresholds. By default they are both set to 100, which means that everything has to pass for the result to be green. You can drag the lines for either of the two thresholds, or simply type into the boxes either side of the control:
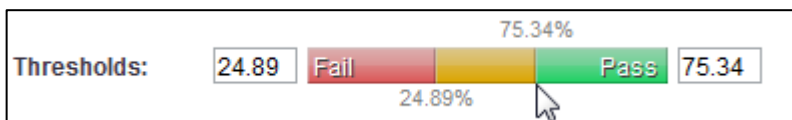
As this is evaluated for display purposes only, you can change this at any time, without the Table requiring re-validation.

Filter

This allows you to either have no filter or choose a previously saved filter to affect which rows are considered for the validation of any of the Rules on this Table. Select **<none>** to remove the filter.
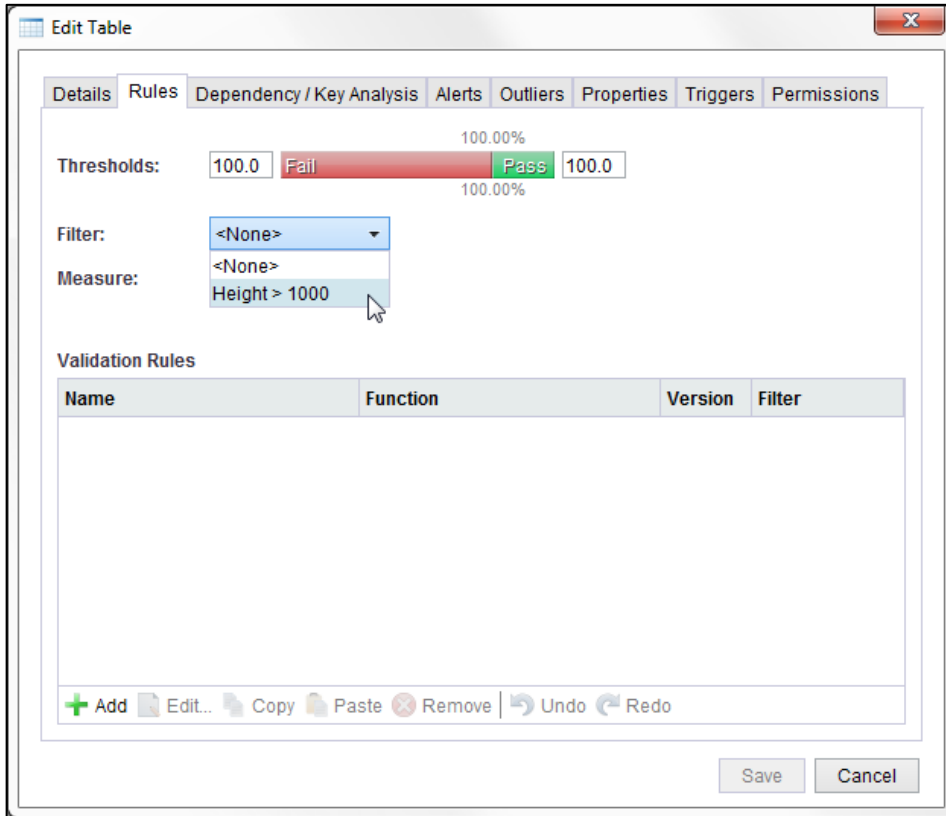
Measure

This allows you to either have no measure or choose any column or Expression in the table to be a measure. Select **<None>** to remove the measure setting.
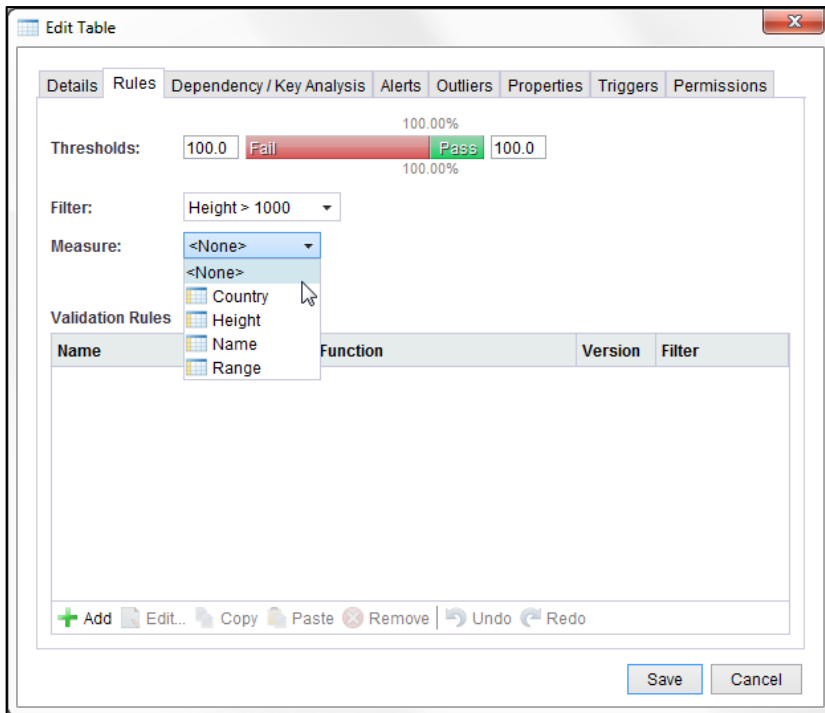
Column Measures

When choosing an existing column as a measure, the edit dialog validation panel will change as follows:

Selecting a Default Measure

When using a Column or Table Expression as a Measure, an additional control appears for the **Default Measure**. This provides a number of options for what Experian Pandora should do when the value in the chosen column for a given row is not a calculable value, i.e. it is null or some form of alphanumeric value.

The possible choices are:

Default Measure Option	Description
Average Measure Value	Use the average of all values across all rows in the table for the measure column.
Minimum Measure Value	Use the smallest value on any row for the measure column.
Maximum Measure Value	Use the largest value on any row for the measure column.
Default Measure of One	Use the number 1. This is the default setting.

Validation Rules Panel

This panel contains a list of all the Validation Rules attached to the table.

Adding a Validation Rule

To add a validation rule, simply click on the green **Add** button. The **New Validation Rule** dialog will appear:

An alternative method of quickly adding validation rule is by right-clicking on the **Rules** folder, under the Table in the explorer, and select **New**.

Once you have the above dialog, you can then name your Rule, give it a description and select which function you wish to use. When selecting a function, the dialog looks like this:
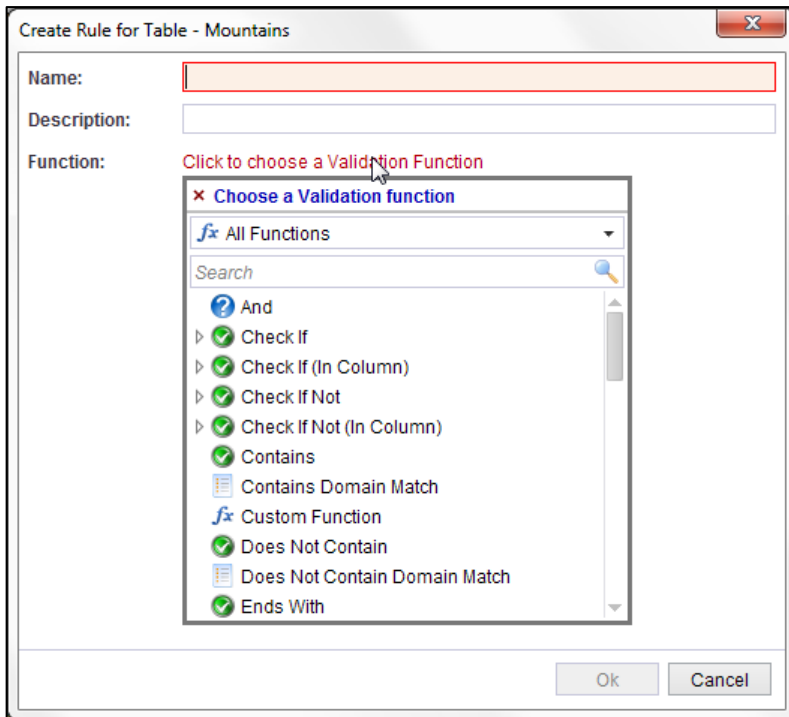
The additional component is a searchable list of all available validation functions, including both built-in and user-defined functions. The top dropdown allows you to switch the displayed list between all of the available functions and specific function libraries. Simply click on one and a new panel will be added to the dialog which will then enable you to select a filter and a rule category for your rule, if you so wish. It will also allow you to input any required parameters. The following is an example of what the dialog looks like if you choose **Contains**:
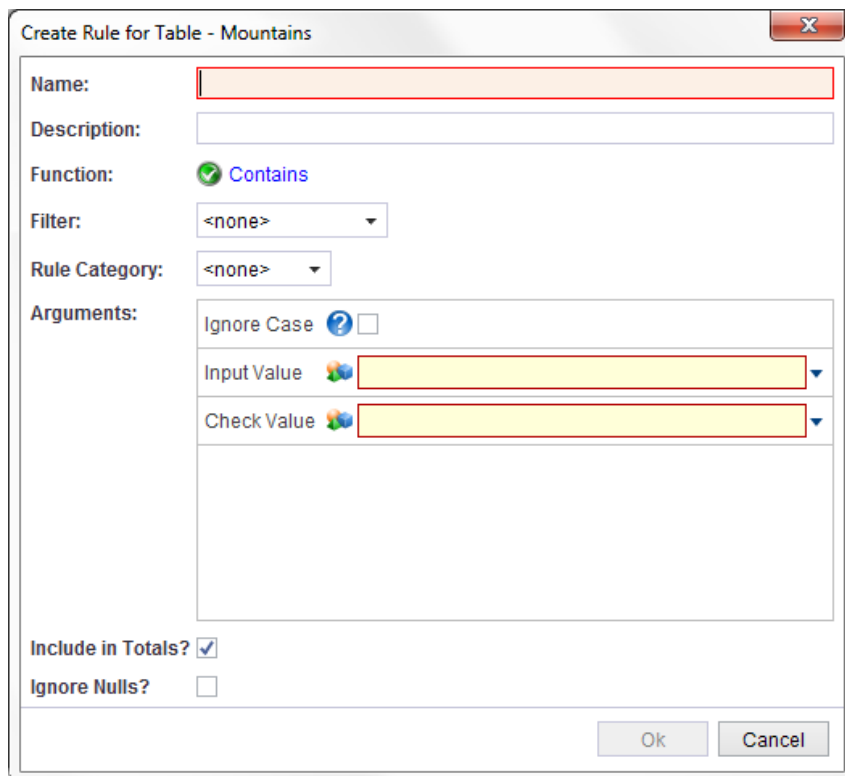
The arguments are the same as the ones you would see if you used this function on the expression editor in a drilldown.

![Experian logo]

Removing a Validation Rule

Simply select the rule in the Validation panel and click on the **Remove** button. You are not prompted when you remove the rule, but no changes are actually saved until you click the **Save** button.
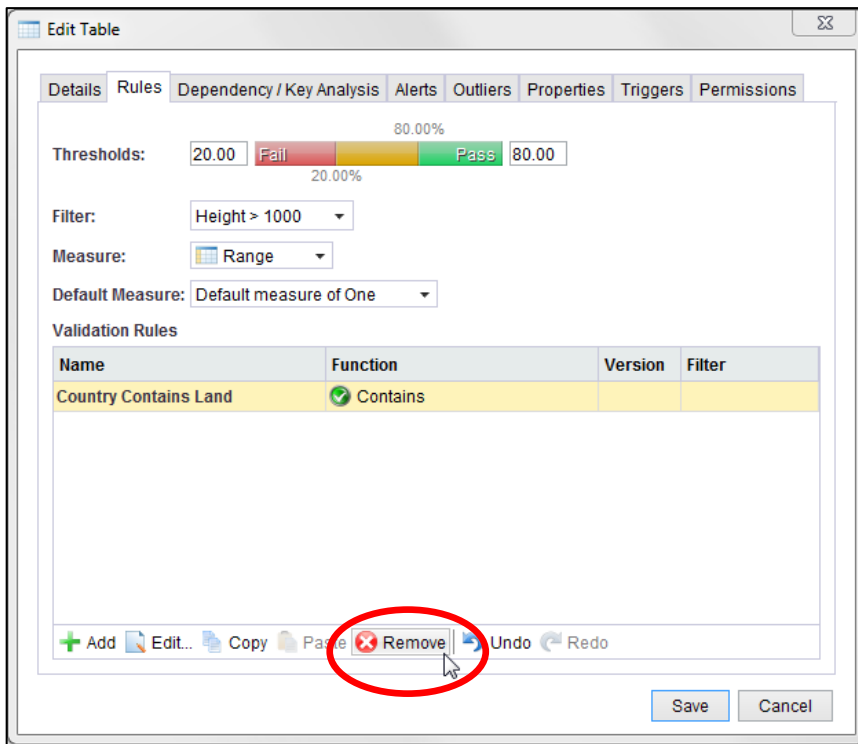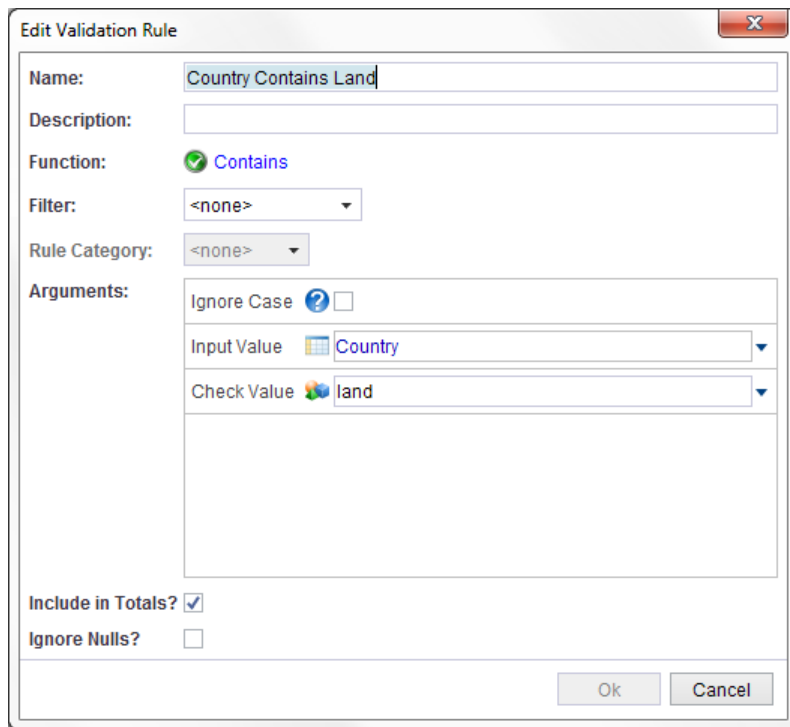
Alternatively, you can right-click on the Rule in the explorer and select **Remove**.

Undo/Redo

You have multiple undo and redo history, allowing you to undo changes or re-apply them. Use the **Undo** and **Redo** buttons, accordingly.

Editing a Validation Rule
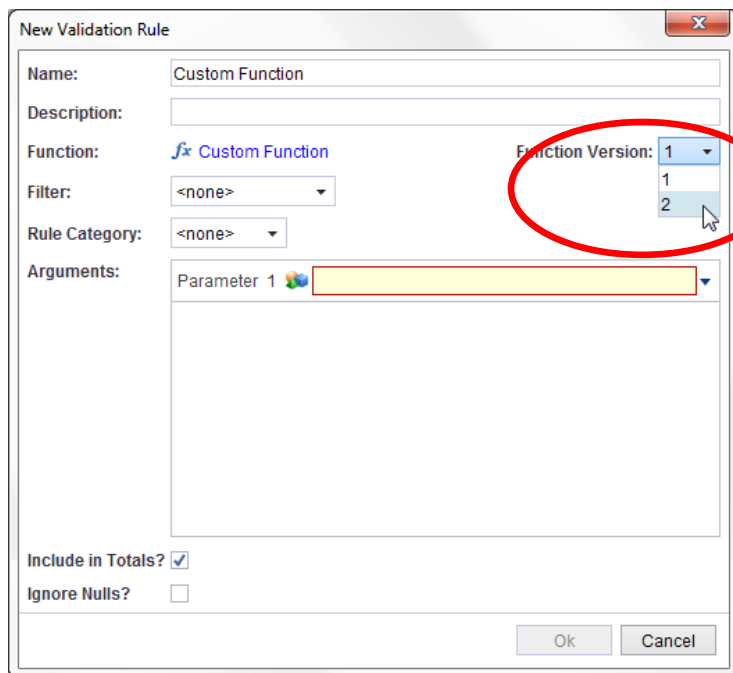
Select the validation rule you want to edit and click on **Edit.** The **Edit Validation Rule** dialog will appear. You can then use this to edit the arguments, filter or Rule Category for the Rule:

Changing Function Versions

If you are using a Custom Validation Function, the function may be versioned. In this case, the use of the Function in the Validation Rule is **at a specific version** ensuring the behaviour of the function does not change as the definition of a Validation Function at a specific version is fixed.

To change the version of the function that is attached, simply bring up the edit validation rule dialog and change the selected function version:

Managing Column Quality Rules

Right-click on any Column of a Table in the Explorer, select **Edit** and click on the **Rules** button.
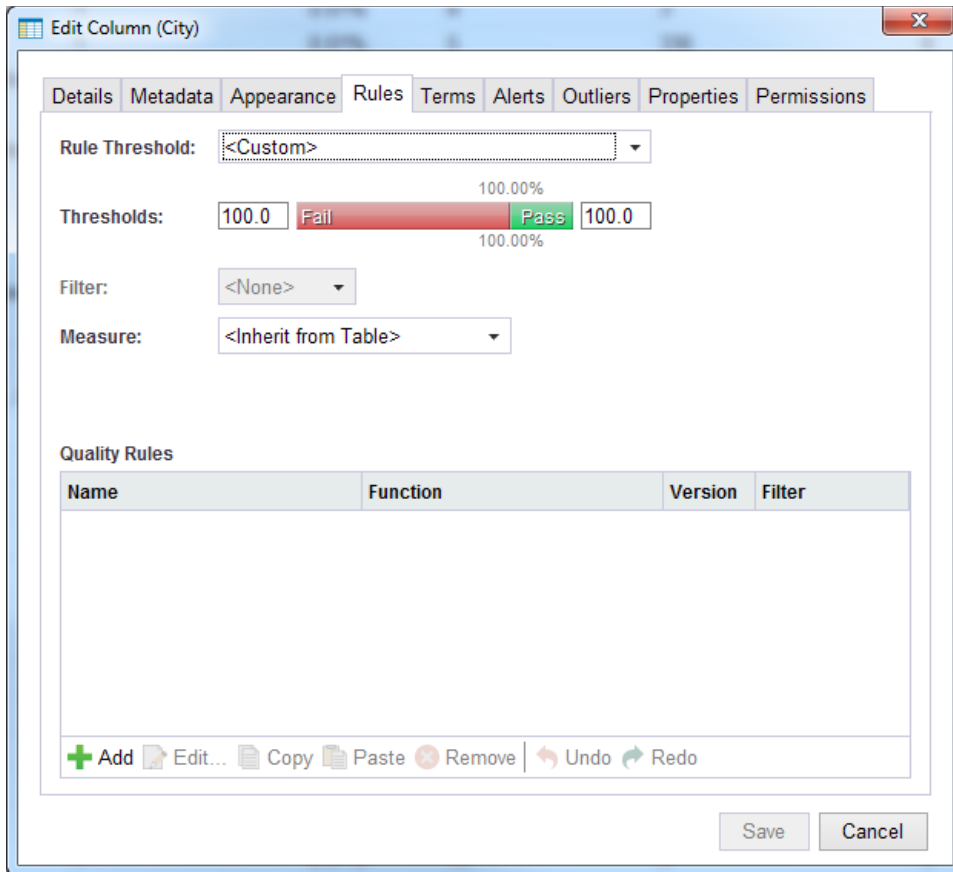
From this dialog you can Add, Edit or Remove Validation Rules as well as set or change the Measure and Filter definitions for the Column as a whole. You can also apply a pre-defined Rule Threshold.

Rule Threshold

This allows you to select a pre-defined Rule Threshold.

Threshold Control

This control allows you to set the lower and upper validation thresholds. By default they are both set to 100, which means that everything has to pass for the result to be green. You can drag the lines for either of the two thresholds, or simply type into the boxes either side of the control:
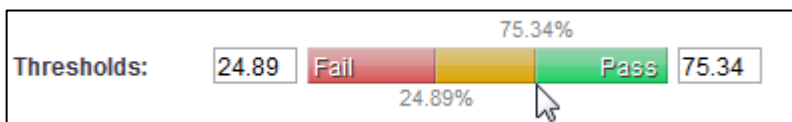
As this is evaluated for display purposes only, you can change this at any time, without the Column requiring re-validation.

Filter

This allows you to either have no filter or choose a previously saved filter to affect which rows are considered for the validation of any of the Rules on this Column. Select **<none>** to remove the filter.
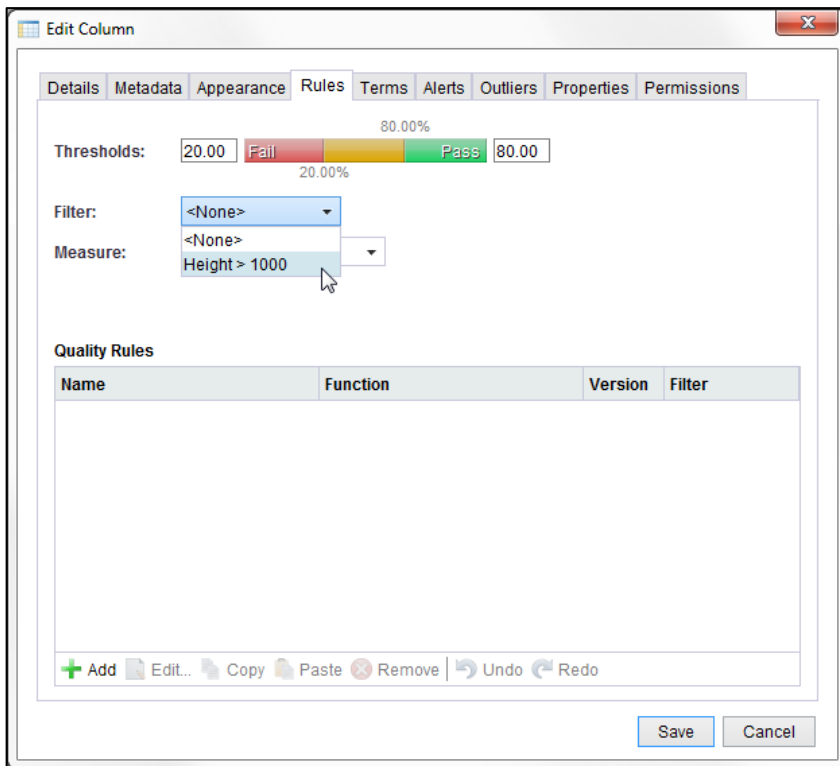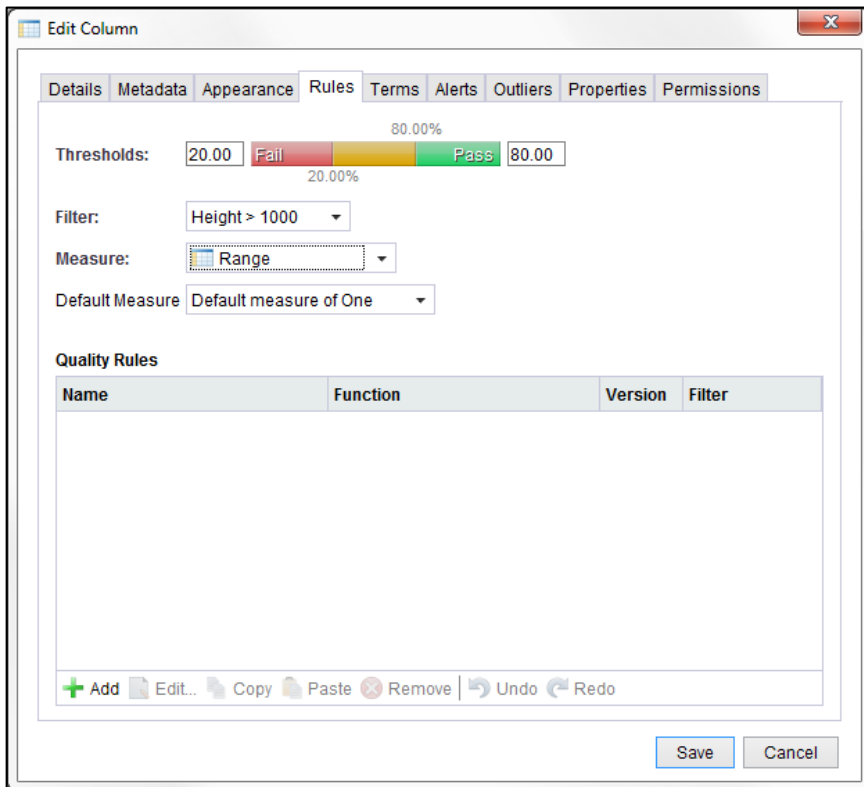

Measures

This allows you to either have no Measure or choose any Column, or Table Expression in the Table to be a Measure. It also allows you to inherit the Table's Measure definition, which is the default setting. Select **<None>** to remove the Measure setting.
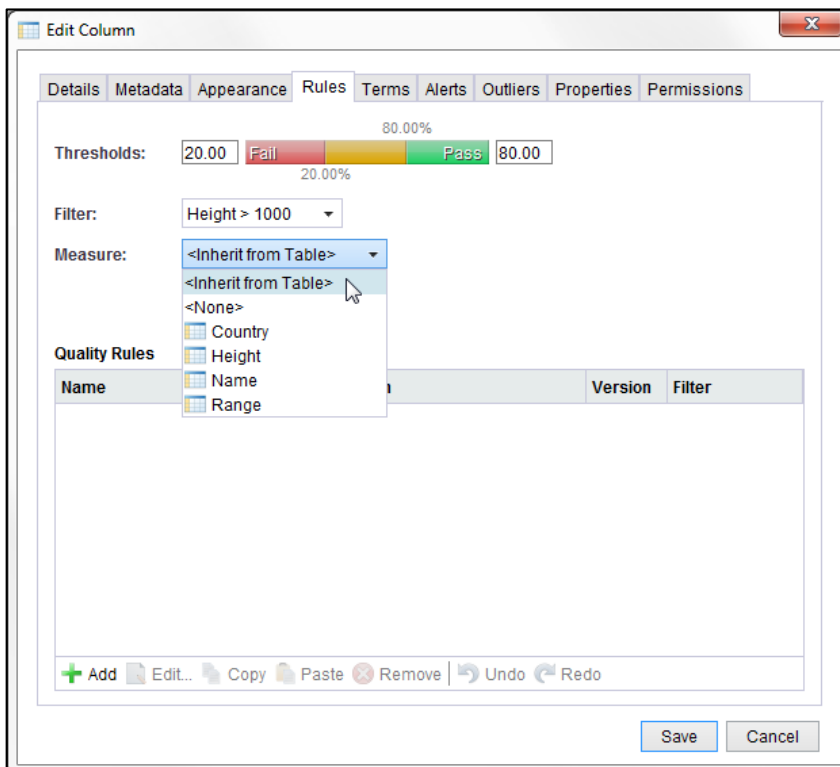
Column Measures

When choosing an existing Column or a Table Expression as a measure, the edit dialog validation panel will change as follows:

Selecting a Default Measure

When using a Column or Table Expression as a Measure, an additional control appears for the **Default Measure**. This provides a number of options for what Experian Pandora should do when the value in the chosen column for a given row is not a calculable value, i.e. it is null or some form of alphanumeric value.

The possible choices are:

Default Measure Option	Description
Average Measure Value	Use the average of all values across all rows in the table for the measure column.
Minimum Measure Value	Use the smallest value on any row for the measure column.
Maximum Measure Value	Use the largest value on any row for the measure column.
Default Measure of One	Use the number 1. This is the default setting.

Quality Rules Panel

This panel contains a list of all the Quality Rules attached to the Column.

Adding a Quality Rule

To add a Quality Rule, simply click on the green **Add** button. The **New Quality Rule** dialog will appear:
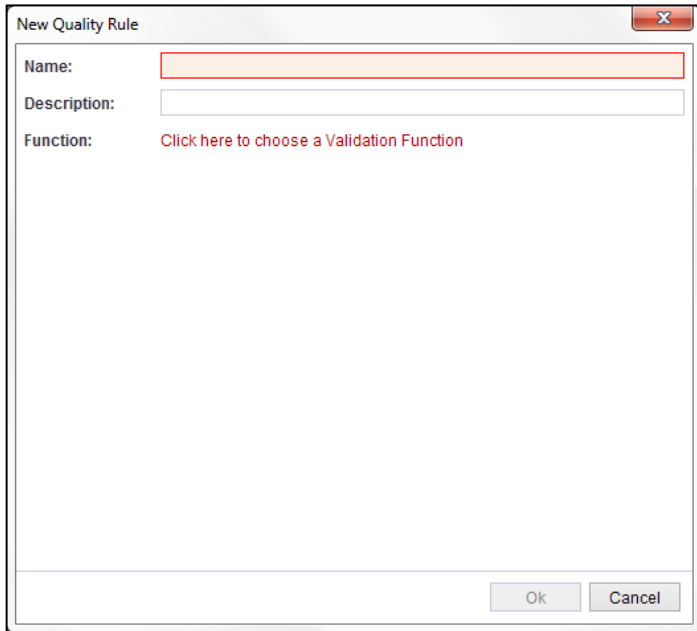
You can also add a Quality Rule to a Column by right clicking on the **Rules** folder, under the Column in the Explorer, and select **New**.

Once you have the above dialog, you can then name your Rule, give it a description and select which function you wish to use. When selecting a function, the dialog looks like this:
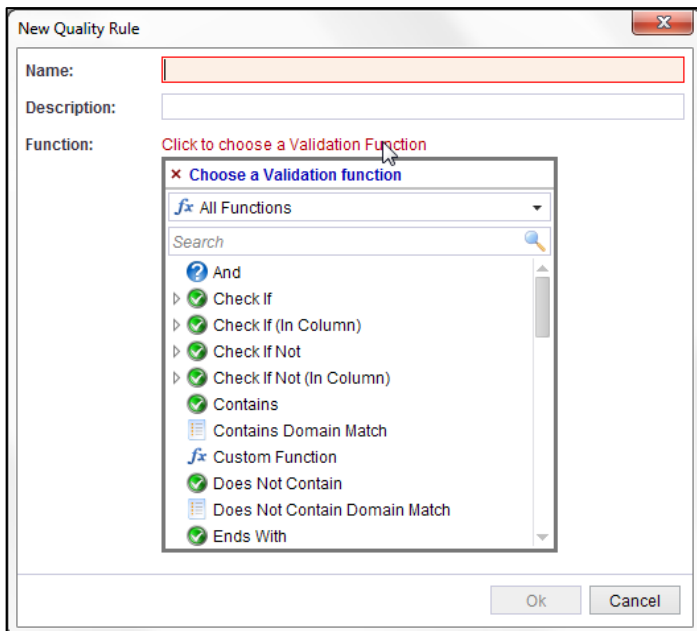
The additional component is a searchable list of all available Validation Functions, including both built-in and user-defined functions. The top dropdown allows you to switch the displayed list between all of the available functions and specific function libraries. Simply click one and a new panel will be added to the dialog which will then enable you to select a filter and a rule category for your rule, if you so wish. It will also allow you to input any required parameters.

The following is an example of what the dialog looks like if you choose **Contains**:
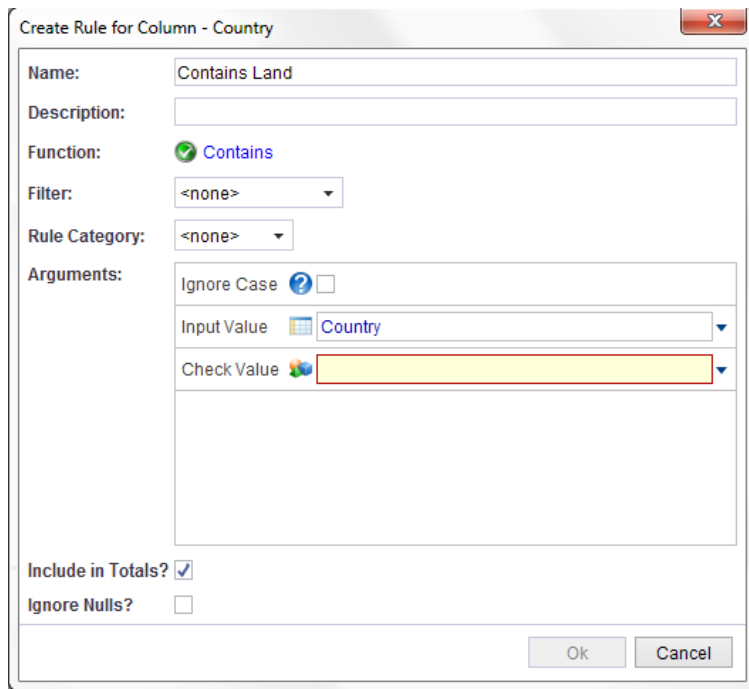
The arguments are the same as the ones you would see if you used this function in the Expression Editor in a drilldown.

Removing a Quality Rule

Simply select the rule in the Rules panel and click on the **Remove** button. You are not prompted when you remove the rule, but no changes are actually saved until you click the **Save** button. Alternatively, you can right-click on the Rule in the explorer and select **Remove**.
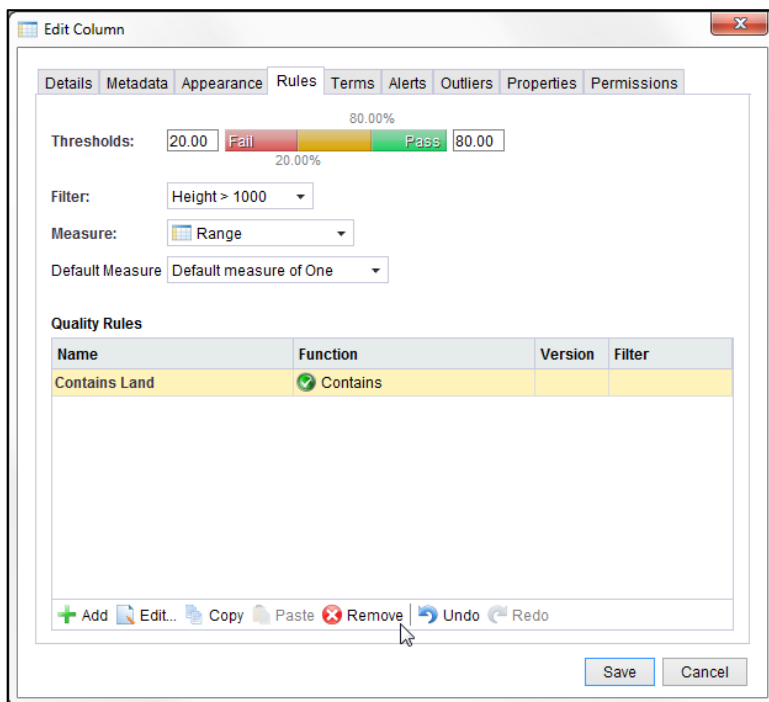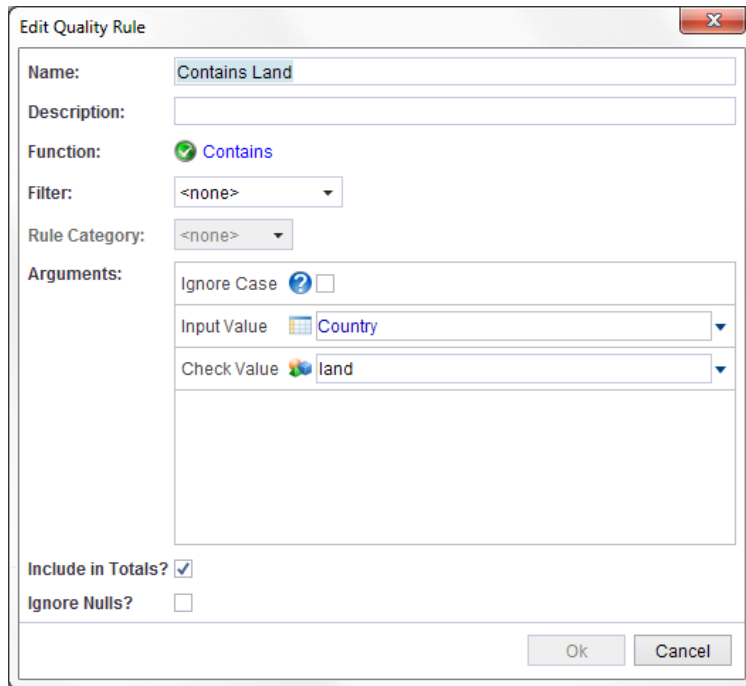
If you want to quickly remove a rule, you can also go to it in the explorer, right click on it and **Delete**.

Undo/Redo

You have multiple undo and redo history, allowing you to undo changes or re-apply them. Use the **Undo** and **Redo** buttons, accordingly.

Editing a Quality Rule

Select the Quality Rule you want to edit and click on **Edit**. The **Edit Quality Rule** dialog will appear. You can then use this to edit the arguments, filter or Rule Category for the Rule:

Changing Function Versions

If you are using a Custom Validation Function, defined in Libraries, the function may be versioned. In this case, the use of the Function in the Quality Rule is **at a specific version** ensuring the behaviour of the function does not change as the definition of a Validation Function at a specific version is fixed.

To change the version of the function that is attached, simply bring up the edit validation rule dialog and change the selected function version:

Executing Validation

Once all the rules have been applied to the table and/or column, you can validate it by selecting the option to **Validate Table**. This is available in numerous right-click menus in the table on the explorer:

From **Tables**, select **Validate Tables**
From any Table, select **Validate Table**
From any Column of a Table, select **Validate Table**
From the **Rules** folder of a Table, select **Validate Table**
From the **Rules** folder of a Column, select **Validate Table**

Visualising Rule Results

Validation Dashboard

To view the Validation Dashboard go into a Table's folder and click on **Rules**. The Validation Dashboard shows you all the relevant information about the Table, Rule Category or Rule you selected as well as a drilldown of the rules related to it:

From here you are able to drilldown onto rows of the table and the results of the rules applied. The buttons to the top right of the dashboard can be used for adding a new rule and re-validating the Table. The **Close** button is also provided in order to view the results without the dashboard.

Individual Validation Rule

By double clicking on one of the rules you are given a similar drilldown but with the results of only that rule as well as the **Score** drilldown underneath:

	Result	Country	Name	Height	Range
1	✔ Yes	Scotland	Ben Nevis	1344	Fort William to Loch Tre
2	✔ Yes	Scotland	Ben Macdui	1309	The Cairngorms
3	✔ Yes	Scotland	Braeriach	1296	The Cairngorms

All Rows for Mountains (Table)
Rule: Contains Land
2,367 Rows

Validate   New Rule...   Close

Score:	98.44% (Ok)	Passed Measure:	N/A
Failed Measure:	N/A	Passed Rows:	253 (98.44%)
Failed Rows:	4 (1.56%)	Considered:	257 (10.86%)
Ignored:	2,110 (89.14%)	Evaluated On:	30-Jul-2014 16:31:01
Validated by:	Administrator	Measure:	N/A
Default Measure:	Default measure of One	Pass Range:	100%
Fail Range:	0% - 90.0%	Filter:	Height > 1000

Overall Table Validation Rule Results

You can expand the **Rules** folder on the table in the explorer to see the list of rules on the table:

Rules
2 Rules – 98.44%
Type here to filter items

Contains Land
Score: 98.44%

Range Not Null
Score: 100%

You can go into a rule in the explorer to see the overall results for that rule:

Quality Dashboard

To view to the Quality Dashboard, go into a column's folder and click on **Rules**. The Quality Dashboard shows you all the relevant information about the Column, Rule Categories or Rule you selected as well as a drilldown of the rules related to it:

From here you are able to drilldown onto rows of the table and the results of the rules applied.

The options in the top left can be used for adding a new rule and validating for convenience. A close option is also given in order to view results in full screen.

Individual Quality Rule

By double clicking on one of the rules you are given a similar drilldown but with the results of only that rule as well as the **Score** drilldown underneath:

Overall Column Quality Rule Results

You can expand the **Rules** folder on the table in the explorer to see the list of rules on the table:

You can go into a rule in the explorer to see the overall results for that rule:

Score

Produces a drilldown of all the rows in the Table, with an additional column for the Rule's result, the value of the Measure for that row and whether the Measure was valid or not.

	Result	Country	Name	Height	Range
127	✔ Yes	Scotland	Creag Meagaidh - Creag Mhor	1069	Loch Lochy to Loch Laggan
128	✔ Yes	Scotland	Glas Maol	1068	Braemar to Montrose
129	✔ Yes	Scotland	White Mounth - Creag a'Ghlas-uillt	1068	Braemar to Montrose
130	✔ Yes	Scotland	Aonach Mor - Stob an Cul Choire	1068	Fort William to Loch Treig and Loch Leven
131	✔ Yes	Scotland	Mam Sodhail - Carn Coulavie	1068	Loch Duich to Cannich
132	✔ Yes	Scotland	Stob Binnein - Stob Coire an Lochain	1068	Loch Lomond to Strathyre
133	✔ Yes	Scotland	Sgurr Fhuaran	1067	Loch Duich to Cannich
134	✘ No	Wales	Crib y Ddysgl	1065	Snowdonia
135	✔ Yes	Scotland	Cairn of Claise	1064	Braemar to Montrose
136	✘ No	Wales	Carnedd Llewelyn	1064	Snowdonia
137	✔ Yes	Scotland	Binnein Mor - South Top	1062	Fort William to Loch Treig and Loch Leven
138	✔ Yes	Scotland	An Teallach - Bidein a'Ghlas Thuill	1062	Loch Maree to Loch Broom

Passed

Produces a drilldown of all the considered Rows in the Table where the result is a pass.

	Result	Country	Name	Height	Range
1	✔ Yes	Scotland	Ben Nevis	1344	Fort William to Loch Treig and Loch Leven
2	✔ Yes	Scotland	Ben Macdui	1309	The Cairngorms
3	✔ Yes	Scotland	Braeriach	1296	The Cairngorms
4	✔ Yes	Scotland	Ben Macdui - North Top	1295	The Cairngorms
5	✔ Yes	Scotland	Cairn Toul	1291	The Cairngorms
6	✔ Yes	Scotland	Braeriach - Carn na Criche	1265	The Cairngorms
7	✔ Yes	Scotland	Sgor an Lochain Uaine	1258	The Cairngorms
8	✔ Yes	Scotland	Ben Macdui - Stob Coire Sputan Dearg	1249	The Cairngorms
9	✔ Yes	Scotland	Cairn Gorm	1245	The Cairngorms
10	✔ Yes	Scotland	Braeriach - Stob Coire an Lochain	1235	The Cairngorms
11	✔ Yes	Scotland	Aonach Beag	1234	Fort William to Loch Treig and Loch Leven
12	✔ Yes	Scotland	Aonach Mor	1221	Fort William to Loch Treig and Loch Leven

Failed

Produces a drilldown of all the considered Rows in the Table where the result is a failure.

	Result	Country	Name	Height	Range
1	✘ No	Wales	Snowdon - Yr Wyddfa	1085	Snowdonia
2	✘ No	Wales	Crib y Ddysgl	1065	Snowdonia
3	✘ No	Wales	Carnedd Llewelyn	1064	Snowdonia
4	✘ No	Wales	Carnedd Dafydd	1044	Snowdonia

Considered

Produces a drilldown of all the considered Rows in the Table whether they passed or failed.

	Result	Country	Name	Height	Range
127	✔ Yes	Scotland	Creag Meagaidh - Creag Mhor	1069	Loch Lochy to Loch Laggan
128	✔ Yes	Scotland	Glas Maol	1068	Braemar to Montrose
129	✔ Yes	Scotland	White Mounth - Creag a'Ghlas-uillt	1068	Braemar to Montrose
130	✔ Yes	Scotland	Aonach Mor - Stob an Cul Choire	1068	Fort William to Loch Treig and Loch Leven
131	✔ Yes	Scotland	Mam Sodhail - Carn Coulavie	1068	Loch Duich to Cannich
132	✔ Yes	Scotland	Stob Binnein - Stob Coire an Lochain	1068	Loch Lomond to Strathyre
133	✔ Yes	Scotland	Sgurr Fhuaran	1067	Loch Duich to Cannich
134	✖ No	Wales	Crib y Ddysgl	1065	Snowdonia
135	✔ Yes	Scotland	Cairn of Claise	1064	Braemar to Montrose
136	✖ No	Wales	Carnedd Llewelyn	1064	Snowdonia
137	✔ Yes	Scotland	Binnein Mor - South Top	1062	Fort William to Loch Treig and Loch Leven
138	✔ Yes	Scotland	An Teallach - Bidein a'Ghlas Thuill	1062	Loch Maree to Loch Broom

Ignored

Produces a drilldown of all the rows in the Table that were not considered for the Rule,

|	Result	Country	Name	Height	Range
1		Scotland	Cairn Bannoch - Fafernie	1000	Braemar to Montrose
2		Scotland	Meall nan Tarmachan - Beinn nan Eachan	1000	Glen Lyon to Glen Dochart and Loch Tay
3		Scotland	Maoile Lunndaidh - Creag Toll a'Choin	1000	Kyle of Lochalsh to Garve
4		Scotland	Stob Ban	999	Fort William to Loch Treig and Loch Leven
5		Scotland	Sgurr Choinnich	999	Kyle of Lochalsh to Garve
6		Scotland	Sgurr nan Ceathreamhnan - Creag nan Clachan Geala	999	Loch Duich to Cannich
7		Wales	Glyder Fawr	999	Snowdonia
8		Scotland	Sgurr Breac	999	The Fannaichs
9		Scotland	Broad Cairn	998	Braemar to Montrose
10		Scotland	Sgurr nan Conbhairean - Creag a'Chaorainn	998	Glen Affric to Glen Moriston
11		Scotland	Stob Diamh	998	Glen Etive to Glen Lochy
12		Scotland	Ben Challum - South Top	998	Glen Lyon to Glen Dochart and Loch Tay

Rule Results over Time

Right-click any Rule, Validation or Quality, and select **Changes over Time > List**. This produces a list of versions for the Rule and how the results have changed over time. Trend icons show whether the trend is up or down. Background colouration shows the status at that version of red, amber or green.

Right-click any Rule, Validation or Quality, and select **Changes over Time > Graph**. This produces a graph of versions for the Rule, which shows how the results have changed over time. The graph only shows one aspect at a time, selectable from the dropdown on the toolbar. The score is shown as a percentage and the bars are solidly coloured according to whether the rule was red, amber or green at that version.
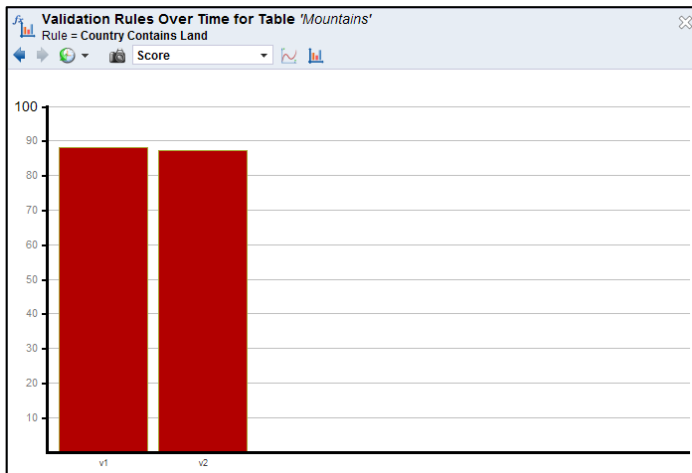
The other results are all shade-coloured as per the other graphs of this type in Experian Pandora.
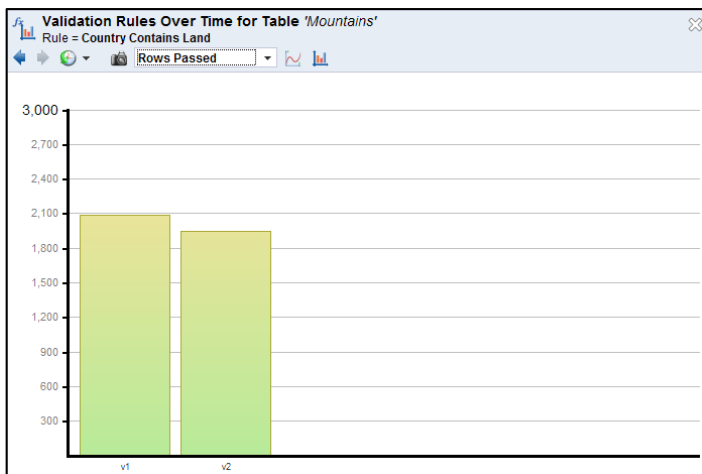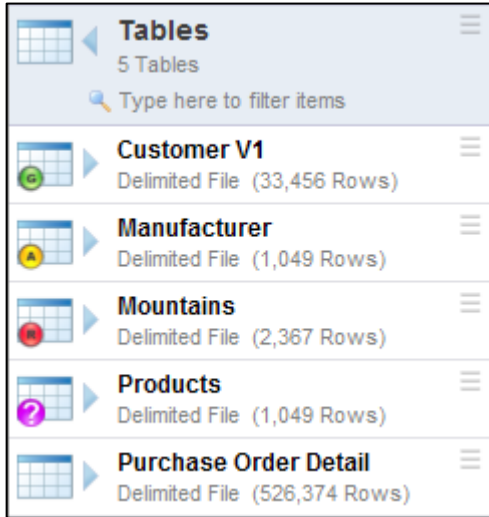

Clicking on any of the bars produces the corresponding drilldown for the version in question.

Rule Categories Results

The dashboard for a particular Rule Category can be found by right clicking on **Rules** and selecting **By Category > "Rule Category"**.

Validation Traffic Lights

Any Table or Column that has validation rules applied will have a traffic light attached to its icon. This traffic light will either be red, amber or green – to signify its status. If it has not been validated, a purple circle containing a question mark will appear instead. The image below shows all the possible states for a table.

The green circle with the **G** symbol signifies the target quality was achieved (pass).
The amber circle with the **A** symbol signifies above acceptable threshold but below target quality (ok).
The red circle with the **R** symbol signifies below acceptable threshold (fail).
The purple circle with the question mark signifies that the table validation is out of date or has not yet been validated.

All the tables and columns have their own highlight threshold levels.

Column Rules

We apply the following rules on the Columns of Customers:

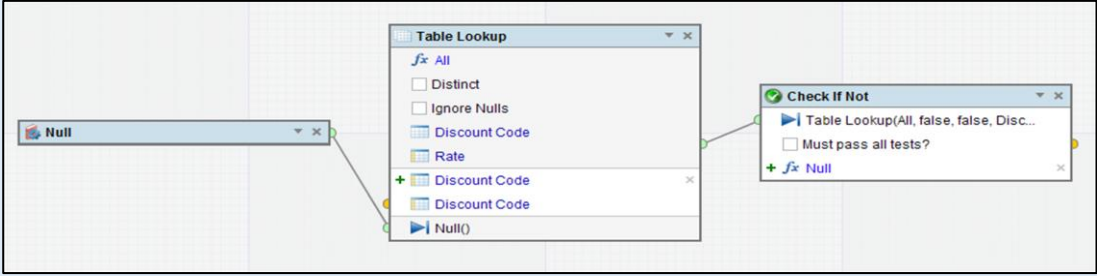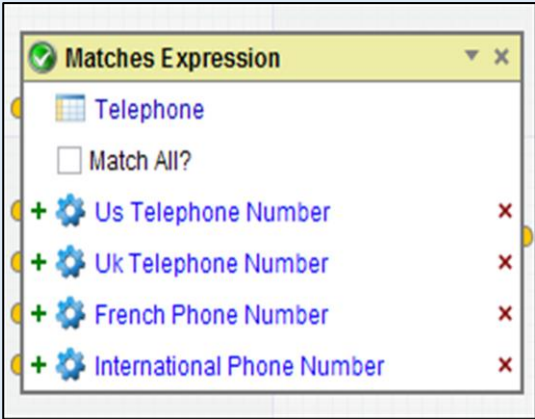| Column	Rule Description
Addressline	Value must be present – use the **Is Not Empty** built-in rule
Company Name	Value must be present – use the **Is Not Empty** built-in rule
Country	Must be a valid country name – use the **Is Valid Country Name** built-in rule
Customer Id	Ensure it is unique. Use the built-in rule **Is Unique**
Discount Code	Ensure it is valid. Use a custom validation function **Is Valid Discount Code** which ensures it is not empty and checks the **DISCOUNT_CODE** reference table for a match
Email	Ensure it is a valid email address. Use the built-in function **Matches Expression** and the supplied Business Constant **Email Address Multi** which is an expression for matching multiple formats of email addresses
Forename	Ensure it is present. Use the built-in function **Is Not Empty**. Ensure it is alphabetic. Use the built-in function **Is Alphabetic.**
Surname	Ensure it is present. Use the built-in function **Is Not Empty**. Ensure it is alphabetic. Use the built-in function **Is Alphabetic.**
Telephone	Ensure it is a valid telephone number. Use a custom function **Valid Phone Number**.

Table Rules

We also want to check whether **both** City and Post Code are valid. We do this by attaching an **Is Not Empty** rule for each column at table level. This is a very simple example that just ensures they are present.

Quality Thresholds

We set the quality thresholds as follows:

Object	Lower Threshold	Higher Threshold
Customers Table	85%	95%
Customer Id	100%	100%
Discount Code	100%	100%
Forename	100%	100%
Surname	100%	100%
Email	89%	100%
Telephone	30%	100%
Company Name	69%	100%
Addressline	93%	100%
Country	100%	100%

Filters

We have set a filter of "Credit Limit is Greater than 1,000,000" on the Table; "Forename is in the range of A to Gz" on the Forename Column; and "Email Starts With a" on the Matches Expression Rule on the Column Email.

Table Expressions as Measure Columns

We have created the **Fullname** Table Expression to be used as a measure column on the table. The columns will all inherit this measure column by default:
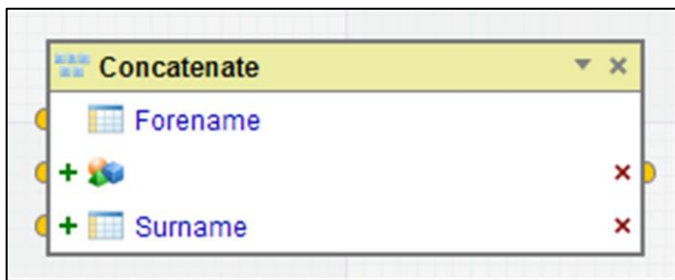
Table Validation Result

If we go into the Table's folder and click on **Rules** we can see the overall result of the Rules on the Table. The overall score is both written and shown in the dial to the side. The Pass Range and Fail Range are also shown. A drilldown of all the rows is shown underneath.

Validation Rules for Customers (Table)
Last Validated 30 July 2014 17:12:38 BST

2 Rows

Validate   New Rule...   Close

Score:	99.99% (Passed)	**Rules:**	2
Passed Measure: N/A		**Failed Measure:**	N/A
Passed Rows:	14,283 (99.99%)	**Failed Rows:**	1 (0.01%)
Considered:	14,284 (39.47%)	**Ignored:**	21,901 (60.53%)
Evaluated On:	30-Jul-2014 17:12:38	**Validated by:**	Administrator
Measure:	N/A	**Default Measure:** Default measure of One	
Pass Range:	95.0% - 100%	**Fail Range:**	0% - 85.0%
Filter:	Credit Limit > 1,000,000		

Pass

	Name	Score	Measure	Passed Measure	Failed Measure	Rule Category	Descriptio
1	City Is Not Empty	99.99%	N/A	N/A	N/A		Tests whe
2	Post Code Is Not Empty	100%	N/A	N/A	N/A		Tests whe

By double clicking on one of the rows we can see the dashboard and the **Score** drilldown for that individual result:

All Rows for Customers (Table)

36,185 Rows

Validate   New Rule...   Close

Score:	99.99% (Passed)	**Rules:**	2
Passed Measure: N/A		**Failed Measure:**	N/A
Passed Rows:	14,283 (99.99%)	**Failed Rows:**	1 (0.01%)
Considered:	14,284 (39.47%)	**Ignored:**	21,901 (60.53%)
Evaluated On:	30-Jul-2014 17:12:38	**Validated by:**	Administrator
Measure:	N/A	**Default Measure:** Default measure of One	
Pass Range:	95.0% - 100%	**Fail Range:**	0% - 85.0%
Filter:	Credit Limit > 1,000,000		

Pass

	Result	Customer Id	Discount Code	Forename	Surname	Company Name
1		A10-E100750	M	Janice	Davies	Summit Securiti
2		A10-E102508	H	Michael	Bowen	Oregon Steel Mi
3		A10-E102659	H	Irene	Reynolds	M S Capital & M
4		A10-E102982	M	Philip	Friesner	Pacific Pearl Fir
5		A10-E103408	M	Jr	Maria	V L S Securities
6		A10-E103489	H	Adelle	Madden	Phone House T
7		A10-E105402	N	Accounts	Schofield	Micah Systems,
8		A10-E105686	N	Paul	Anderson	NCCW-CASA
9	✔ Yes	A10-E105688	M	Nigel	Horrocks	Music World Ltd
10		A10-E106043	N	Reginald	Vorst	infoUSA Inc.
11		A10-E108032	N	Roger	Borgt	Hottinger Baldw
12	✔ Yes	A10-E108768	N	Greg	Guy	Pennsylvania D
13	✔ Yes	A10-E108908	M	Lisa	Colston	Mahanadi Coalf
14		A10-E109669	M	Linda	Kemp	Spridick Garme

Clicking on the information in the dashboard will produce the relevant drilldown underneath.

The individual rules can also be found in the explorer by going into the **Rules** folder:

Rules
2 Rules - 99.99%
Type here to filter items

City Exists
Score: 99.99%

Post Code Exists
Score: 100%

By going into the folder of one of the rules we are given its results:

Post Code Exists
Score: 100%
Type here to filter items

Score
100%

Passed
14,284 (100.0%)

Failed
0 (0.0%)

Considered
14,284 (39.47%)

Ignored
21,901 (60.53%)

Filter
Credit Limit > 1000000

Included
Yes

Library for Function
Validation

Validation Function
Check If Not Date/Time

Description
Tests whether a value is not v...

Id
3
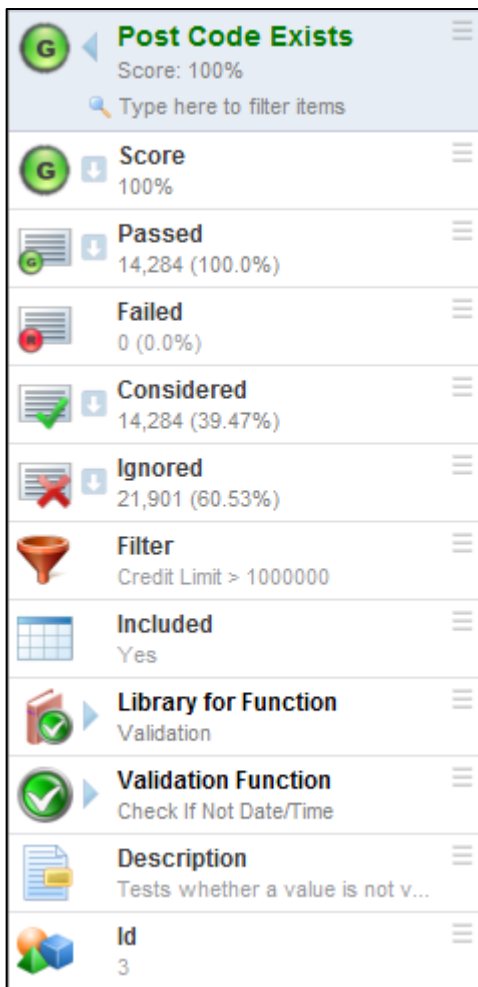
Column Quality Results

If we click on **Columns**, we can see all the overall scores for the individual columns.

	Name	Table	Uniqueness	Unique Count	Completeness	Score	Row Count	Null Count
	Columns for Customers (Table)						18 of 216 Columns	
								12 Rows
1	Customer Id	Customers	99.72%	36,083	100%	99.44%	36,185	0
2	Discount Code	Customers	0.02%	8	100%	99.99%	36,185	0
3	Forename	Customers	9.55%	3,455	100%	99.39%	36,185	0
4	Surname	Customers	29.11%	10,532	100%	99.75%	36,185	0
5	Company Name	Customers	82.11%	29,711	99.99%	99.99%	36,180	5
6	Addressline	Customers	75.33%	27,259	99.99%	99.99%	36,184	1
7	City	Customers	6.33%	2,290	99.99%		36,180	5
8	Country	Customers	0.02%	6	100%	100%	36,185	0
9	Post Code	Customers	82.57%	29,879	100%		36,185	0
10	Telephone	Customers	99.24%	35,910	100%	98.33%	36,185	0
11	Email	Customers	99.71%	36,081	100%	94.51%	36,185	0
12	Credit Limit	Customers	1.4%	507	100%		36,185	0

If we right-click on the Column and select **Dashboard** we can see the overall result of the Rules on the Column. The overall score is both written and shown in the dial to the side. The Pass Range and Fail Range are also shown. A drilldown of the rules appears underneath.

Quality Rules for Customers.Surname (Column)
Last Validated 30 July 2014 17:12:38 BST
25 of 28 Columns
2 Rows

Validate   New Rule...   Close

Score:	99.75% (Failed)
Passed Measure:	N/A
Passed Rows:	36,095 (99.75%)
Considered:	36,185 (100%)
Evaluated On:	30-Jul-2014 17:12:38
Measure:	N/A
Pass Range:	100%
Filter:	N/A

Rules:	2
Failed Measure:	N/A
Failed Rows:	90 (0.25%)
Ignored:	0 (0%)
Validated by:	Administrator
Default Measure:	Default measure of One
Fail Range:	0% - 100.0%

	Name	Score	Measure	Passed Measure	Failed Measure	Rule Category	Description
1	Is Alphabetic	99.75%	N/A	N/A	N/A		Tests whe
2	Is Not Empty	100%	N/A	N/A	N/A		Tests whe

Clicking on the information on the dashboard will produce the relevant drilldown underneath.

The individual rules can also be found in the explorer by going into the **Rules** folder:
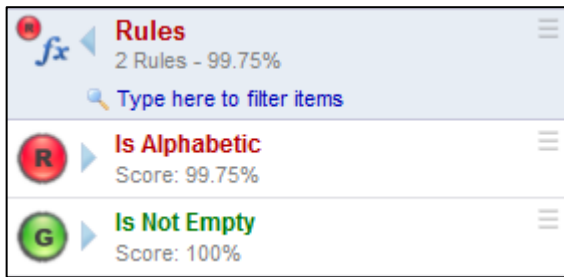
Viewing Overall Score and Passed, Failed, Considered and Ignored Rows

From the Table, a Column, a Rule Categories or a Rule, we can view the overall passed, failed, considered and ignored rows at that level. We can also view all the rows with their scores. For example go onto a Table's dashboard and you are given the **Score** drilldown by selecting **Score**:

	Result	Customer Id	Discount Code	Forename	Surname
1		A10-E100750	M	Janice	Davies
2		A10-E102508	H	Michael	Bowen
3		A10-E102659	H	Irene	Reynolds
4		A10-E102982	M	Philip	Friesner
5		A10-E103408	M	Jr	Maria
6		A10-E103489	H	Adelle	Madden
7		A10-E105402	N	Accounts	Schofield
8		A10-E105686	N	Paul	Anderson
9	✔ Yes	A10-E105688	M	Nigel	Horrocks
10		A10-E106043	N	Reginald	Vorst
11		A10-E108032	N	Roger	Borgt
12	✔ Yes	A10-E108768	N	Greg	Guy

From the dashboard, select **Passed Rows** to see all the rows that passed:

	Result	Customer Id	Discount Code	Forename	Surname
1	✔ Yes	A10-E105688	M	Nigel	Horrocks
2	✔ Yes	A10-E108768	N	Greg	Guy
3	✔ Yes	A10-E108908	M	Lisa	Colston
4	✔ Yes	A10-E111093	N	Michelle	Bertovich
5	✔ Yes	A10-E115541	H	Philip	Saddler
6	✔ Yes	A10-E117853	N	Jane	Pritchard
7	✔ Yes	A10-E121594	M	Jerry	Williams
8	✔ Yes	A10-E123286	L	Anthony	Joseph
9	✔ Yes	A10-E124690	H	Antonius	Ash
10	✔ Yes	A10-E125466	N	Clare	Faulkingham
11	✔ Yes	A10-E126564	N	Alan	Chovaz
12	✔ Yes	A10-E129546	N	James	Bhogal
13	✔ Yes	A10-E129689	H	Lesley	Sweeney
14	✔ Yes	A10-E130309	M	Raymond	Jones

From the dashboard, select **Failed Rows** to see all the rows that failed:

	Result	Customer Id	Discount Code	Forename	Surname
1	✖ No	D60-E149816	N	Brian	Potts

From the dashboard, select **Considered Rows** to see all the rows that passed the filter:

	Result	Customer Id	Discount Code	Forename	Surname
1	✔ Yes	A10-E105688	M	Nigel	Horrocks
2	✔ Yes	A10-E108768	N	Greg	Guy
3	✔ Yes	A10-E108908	M	Lisa	Colston
4	✔ Yes	A10-E111093	N	Michelle	Bertovich
5	✔ Yes	A10-E115541	H	Philip	Saddler
6	✔ Yes	A10-E117853	N	Jane	Pritchard
7	✔ Yes	A10-E121594	M	Jerry	Williams
8	✔ Yes	A10-E123286	L	Anthony	Joseph
9	✔ Yes	A10-E124690	H	Antonius	Ash
10	✔ Yes	A10-E125466	N	Clare	Faulkingham
11	✔ Yes	A10-E126564	N	Alan	Chovaz
12	✔ Yes	A10-E129546	N	James	Bhogal
13	✔ Yes	A10-E129689	H	Lesley	Sweeney
14	✔ Yes	A10-E130309	M	Raymond	Jones
15	✔ Yes	A10-E134505	N	Joseph	Dhadwal
16	✔ Yes	A10-E135749	M	Mike	Perkins

From the dashboard select **Ignored Rows** to see all the rows that did not pass the filter:

	Result	Customer Id	Discount Code	Forename	Surname
1		A10-E100750	M	Janice	Davies
2		A10-E102508	H	Michael	Bowen
3		A10-E102659	H	Irene	Reynolds
4		A10-E102982	M	Philip	Friesner
5		A10-E103408	M	Jr	Maria
6		A10-E103489	H	Adelle	Madden
7		A10-E105402	N	Accounts	Schofield
8		A10-E105686	N	Paul	Anderson
9		A10-E106043	N	Reginald	Vorst
10		A10-E108032	N	Roger	Borgt
11		A10-E109669	M	Linda	Kemp
12		A10-E111267	N	Tina	Fletcher
13		A10-E112904	N	Chagger	Wharrad
14		A10-E113663	L	Michael	Haq

These produce a standard rows drilldown with an extra **Result** column which is the overall result at that level. The interesting aspect is that this contains a dynamically built expression that is the combination of all the rules at that level, a composite validation rule. Viewing this expression provides a complete end-to-end definition of what causes a row to pass or fail, which is extremely useful for auditing the process.

Adding Versions

You can load additional versions of the table. The validation rules are automatically attached and evaluated and we can see how the quality changes over time.

Table Changes over Time

Go into a table's folder, right-click on **Versions** and select **Changes over Time > List** to see what aspects have changed. Note how the validation score has changed over the three versions.

Table Comparison over time

Value	Customers v1	Customers v2	Customers v3
Row Count	36,185	▼ 26,040	▲ 36,185
Column Count	12	12	12
Columns Added	0	0	0
Columns Moved	0	0	0
Columns Removed	0	0	0
Unique Values	210,874	▼ 157,369	▲ 210,874
Unique Formats	19,135	▼ 14,843	▲ 19,135
Alphanumeric Values	207,671	▼ 130,806	▲ 207,671
Integer Values	3,203	▲ 26,563	▼ 3,203
Decimal Values	0	0	0
Number Values	3,203	▲ 26,563	▼ 3,203
Money Values	0	0	0
Date Values	0	0	0
Alphanumeric Formats	19,125	▼ 14,832	▲ 19,125
Integer Formats	10	▲ 11	▼ 10
Decimal Formats	0	0	0
Number Formats	10	▲ 11	▼ 10
Money Formats	0	0	0
Score	0	▲ 100	▼ 99.99
Rules	0	▲ 2	2
Rows Passed	0	▲ 10,326	▲ 14,283
Rows Failed	0	0	▲ 1
Rows Considered	0	▲ 10,326	▲ 14,284
Rows Ignored	0	▲ 15,714	▲ 21,901

This can also be viewed graphically, viewing one aspect at a time by right-clicking on **Versions** and selecting **Changes over Time > Graph**.
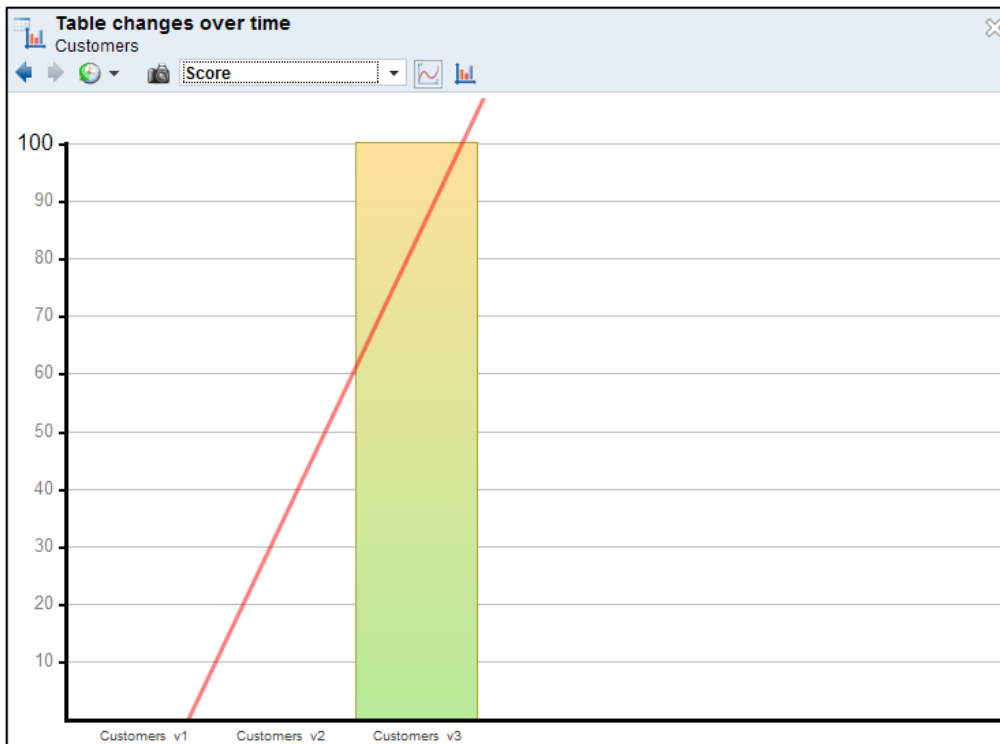
Column Changes over Time

Right-click on any column and select **Changes over Time > List** to see what aspects have changed. Note how the Quality Scores have changed over the three versions.

Changes to Forename over time

Value	Forename v1 Customers	Forename v2 Customers	Forename v3 Customers
Value Count	36,185	▼ 26,040	▲ 36,185
Distribution	9.55	▼ 6.92	▲ 9.55
Unique Value Count	3,455	▼ 1,802	▲ 3,455
Unique Format Count	53	▼ 25	▲ 53
Null Count	0	0	0
Blank Count	0	0	0
Zero Count	0	0	0
Negative Count	0	0	0
Shortest Length	1	▲ 2	▼ 1
Longest Length	16	▼ 15	▲ 16
Precision	2	2	2
Scale	0	0	0
Sum	41	41	41
Average	0	▲ 0	▼ 0
Average Length	5	5	5
Average Frequency	10.47	▲ 14.45	▼ 10.47
Alphanumeric Value Count	36,184	▼ 26,039	▲ 36,184
Alphanumeric Unique Value...	3,454	▼ 1,801	▲ 3,454
Alphanumeric Distribution	9.55	▼ 6.92	▲ 9.55
Integer Distribution	100	100	100
Decimal Distribution	0	0	0
Money Distribution	0	0	0
Date Distribution	0	0	0
Alphanumeric Unique Form...	52	▼ 24	▲ 52
Alphanumeric Shortest Len...	1	▲ 2	▼ 1
Alphanumeric Longest Leng...	16	▼ 15	▲ 16

This can also be viewed graphically, viewing one aspect at a time, by right-clicking on the column and selecting **Changes over Time > Graph**

Column changes over time — Customers.Forename
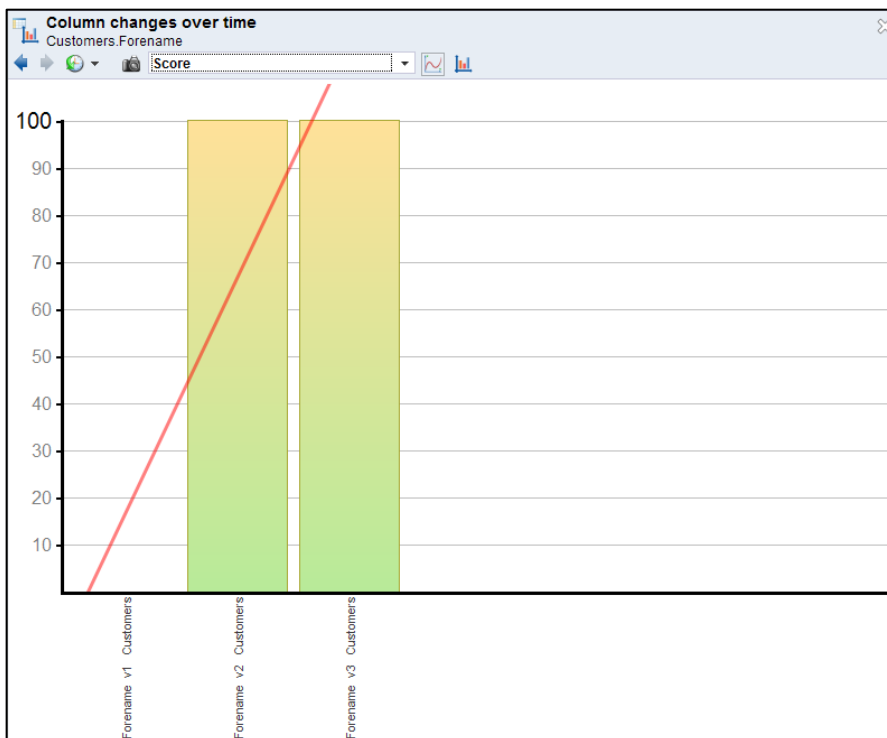
Viewing all column version traffic lights

Right-click on the **Versions** folder of the table and select **Versions.** Highlight the tables with columns you want to look at, right click and select **Columns.**

	Name	Version	Uniqueness	Unique Count	Completeness	Score	Row Count
	Columns for All Versions of Customers (Table)				18 of 216 Columns		36 Rows
1	Addressline	3	86.59%	15,284	99.99%	99.99%	17,651
2	Addressline	2	76.16%	26,487	99.99%	99.99%	34,775
3	Addressline	1	75.33%	27,259	99.99%	99.99%	36,184
4	City	3	8.91%	1,573	99.98%		17,649
5	City	2	6.5%	2,260	99.99%		34,771
6	City	1	6.33%	2,290	99.99%		36,180
7	Company Name	3	91.08%	16,077	99.99%	99.99%	17,650
8	Company Name	2	82.67%	28,750	99.99%	99.99%	34,771
9	Company Name	1	82.11%	29,711	99.99%	99.99%	36,180
10	Country	3	0.03%	5	100%	100%	17,652
11	Country	2	0.02%	6	100%	100%	34,776
12	Country	1	0.02%	6	100%	100%	36,185
13	Credit Limit	3	2.87%	507	100%		17,652
14	Credit Limit	2	1.46%	507	100%		34,776
15	Credit Limit	1	1.4%	507	100%		36,185
16	Customer Id	3	99.71%	17,601	100%	99.42%	17,652
17	Customer Id	2	99.71%	34,675	100%	99.42%	34,776
18	Customer Id	1	99.72%	36,083	100%	99.44%	36,185
19	Discount Code	3	0.01%	1	100%	100%	17,652
20	Discount Code	2	0.02%	7	100%	99.99%	34,776
21	Discount Code	1	0.02%	8	100%	99.99%	36,185
22	Email	3	99.71%	17,600	100%	94.29%	17,652
23	Email	2	99.7%	34,673	100%	94.49%	34,776
24	Email	1	99.71%	36,081	100%	94.51%	36,185
25	Forename	3	13.47%	2,377	100%	99.32%	17,652
26	Forename	2	9.8%	3,407	100%	99.39%	34,776
27	Forename	1	9.55%	3,455	100%	99.39%	36,185
28	Post Code	3	89.59%	15,814	100%		17,652

Controlling Access

A User is someone with a defined role who uses the Experian Pandora application and is managed independently of any host platform security, in that they are separate from Operating System Users. This ensures that the security for data held within Experian Pandora is self-managed and self-contained and transportable with the database. All user activity is co-ordinated by Experian Pandora in real-time.

The **Administrator** is a special user that always exists and has the ability to perform any action and have total access to all data and metadata. There is only one Administrator, but it is possible to give additional users administrator rights via their User Role.

Information about Users can be accessed from the explorer from **Settings > Security > Users**. Expanding a User in the explorer will display information about the user as well as detailing whether that user is currently logged in or not. From the right-click menu for a User there is a variety of options covering all aspects of interaction with the user, as well as facilities for editing the definition and preferences of a User.

Permissions and Ownership

The Experian Pandora database provides a comprehensive security model for database objects, such as Tables and Columns, with multiple access levels and security rights.

There are 3 basic levels of security:

Owner

Every object in the database is owned by a User, generally the User that created it unless ownership has been changed. Every object can have specific rights for the owning User.

Team

Every object in the database is owned by a User who is a member of a Team. Each object can have specific privileges set for other Users in the same Team as the owning User.

Public

This defines the privileges for all Users who are **not** the owning User or in the Team of the owning User.

For special access cases, all database objects also support **Access Control Lists**. Specific permissions are given to a specific user regardless of whether they are in the same User Group or not. The available security privileges are:

Privilege	Description
Read	Allows a user the ability to read data in the object. For actual data in Columns and Tables, the lack of this privilege will result in all data access being Obfuscated.
Modify	Allows a user the ability to modify an object, for example changing the name of a Table.
Delete	Allows a user the ability to delete an object, apart from loaded or Saved Tables which can only be deleted by an Administrator.
Export	Allows a user the right to export data from drilldowns, notes or reports to external targets such as export files.
Annotate	Allows a user the right to annotate an object, provided they have the Annotate capability.

The Permissions dialog can be accessed from the *Edit* or *Create* dialogs for any database object that supports changes to its permissions, which is most of them. This is available from the Security button on dialogs.

Obfuscation

Password-protected user access goes a long way to protecting sensitive data, however a common problem in enterprise Data Analysis is that the people who need to analyse the data for quality purposes are often not actually allowed to see some or all of the values! Experian Pandora provides the unique capability to automatically *obfuscate* the values related to Database objects that the User is not allowed to see.

Obfuscation Example

An EMPLOYEE file needs to be analysed, but the analyst is not allowed to see sensitive data such as the SALARY Column. Obfuscation scrambles the value making it meaningless on its own but retains information such as statistics and format patterns. Drilling down from values to their usages or to rows in tables continues to function. Of course with obfuscated values the phonetic patterns are rendered useless as these cannot sensibly be encrypted and still have meaning. Also, analysts cannot drilldown onto Dependency detail for discovered Dependencies if they do not have full read access. The reason for this is that drilling down onto a discovered Dependency turns it to a *tested* state, which implies the discrepancies have been viewed and understood - something which is not possible if the data was obfuscated!

Experian Pandora maintains obfuscation at every level, even in the Global table of values (list of values across all tables and columns). The encryption is dynamic to ensure that the obfuscation cannot be easily decrypted by the Analyst.

This unique capability facilitates cross-enterprise analysis of sensitive data, a typical activity on Compliance initiatives and Master Data Management projects.

The encoding for Obfuscation changes continuously ensuring that a user cannot reverse engineer the true value. Obfuscation is also carried through transformation rules, exports and anywhere that data can either be visualised or its meaning inferred via calculation.

How to enable obfuscation

Simply remove the permission to read from any Column of a loaded Table in Experian Pandora, by right-clicking on a column in the Explorer, selecting **Edit** then choosing the permissions tab. The user who does not have read permission will see everything related to that Column as obfuscated. If you are loading a table with sensitive data we suggest you load it as unreadable and then protect individual columns before giving read permission on the remaining data.

Advanced Topics

In this section we will look at some of the more advanced functionality in Experian Pandora

Multi-Dimensional Grouping

Experian Pandora also provides 3 additional multi-dimensional groupings where more than one level of aggregation can be viewed at one time. This can be changed by setting the grouping type, which can be found on the Group panel of the Drilldown Manager.
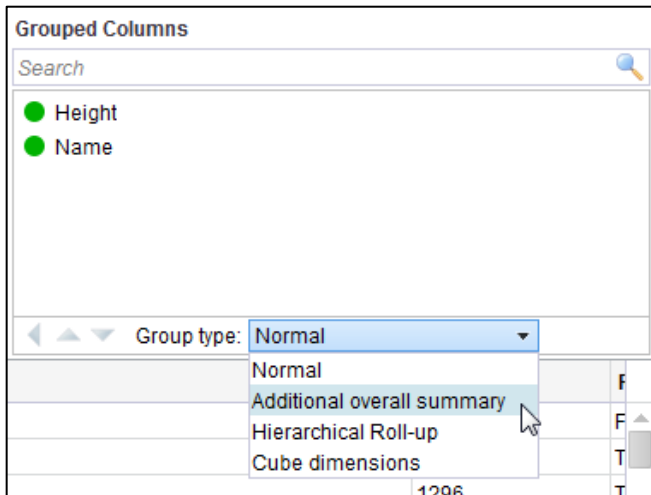
Grouping type	Description	Requires multiple grouping columns
Normal	Normal grouping operation	No
Additional Overall Summary	Creates an additional total line at the end of the result set which aggregates across all records	No
Hierarchical Rollup	Creates additional summary totals at each level of a multi-column grouping	Yes
Cube Dimensions	Creates every possible permutation of grouping summary totals for a multi-column grouping.	Yes

Additional Overall Summary

This is simply a total line added to the end of the group drilldown. Using the previous Mountains example, if we wanted to know the total number of mountains and the highest mountain across all the countries, we simply select this option.

We now have an additional total line in blue.

You can now see the total symbol on the total line. Note the output is more report like when these different grouping types are used, and that you can no longer sort the output as it is kept in a specific order to be able to show totals.

Hierarchical Rollup

This multi-dimensional grouping type is widely used to provide multiple levels of aggregation in a single drilldown. Using our Mountains example, we have been grouping by **Country**. If we also grouped by **Range** then we would get a finer grain level of information about the number of mountains and highest mountain by Range within Country. However, we lose the Country level aggregation results in the process. By turning on Hierarchical Rollup, Experian Pandora will treat the groupings as a hierarchical ordering, so will group by Country and Range together, but also by Country on its own, and order the output to provide additional sub totals for Country, nestled within the Country/Range combined results for that Country, and finally an overall total line.

There is a function called **Group Set Name** which will appear as a Key Column while a grouping is active. This function will mark the end of each group and will insert a row which will give the totals in each column for that group set.

The following example groups Mountains by **Country** and then **Range**, with Hierarchical Rollup set as the type of grouping. The image below shows the tail end of the drilldown. We now have 100 result rows, because grouping by Country and Range results in a finer grain level of detail. The tail end of the results show:

- The results for the ranges in Scotland
- The summary line for Scotland
- The results for the ranges in Wales
- The summary line for Wales
- The overall total for all of the mountains

	Country	Number of Mountains #	Highest Mountain V
84	Scotland	23	871
85	Scotland	13	706
86	**Scotland**	**1703**	**1344**
87	Wales	15	752
88	Wales	10	564
89	Wales	42	905
90	Wales	27	854
91	Wales	9	481
92	Wales	17	893
93	Wales	43	886
94	Wales	10	578
95	Wales	13	600
96	Wales	67	1085
97	Wales	15	645
98	Wales	17	660
99	**Wales**	**285**	**1085**
100		2367	1344

Σ Rows for Mountains (Table)   3 of 8 Columns

100 Groups
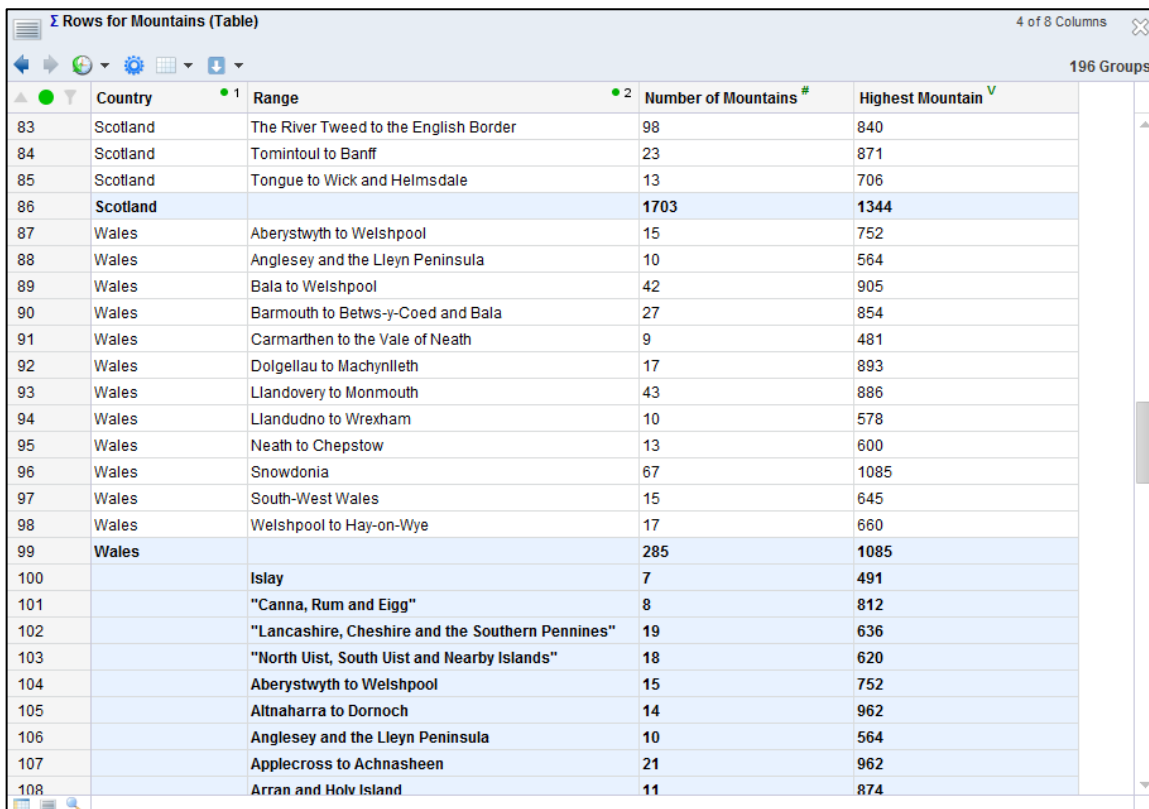
Cube Dimensions

This is very similar to Hierarchical Rollup except that Experian Pandora does not generate the additional groupings hierarchically, but instead generates every permutation of grouping.

To illustrate the differences for the Mountains example, grouping by **Country** and **Range:**

Grouping type	Results Generated
Normal	**Country, Range** grouping
Additional Overall Summary	**Country, Range** grouping Overall total line
Hierarchical Rollup	**Country** grouping **Country, Range** grouping Overall total line
Cube Dimensions	**Country** grouping **Country, Range** grouping **Range** grouping Overall total line

Simply by changing the grouping type to **Cube Dimensions** with the Mountains example, the new result has an additional 96 results in it, which are the additional group results where **Range** on its own has also been grouped. The screenshot below shows how these additional groupings appear after the Hierarchical results, but before the overall total line.

Settings

System settings and user preferences enable configuration and personalisation of how Experian Pandora operates. The Administrator has access to all the setting options while Users have access to those related to their User Capabilities.

My Preferences

Users can set the options they prefer, e.g. single drilldown mode, show tooltips, expression editor function to RHS etc.

Security

The Administrator or a user with Administration capabilities has access to this area.

Roles

The capabilities of a User within Experian Pandora are dictated by their defined User Role. Experian Pandora ships with several pre-defined roles which can be modified or added to.

User Capabilities

Capability	Description
Admin	Allows users to have similar privileges to the Administrator
Alert Management	Allows users to unsubscribe to receiving Alerts to particular events.
Analyse	Allows users to run analysis tasks such as Dependency Analysis
Annotate	Allows users to annotate with Notes
Create	Allows users to create new metadata objects, such as Projects
Delete	Allows users to delete metadata objects
Export	Allows users to export data and metadata
Glossary	Allows users access to the Business Glossary
Glossary Admin	Allows users to manage objects that reside in the Glossary folder
Library Admin	Allows users to manage objects that reside in the Libraries folder
Load	Allows users to load data into the database
Reporting	Allows users to consume reports via ODBC/JDBC and REST
SQL	Allows users to drilldown using ad-hoc SQL queries
Script	Allows users to execute scripts through alerts

Users

Create and delete users, on the Details tab set their capabilities and passwords and on the Visibility tab set more detail for which elements individuals can see within Experian Pandora.

Teams

A Team is simply a collection of Users. Within Experian Pandora are 3 levels of permission access – User, Team and World. Grouping Users into Teams allows easy subdivision of permission access to shared objects for members of the same team. All Users in the Team will be subject to the team level privileges for any database object owned by any User within that Team.

Audit Types

Set the auditable events that you want Experian Pandora to record in the Audits.

Alerts

Create Rosters and Templates for Alerts. Administrators and users with Administration rights can view and edit all the Default setting for Alerts. Please refer to the section on Alerts for more information.

Notes

Create, delete, and amend the Notes Categories, Impacts and Status.

Rule Categories

Create, delete, edit Rule Categories which are used to categorise the use of Validation Functions to group them together under a common definition.

Export Styles

Set the Export styles to be used when exporting files

Chart Appearances

Set colour schemes for use with Charts. See the section on Charts for more information.

Server Settings

The Server Settings are accessible from within Experian Pandora and are divided into groups of settings. All Server Settings are available in one drilldown, to view this, click on **All Server Settings** from the **Settings > Server Settings** item in the Explorer.

All settings have a default which can be accessed from the right-click menu on the individual item in the Explorer or the Drilldown. The defaults are the standard settings used when you first install Experian Pandora. Only a User who has the Admin capability enabled will be able to edit the server settings. These cover a wide array of aspects from particular behaviours and policies to file locations, debug settings and much more.

System Administration

Scheduled Jobs

Many activities in Experian Pandora result in a job being scheduled. You can view scheduled jobs by clicking on the **Jobs** icon in the Quick Bar. There are a number of other options to manage jobs from the right-click menu, most of which will operate on multiple selections of jobs as well as single selections. Each job has its own log which can be accessed by double-clicking on the job within the Scheduled Jobs Drilldown.

Job States

State	Description
Creating	The job is in the process of being created
Initialising	The job is starting up
Running	The job is running and active
Waiting	The job is running but waiting for some other activity to complete or some state to be reached
Finishing	The job is in the process of completing
Tidying	The job is tidying up which occurs both in cases of success and failure
Aborted	The jobs was aborted usually as a result of a failure outside the job or in an a sub-task
Failed	The job failed in an easily determinable way
Completed	The job finished successfully
Stopped	The job was stopped typically by the user
Queued	The job is queued awaiting for start conditions to be met or resources to be made available
Logging	The job is in the process of logging
Queue Full	The job uses a data queue which it is attempting to write to, and it has had to suspend as the queue is currently full

Processes

Experian Pandora provides a real time view of current process activity on the server. It shows a breakdown of current activity particularly during long running scheduled tasks such as the loading of Tables. The display automatically refreshes as events occur. This can be invoked by right-clicking on **Home** from the explorer, and selecting **System > Processes** from the right-click menu or by navigating to the **System** folder from the Explorer and clicking on **Processes**.

Sessions

Experian Pandora provides a drilldown to display the active sessions for all users. This includes all server-side sub sessions and every session corresponding to the login for a user (Primary Session) or a drilldown (Client Session, non-primary). These can be individually terminated should a user suffer from any issues, or monitored for activity, including displaying live progress of those individual drilldowns. Options on this view include the ability to drilldown to a replica of the actual activity of other users by double clicking on a session with an associated drilldown. This drilldown can be invoked by navigating to the **System** folder from the Explorer and clicking on **Sessions**.

Services

Services are critical background processes, such as the Scheduler, which form the running server. These can be viewed by navigating to the **System** folder from the Explorer and clicking on **Services.**

Watchdog Tasks

Tables that have source or trigger event monitoring will result in a new Watchdog Task being created. The Watchdog is a Experian Pandora system task that handles the entire event monitoring management for file arrival and other events that result in a reload of data. To view the current watchdog tasks, go into the System folder in the Explorer and select **Watchdog Tasks**. The Watchdog Tasks drilldown appears. This will display all currently live tasks along with information about their status and activity.

Log Messages

Virtually every activity in Experian Pandora will result in messages being written to the full system log. This can be viewed by right-clicking on **Home** in the explorer and selecting **System > Get Log Files** to see the Activity Log Messages drilldown. The Activity Log can be viewed by navigating to the **System** folder in the Explorer and clicking on **Log Messages**. This contains the most relevant messages as a Drilldown. Once this Drilldown is displayed, you can use the Drilldowns sub-menu on the Drilldown window activity menu to flip between different log message views such as Errors, Debug etc.

System information

Right-click on **Home** and select **System > System Information**. This produces a drilldown of system related information. This varies from platform to platform. For example, on Linux systems this produces much more information than Windows systems.

Also particularly useful is the **Free Memory** statistic. Experian Pandora is configured to use a certain amount of system memory at start time, and then manages that amount by itself. This figure goes up and down all the time as Experian Pandora automatically releases memory. It is useful for checking how much memory has been committed from the operating system to Experian Pandora, and how much the server is currently using.

System Information can also be found in the System section of the explorer.

Storage Information

Experian Pandora has a number of logically defined directories for data storage which form the Experian Pandora database. These are continuously monitored by the Housekeeper task. Right-clicking on **Home** and selecting **System > Storage Information** to view the Storage Information drilldown which produces information about the directories, including their purge intervals, and free available space. Any entries marked in red are currently running below the configured minimum space threshold percentage (10% by default).

Storage Information can also be found in the System section of the explorer.

Repository Objects

An entire set of objects managed in the Experian Pandora repository is accessible from the Repository Objects drilldown. Go into the **System** section of the explorer and click on **Repository Objects**

You can right-click on any object in the drilldown to get its standard right-click menu with all the available options. This drilldown allows an enterprise-wide metadata search using all of the standard filter techniques of the drilldown.

Note: the Repository Objects drilldown is only available to administrators.

Audits

Audit Overview

Experian Pandora supports a number of categories for auditing events. These may be enabled or disabled by an Administrator. By default no audit events are logged and will need to be defined by the Administrator. Only an Administrator may change the Auditable event types. The one class of audit events that is always enabled, (and cannot be disabled) is auditing changes to the Audit Types, i.e. whenever an Audit Type is enabled or disabled. Enabling all audit types may result in substantial disk storage requirements over time.
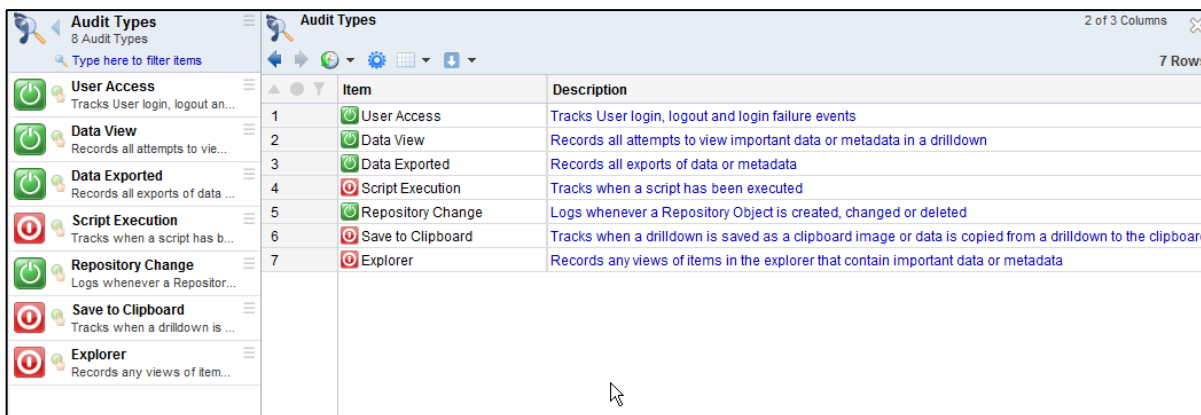
Audit Types

- User Access – User and Script Logins, Logouts and Login Failures
- Data View – Viewing Tables and Drilldowns
- Data Exported – Data exports to an external file or Table
- Script Execution – Execution of JavaScript and SQL scripts
- Repository Change – Changes to a repository object (Created, Modified, Deleted)
- Save To Clipboard – Saving data views or data selections to the system clipboard
- Explorer – Views of Explorer nodes that expose data

Enabling/Disabling Audits

Navigate to **Home > Settings > Security > Audit Types**
Clicking on an Audit Type in the Explorer will toggle its state between enabled and disabled, and double-clicking on an Audit Type in the drilldown will toggle its state.

Viewing Audits

Navigate to **Home > System > Audits**

Audit log entries are displayed with the most recent at the top of the drilldown.

At row 15 is an example of a modification of a Repository Object (The Schema named Test). The number of changes in this instance is 1. In order to view the changes you must right-click on the row and select the option "View Changes".

Administration

Deleting tables

Deleting whole tables from Experian Pandora is an irreversible process that could take a similar time to loading the data in the first place. Any loaded table, including tables created via the **Save As > Table** drilldown save option, can be deleted by right-clicking on the table and selecting **Advanced > Delete**.

It is not possible to delete a table if it has references, such as:

- Saved Tables
- Saved Drilldowns
- Saved Views
- References from table measures
- References from Table Lookups
- References from user-defined functions

For example, if we create a View from a loaded table, then try to delete the loaded table, the following error dialog will be shown:
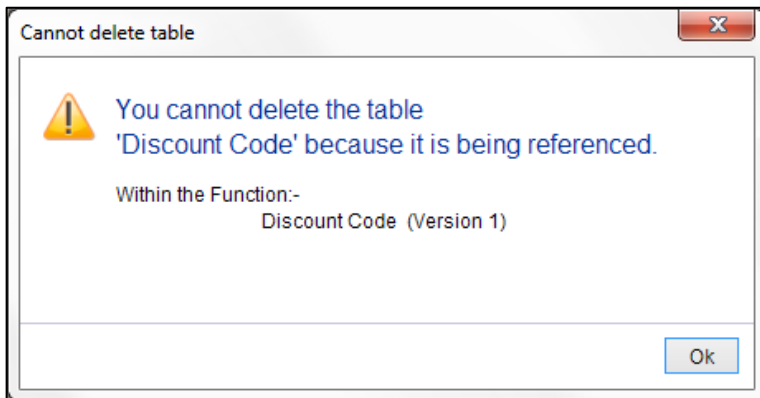

This check can only be completed at point of delete as it is not a rapid check to perform. Therefore even if the delete option is enabled for a table, it does not mean that you can actually delete it!

Deletion is a restricted activity. It can also be configured to not happen during the working day. When a table has been requested for deletion, a red cross will appear on it.

Remotely managing the server

Experian Pandora provides facilities to remotely shutdown the server.
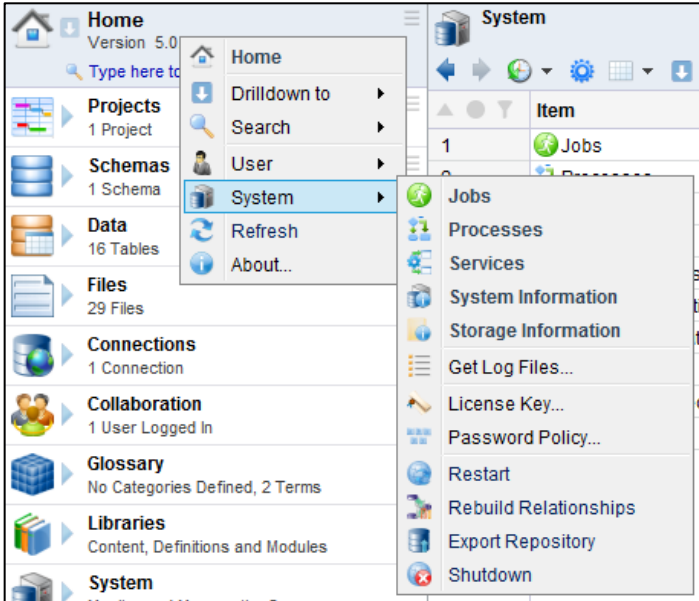
Refreshing metadata and data

Selecting the **Refresh** item forces a manual refresh of information with the server, which should be entirely unnecessary as the client manages the synchronisation automatically. This is a tool for the impatient where the administrator has set the automatic refresh polling to be relatively delayed due to poor network bandwidth for example. You should almost never need to use this. You can also press the **F5** key to achieve the same result.

License management

Selecting **License Key** presents the license entry window for entering new license keys which you may receive from time-to-time. You can also update your license by simply dragging the license file (with a **.lic** extension) that was sent to you by your support representative onto the Experian Pandora application window. You can copy and paste the license key by using CTRL+V, or simply type the new key in. When you click Apply, Experian Pandora will validate the license key. Should your license key expire or become invalid, you will be presented with this option when you log into Experian Pandora. Until you enter a valid key you will not be able to perform any actions in Experian Pandora.

Password Policy

Here you can set a Global Password Policy to force users to change their passwords on a regular basis. You can specify the type of password you require, minimum length, minimum upper-case letters, minimum digits, expiry date and whether you require it to be changed at first login. When the user is changing their password a message will appear advising them as to the password requirements.

Restart

This will do a full restart of the server, but will not close the client window down to do it.

Rebuild Relationships

To rediscover relationships you need to do a soft restart. Selecting **Rebuild Relationships** and then clicking on Yes in the dialog box which asks "Are you sure you want to restart the server and rebuild all discovered relationships". Experian Pandora will show it is Reconnecting. This can take some time to complete.

Shutdown menu item

Select the **Shutdown** item to instruct Experian Pandora to shut down the server. Any clients connected will automatically close. This option can only be invoked by the Administrator. The server restart may take longer than the automatic retry period for the client in which case the user can manually retry their connection. In either event they will not lose their current session state of open drilldown windows.

Secure communications

The server ships with a built-in self-signed SSL certificate that is used by the Web Server and server-client communications.

The Server Certificate, which can be found in **Home > Settings > Server Settings > Communications**, can be used to configure an alternative certificate. To set up an alternative SSL certificate, **right click > Edit** on the setting:
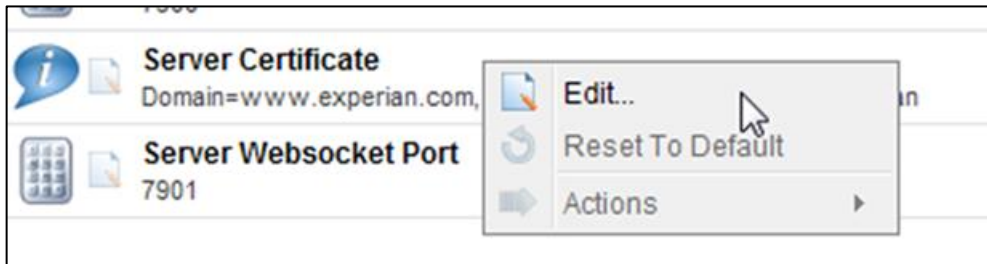
A new dialog will appear. To load a new key and certificate, you will need to put in a Key file and a Certificate file. If a passphrase is required then you need to enter that as well.
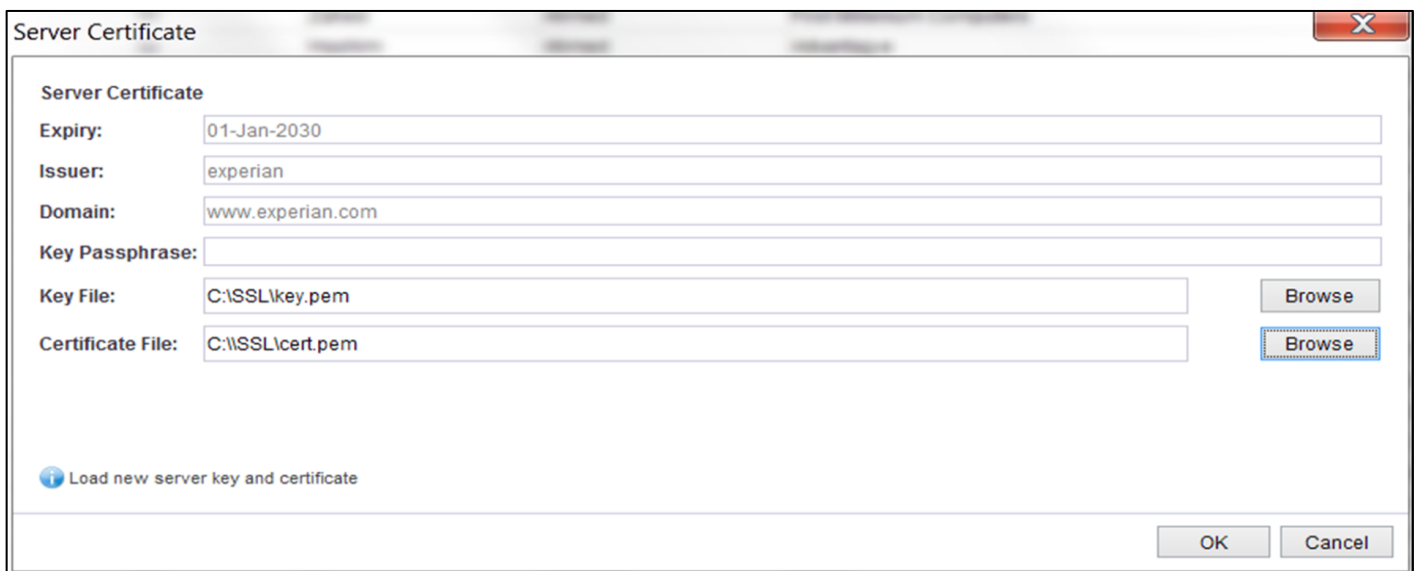
The new certificate will be sent to the server when "OK" is pressed. The server must be restarted so that the new certificate is used.

If you wish to go back to the default Server Certificate settings, then **right click > Reset To Default** on the **Server Certificate** setting. You will need to shut down and restart the server after doing so.

Private Key and Certificate files come in a variety of formats. The server handles the following common formats:

- PKCS12
- PKCS8 (PEM and DER)
- PKCS 7
- PFX (Microsoft)

Login and password management

Experian Pandora allows administrators to select the authentication mechanism used at login time.

Internal authentication

This is the default mechanism (known as INTERNAL). The password is managed internally and is subject to history checks. Note that the default Administrator username can only use the INTERNAL mechanism. This is to protect the account from lockouts should the external mechanisms become unavailable.

External authentication

Experian Pandora currently provides access to two external authentication mechanisms:
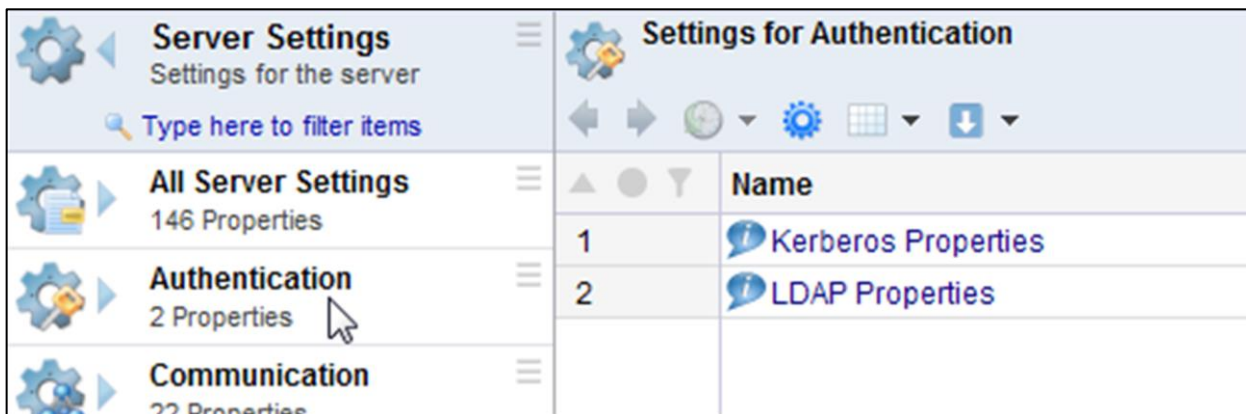
- LDAP
- Kerberos

When you use External Authentication the password is never managed by Experian Pandora and you cannot change the password from within Experian Pandora. Since the authentication is performed externally the length of time it takes to perform the authentication will depend on your network usage etc. The timeout for a response from the authentication server is set at 10 seconds. If your users experience login failures you should check the server log file for the reason.

When using an External mechanism, the Administrator should create an Experian Pandora account in the usual way, but additionally you must also specify the external username.

LDAP servers will always use a Distinguished Name (DN) to identify a user, however Microsoft's Active Directory will also allow you to specify a Kerberos-like username@domainname for simplicity.

Before you can use an external mechanism it must first be enabled in the server settings panel. If you change these properties you must perform a full shutdown and restart.   Performing a **Home** > **System** > **Restart** will <u>not</u> enforce the changes. The Authentication modules are loaded at server start-up and any communication failures with the authentication servers will be reported in the server log file.
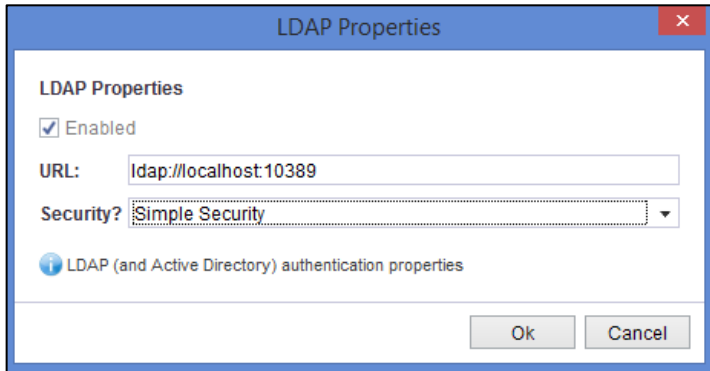

LDAP

To use LDAP password authentication you must first enable it in **Home > Settings > Server Settings > Authentication > LDAP properties** The URL scheme may be:

- ldap - you must specify the port associated with an insecure transport. Note that the connection will however still be secured with SSL.
- ldaps – you must specify the SSL port associated with the secure (SSL) protocol

You should obtain the LDAP server's URL from your LDAP server administrator. Similarly the Administrator will provide you with the appropriate Distinguished Name for your users.

The security setting must match the security used for users' passwords on the LDAP server. The LDAP server administrator will be able to tell you which one is appropriate.


Kerberos

To use Kerberos password authentication you must first enable it in **Home > Settings > Server Settings > Authentication > Kerberos properties**

The URL refers to the url of the Kerberos Key Distribution Centre (KDC) server.
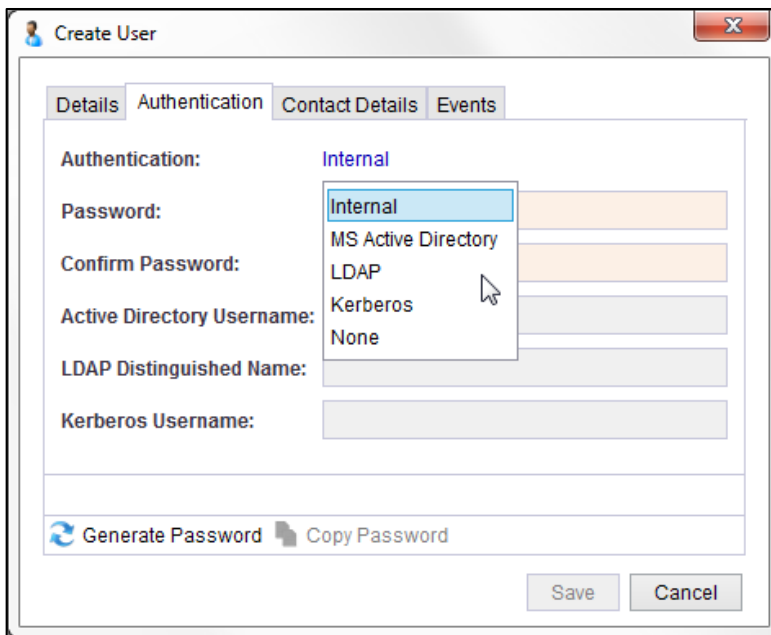The Realm is the Kerberos realm.
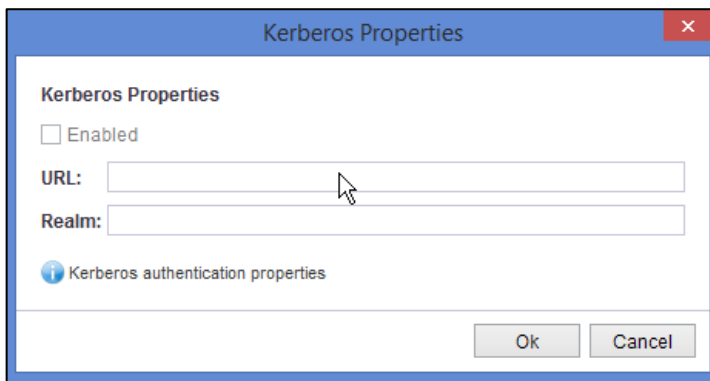
Creating a new user

Select the appropriate Authentication mechanism required when creating a new user by going to **Settings > Security**, right clicking on **Users** and selecting **New.**
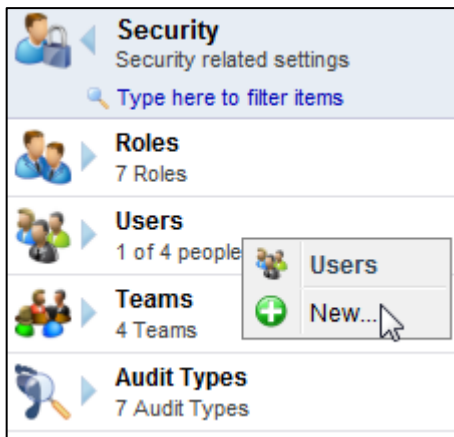
On the **Authentication** tab you can select the Authentication and provide the required information.

Appendix

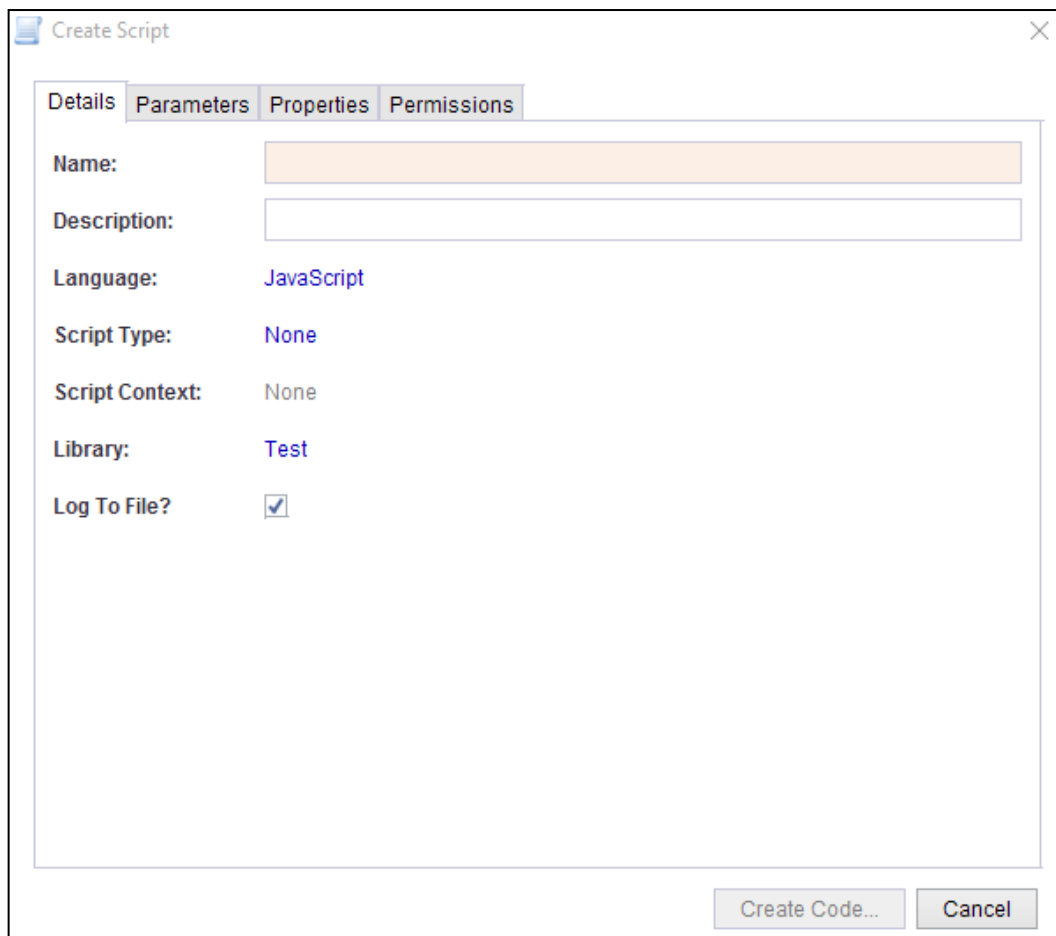## Writing Alert Scripts

Overview

Experian Pandora supports six Script Engines – JavaScript, Scala, Groovy, Python, Ruby and R. Bear in mind that the OS facilities available to the script will depend on the architecture the Experian Pandora server is running on.

Creating an Alert Script

You will need to create a Script Library prior to creating a script.

Navigate to **Home > Libraries > Script Library** and right-click to create a new Script Library

Then navigate to **Home > Libraries > Script Library > My_New_Script_Library** and right-click to create a new script.

Name the Script, select the Language and Script Type. You should ensure that the Language selection matches the language you have used in the script (one of JavaScript, Scala, R, Python, Ruby or Groovy).

If you want to debug the script tick the **Log To File** box. When this box is ticked any execution errors or output from your script will be directed to the script's log, accessed by right-clicking on the script and selecting **View Logs**.

In the Edit Code window, you may either type in the script's code, import a script file you have prepared earlier, or use the **Template** button to display a working template script in your chosen language. You may then save this script by pressing the **Save** button.

```
33  //
34  // The following methods are available for the parameters variable
35  //
36  // public boolean contains(String name)
37  // public TypedProperty get(String name)
38  // public String getString(String name)
39  // public long getNumber(String name)
40  // public boolean getBoolean(String name)
41  // public double getDecimal(String name)
42  // public int size()
43  // public boolean isEmpty()
44  // public String getPropertiesString()
45  // public Iterator<TypedProperty> iterator()
46  //
47  // Note that any output you write to stdout/stderr (or System.out/System.err) (i.e. print stateme
48  // will be written to a log file and this may be viewed by right clicking on the script name in
49  // explorer panel and selecting "View Output"
50  //
51  var user =  (psession) ? psession.getUsername() : "Administrator";
52  var scriptName    = parameters.getString("scriptName");
53  var scriptLanguage = parameters.getString("scriptLanguage");
54  var scriptAction   = parameters.getString("scriptActionType");
55  print("running the " + scriptLanguage + " " + scriptAction + " script (" + scriptName + ") on bel
56
57  // Print the type of the objects passed to the script
58
59  print("There are " + pobjects.size() + " elements in the pobjects list");
60  var idx = 0;
61  for each (var pobject in pobjects) {
62      var className = (pobject) ? pobject.getClass().getName() : "null";
63      print("pobjects[" + idx++ + "] class: " + className);
64  }
65
```

Please refer to the 'Scripting and REST Guide' for more information.