User Guide

ReQuon 3

Document Version 3.3.11● 22 May 2026

SINGLE
STEP
CONSULTING

+61 403241096

enquiries@singlestepconsulting.com.au

Trademarks

ReQuon is a trademark of Single Step Consulting. All other trademarks or registered trademarks are the property of their respective owners.

Disclaimer

The information provided in this document is provided "as is" without warranty of any kind. Single Step Consulting disclaims all warranties, either express or implied, including the warranties of merchantability and fitness for a particular purpose. In no event shall Single Step Consulting be liable for any damages whatsoever including direct, indirect, incidental, consequential, loss of business profits or special damages, even if Single Step Consulting or its suppliers have been advised of the possibility of such damages.

Document Lifetime

Single Step Consulting may occasionally update online documentation between releases of the related software. Consequently, if this document was not downloaded recently, it may not contain the most up-to-date information. Please refer to https://singlestepconsulting.com.au/requon-3 for the most current information.

From the Web site, you may also download and refresh this document if it has been updated, as indicated by the change date.

Where to get help

Single Step Consulting support, product, and licensing information can be obtained as follows.

Product information — Documentation and information about Single Step Consulting products and licensing can be found on the Single Step Consulting website at:

http://www.singlestepconsulting.com.au/products

Technical support

Option 1: Go to http://www.singlestepconsulting.com.au/support, click the “Log a Support Request button” and then follow the prompts.

Option 2: Send an email to support@singlestepconsulting.com.au providing details of your query / request.

Note: to open a service request, you must have an active paid subscription / valid support agreement.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Please send your feedback for this document to:

requon3@SingleStepConsulting.com.au

If you have issues, comments, or questions about specific information or procedures, please include the title and, if available, the revision, the page numbers, and any other details that will help us locate the subject that you are addressing.

Table of Contents

Appendices

1 Introduction

ReQuon 3 is a Data Reconciliation and Quality Tracking tool developed by, and for, Data Migration professionals. Having been actively used across a variety of projects and industries, the solution encapsulates the expertise and learnings of the author across many years and includes many time-saving features to fast track any data reconciliation effort.

1.1 Purpose

The purpose of this document is to provide guidance and practical advice on how to get the most out of your implementation of ReQuon 3.

1.2 Scope

This document covers the use of the ReQuon software, including explanations of the key concepts involved. While basic reconciliation tasks can be accomplished with little or no SQL experience, it is expected that practitioners with moderate to strong SQL skills will be required to perform more advanced reconciliation and data quality extraction tasks.

1.3 Intended Audience

This document supports a range of target audiences.

- Project Managers and Data Migration Managers will benefit from a shared understanding of the key concepts and reconciliation techniques supported by the tooling. The shared understanding will assist in communication and scope definition between the project management team and the data migration / reconciliation practitioners.
- Data Engineers will get practical advice and guidance on how to:
    - Scope and plan out the data reconciliation and data quality tracking activities.
    - Setup and configure the software.
    - Operate the software to define reconciliation checks, execute and test the reconciliation checks and report and publish the reconciliation results.

2 System Overview

2.1 Key Features

ReQuon 3 supports 2 complementary sub-processes within a typical data migration exercise:

1) data reconciliation between the source and target systems, and
2) data quality measurements over time (either at the source and/or the target) to track data quality improvement initiatives.

ReQuon 3 includes the following key features and capabilities:

- Supports connectivity to ODBC compliant data sources, as well as Excel files and delimited or fixed width text files.
- Supports both synchronous and asynchronous data capture (where supported by the relevant ODBC/OLEDB drivers)
- Supports a range of different reconciliation types including:
    - Type 1 – Record Counts by table / file
    - Type 2 – Group-By Aggregates
    - Type 3 – Entity-Dimension Sampling
- Provides the ability to import table and column metadata, from source and target systems, to increase productivity and reduce errors in data capture definitions.
- Supports the bulk generation of data capture skeleton definitions based on the table and column metadata to fast track the data reconciliation exercise.
- Provides a data storage repository for data capture definitions, promoting knowledge capture and potential reuse across projects utilizing similar source or target systems.
- Supports both Ms Access and SQL Server as the backend data storage repository
- Supports reporting of reconciliation results across different temporal baselines (e.g., Dry Run 1, Dress Rehearsal, Final Go-Live, etc.)
- Supports parameter substitution within capture definitions, to minimize / eliminate the need to adjust the capture definitions when moving from one baseline to another.
- Supports the definition, capture and reporting of key data quality metrics that need to be improved and tracked to support a successful data migration.
- Provides the ability to execute data extraction queries from source and/or target systems directly into Excel files (Pull-Thru Queries). Typically used in conjunction with the Data

Quality Metrics analysis to extract erroneous records for distribution to Data Quality Owners to action the required data quality remediations.

- Supports the recording and linking of reconciliation anomalies to reconciliation results for reporting and tracking purposes.
- Supports the generation of consolidated Excel based Reconciliation Report(s) for distribution to interested parties, using customizable templates.

2.2 Key Concepts / Terminology

2.2.1 Stages

The technical reconciliation approach within ReQuon employs a concept of "**Stages**" through which the migrated data moves from source to target (refer to the diagram below).

Each Stage typically represents a physical layer from which reconciliation metrics can be captured and subsequently compared (reconciled) against metrics captured from a different Stage. For the set of metrics to be reconcilable, the metric definitions need to consider any selection criteria and/or transformation rules that have been applied between one stage and any subsequent stages.

Understanding the relevant data migration approach and environments architecture applicable to a particular project will help determine:

- What **Stages** are involved.
- What reconciliations may be required, between different **Stages**.

The following diagrams illustrate a set of typical data migration environment architectures and how the **Stage** Numbers may be allocated to various systems / layers.

In its simplest form, a reconciliation is performed between a Source Database and a Target Database. The target database may be a straight one-for-one copy of the source database, or it may have been heavily transformed.

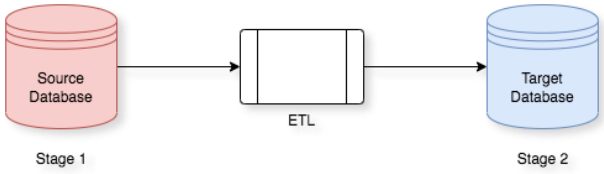


Figure 1 - Simple Source to Target Architecture

In more complex architectures, a dedicated migration staging environment with 1, 2 or more intermediate layers are employed to simplify the transformation (and reconciliation) process between various layers / stages of the data migration end-to-end processing.

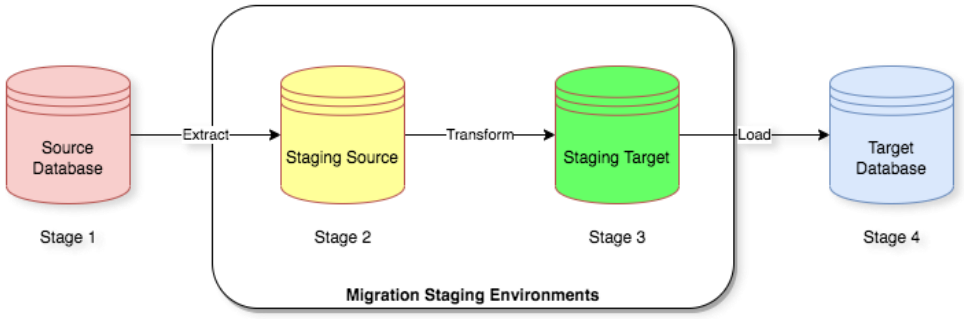


Figure 2 - Two-Layer Migration Staging Architecture

In Figure 2, Staging Source (Stage 2) would typically represent a selective one-to-one copy of the Source Database (Stage 1). The Staging Target (Stage 3) layer would typically represent a transformed "load ready" data set awaiting final loading into the Target Database (Stage 4).

Note: A **Stage** may represent a separate database, a particular schema within a single database or a collection of Excel or flat files.

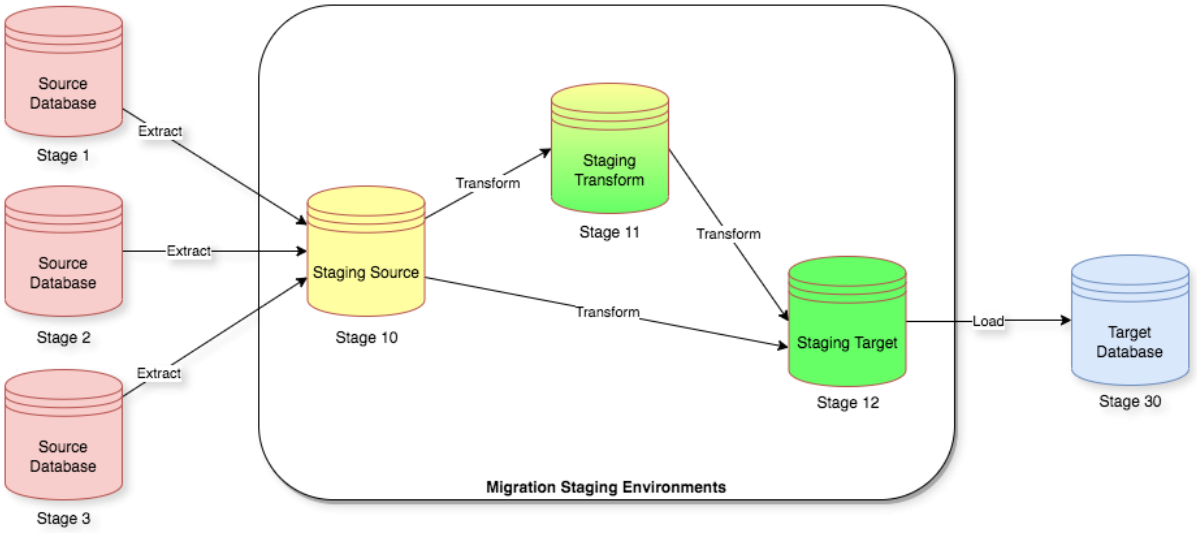


Figure 3 - Multi-Source + 3-layer Migration Staging Architecture

In Figure 3, the introduction of multiple data sources may lead to increased complexity in the transformation processes (e.g., for merging, deduplication, etc.) requiring the introduction of an intermediate Staging Transform layer for a subset of tables. Regardless of the structure involved, the overall reconciliation exercise can be decomposed into a set of reconciliations between 2 stages.

Note: Stage number 0 is reserved for storing the default values for substitution-parameters – these will be discussed in more detail in section 4.1.10, Shortcut Menus and Miscellaneous Popups.

2.2.2 Load Id

Reconciliation metrics need to be captured at an appropriate point in time to be effective. The concept of a **Load ID** is used to define a temporal reference point / baseline against which a set of metrics can be captured and subsequently reported.

For example: -

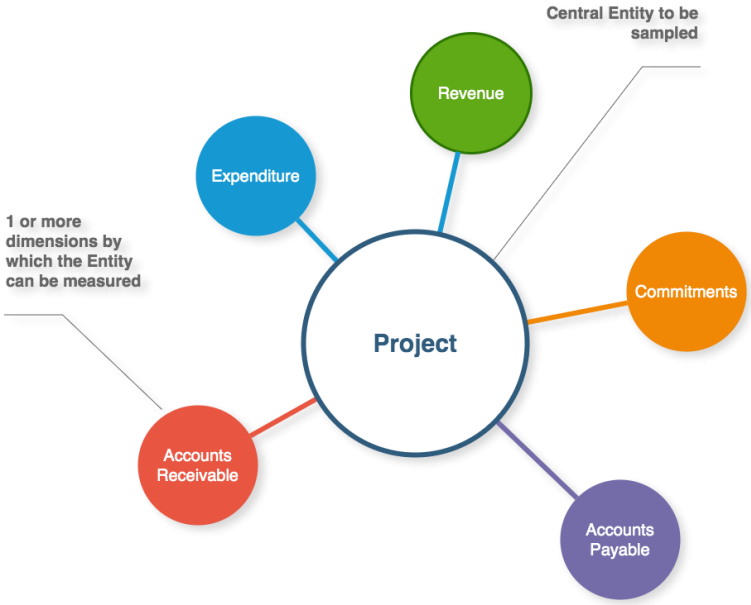
Load 1 might refer to an internal Dry Run
Load 2 could refer to Dry Run 1
…
Load 5 could refer to formal Dress Rehearsal 1
Load 6 could refer to formal Dress Rehearsal 2
Load 7 could refer to the final Go-live conversion

The number of reconciliation baselines required will usually be defined in the project's Data Migration Strategy and informed by the acceptability of the results of prior reconciliations.

2.2.3 Reconciliation Types

Three main (3) reconciliation types are supported in ReQuon as defined in the table below.

Type	*Definition*
1	**Record Counts.** This is the simplest reconciliation method and is generally performed first before embarking on the more complex type 2 and type 3 reconciliations. This type of reconciliation is generally performed for all tables.
2	**Group-By Aggregates.** This type of reconciliation groups the data by one or more key attributes within the data set (e.g., Year and Month associated with a transaction date, Customer Types, etc.) and then aggregates one or more associated attributes (e.g., sum of invoice amount, etc.) for each grouping. Comparisons (reconciliations) are then performed for each data grouping / aggregate. The purpose of this type of reconciliation is to verify the correct distribution of key values (e.g., totals by month, etc.) within a chosen data set. As such, this type of reconciliation is typically applied to key transactional tables / data sets (such as Purchase Orders, Invoices, etc.).
3	**Entity – Dimension Sampling**. The purpose of this type of reconciliation is to provide a more business-oriented view of the reconciliation data. It does this by executing a set of Type 2 Group-By Aggregates (referred to as Dimensions) around a central business entity (such as a customer, supplier or project, etc.) to provide a holistic view of that business entity.

Type	*Definition*
	This type of reconciliation is often complex and time consuming and hence is generally used in sampling context i.e., a subset of business entity ID's (such as the Project Numbers associated with the Top *n* projects based on some criteria) are selected for analysis. 

The progressive application of the different types of reconciliations, beginning with Type 1's and moving into Type 2's and Type 3's, supports a layered, risk-based approach.

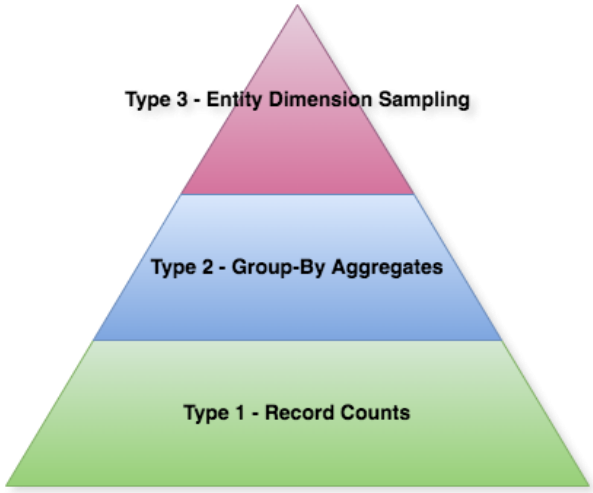


Figure 4 - Layering of Technical Reconciliation Types

2.2.4 Processing Statistics

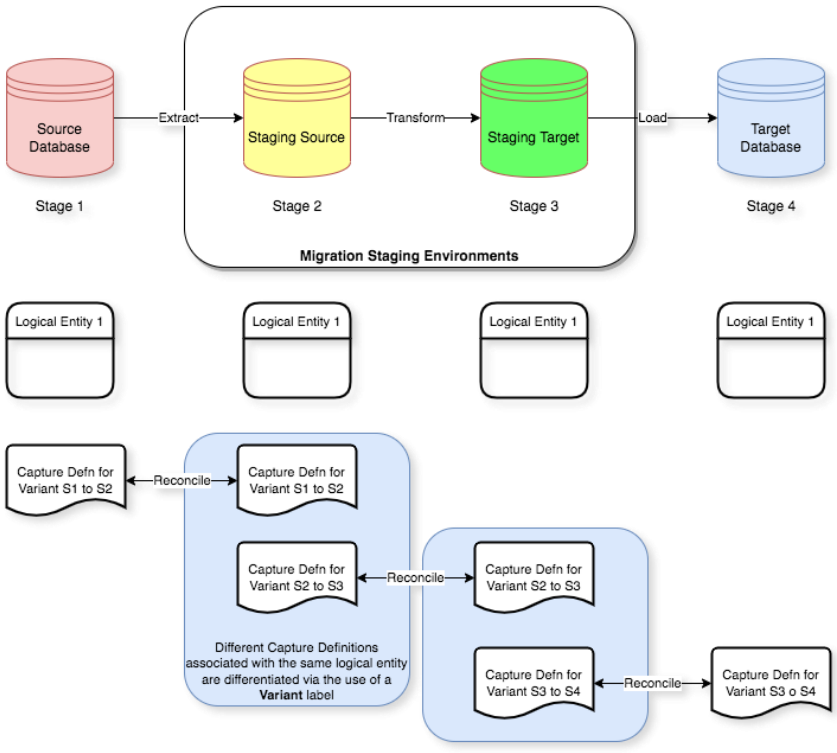
In addition to the 3 types of reconciliations, a complimentary set of metrics (referred to as Processing Statistics) can also be captured. Progressing Statistics are not used in Stage-to-Stage comparisons, but are rather used to provide summary reports such as: -

- The number of staging records that passed or failed validation
- The number of staging records that are unprocessed vs processed vs failed.

Processing Statistics help identify exceptions that may have led to a reconciliation variance detected in the Type 1, 2 and/or 3 reconciliation results.

2.2.5 Variants

Often, different data reconciliation metrics will be captured from the same table for different purposes. For example, the metrics captured from a **Staging Source** (Stage 2) table for comparison with a **Source Database** (Stage 1) table may be different from the metrics that need to be compared (reconciled) with a **Staging Target** (Stage 3) table where certain records have been excluded based on the defined selection criteria.



To differentiate between these metrics, the keys used to identify a specific metric definition include an additional label referred to as the "**Variant**". E.g., a variant of "S1-2" may be used to refer to a metric designed to reconcile between Stages 1 and 2 while "S2-3" would be used to refer to a metric designed to reconcile between Stages 2 and 3.

Similarly, if different selection criteria apply to different releases, the Variant label may include an indication of which release the reconciliation definition applies. E.g., S1_4_Rel1 may be used when reconciling a table between Stage 1 and Stage 4 for the R1 release, while S1_4_Rel2 could be used for the R2 release.

Another use of the **Variant** label occurs where data from a single source table gets split into 2 or more tables in a downstream stage, or vice versa i.e., a One-To-Many or a Many-To-One scenario. For example, a Party table from the source system may get split into Supplier, Customer and Employee tables as part of the transformation rules. In such a situation, three variants of the same metric may be defined such as "S1-2_SUPP", "S1-2_CUST" and "S1-2_EMPL".

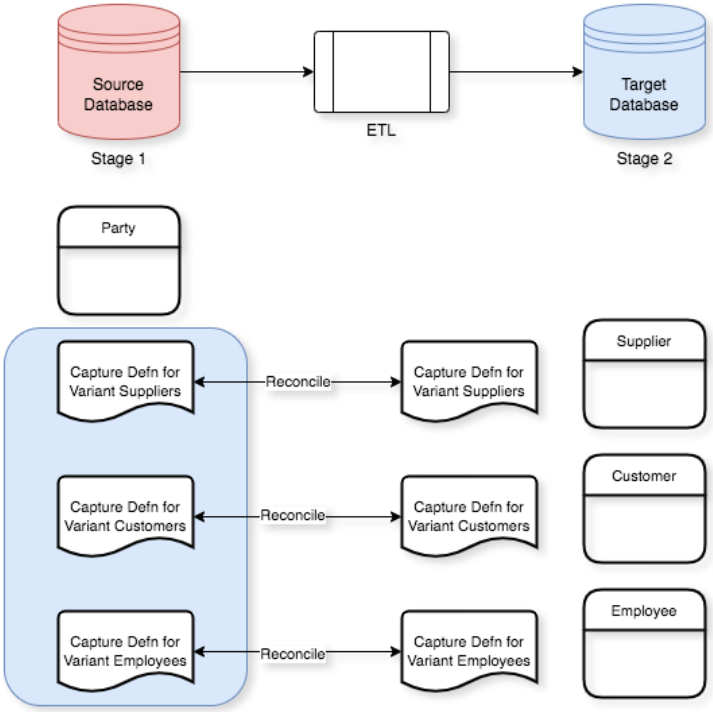


Figure 5 - Use of Variants in One-to-Many (or Many-to-One) Scenarios

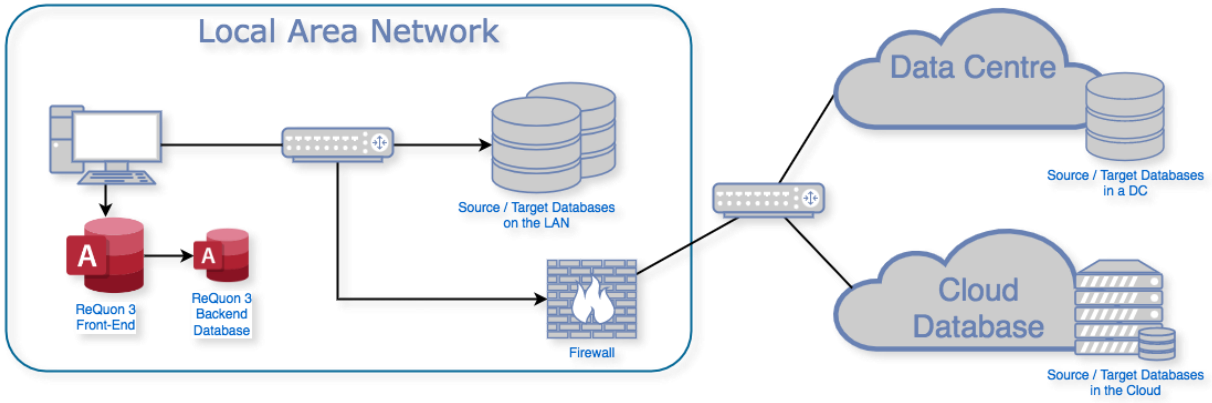
2.3 Environment

ReQuon can be setup in a few different environment configurations to suit the size and needs of the data reconciliation team. Typical configurations include:

- Single-user standalone
- Multi-user shared Ms Access backend on a network drive
- Multi-user shared SQL Server backend (local network or cloud hosted)

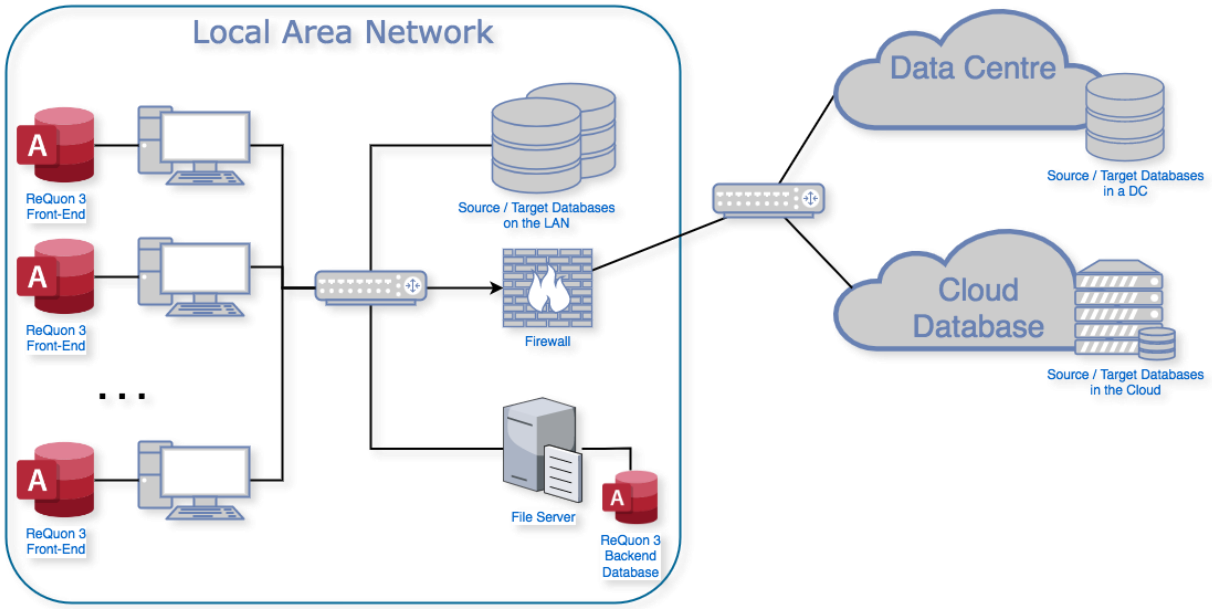
2.3.1 Single-User Standalone

In this configuration, ReQuon is installed on a single Windows PC using a local Ms Access backend database.



2.3.2 Multi-User, Shared Ms Access Backend

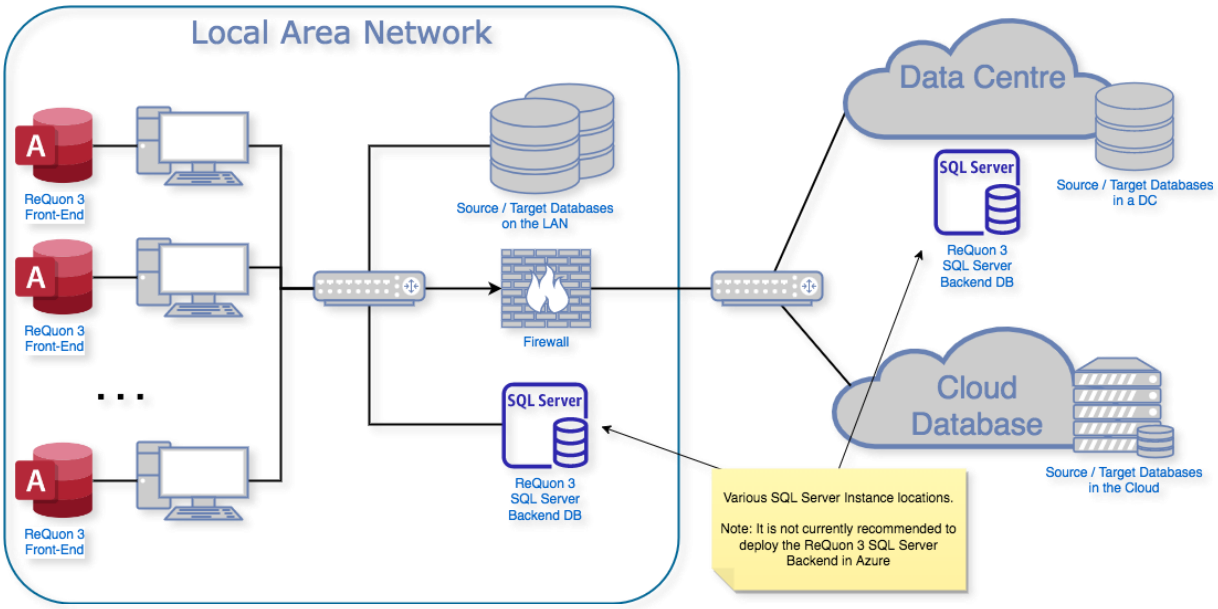
The configuration is suitable for small project teams (e.g., up to 1 to 10 users)



2.3.3 Multi-User, Shared SQL Server Backend

The configuration is suitable for small to large team sizes. The SQL Server database instance where the ReQuon 3 backend will be hosted can reside in various locations, including:

- On-Premise within the Local Area Network, or
- In a company Data Centre



3 Installation

3.1 Hardware Requirements

The following **minimal** hardware requirements are generally based around the Microsoft recommendations for Ms Office which include a Windows capable PC with:

- 4 GB RAM; 2 GB RAM (32-bit)
- 4 GB of available disk space
- 1280 x 768 screen resolution

3.2 Software Requirements

The following additional software components are required in a typical ReQuon installation.

- Windows 11, Windows 10, Windows 8.1, Windows Server 2019, Windows Server 2016.
- Microsoft Access (2016, 365) – Full version or the free runtime edition. ReQuon supports both 32 bit and 64-bit editions.
- Microsoft Excel (2016, 365)
- ODBC and/or OLEDB drivers for the relevant source and target systems to be reconciled.
- Optional: SQL Server backend database.

3.3 Pre-Installation Checklist

Prior to commencing the installation, you should decide which environment configuration best suits your needs. The following table identifies the relevant items applicable to each configuration.

Step	Instructions	Single-User Standalone	Multi-User, Access Backend	Multi-User, SQL Server Backend[1]
1	Identify the end-user PC's and ensure they have at least Windows 7 (preferably Windows 10 or 11)	✓	✓	✓
2	Ensure you have an Administrator password for the PC(s) (required to register some libraries as part of the installation)	✓	✓	✓
3	Ensure each PC has an up-to-date version of either: • the full version of Microsoft Access 2016 / 365, or • the equivalent run-time edition	✓	✓	✓
4	In multi-user environments using MS Access backends, ensure all end-user PCs are running the same bitwise version of Ms Access (e.g., 32 bit or 64 bit). **Warning: DO NOT** mix bitwise versions as this will lead to corruption of the backend database. It is also **highly recommended** to maintain all users on the same version of Ms Access to eliminate the possibility of incompatibility between versions.	N/A	✓	N/A
5	Confirm the network directory for the Ms Access shared database and ensure the end-users have read, write and delete permissions in the directory.	Optional	✓	N/A
6	Identify the SQL Service instance on which the ReQuon backend database(s) – one for each project	N/A	N/A	✓

[1] ReQuon 3 Pro Edition only

3.4 Installing the system

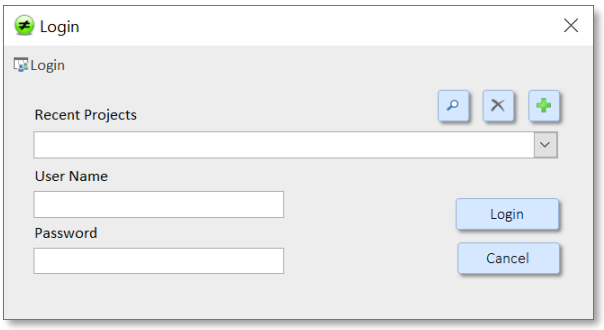
3.4.1 Client Installation

Step	Instructions	Actioned
1	Download the ReQuon 3 Client installation program from the link provided in your subscription confirmation	
2	Execute the installer and following the instructions provided. This will create the application directory structure and copy the application files into the relevant directories.	
3	Download your subscription licence file from your subscription confirmation email and save a copy to the application folder (e.g. C:\ReQuon)	

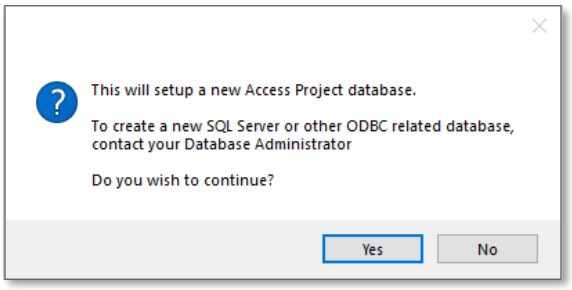
3.4.2 Ms Access Backend Installation

3.4.2.1 On a Local PC

Upon initial start-up of the application, the following popup will be displayed.



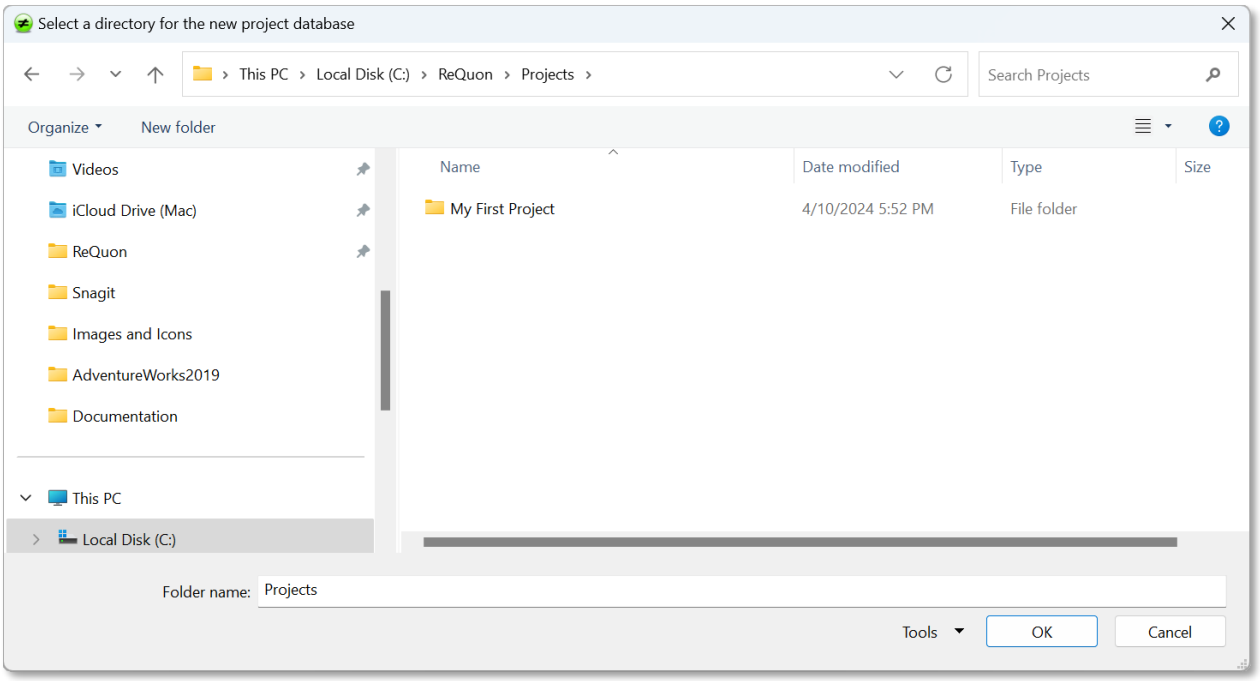
To create a new Ms Access based backend database, click on the  button. On the subsequent dialog box, click on the Yes button to continue.



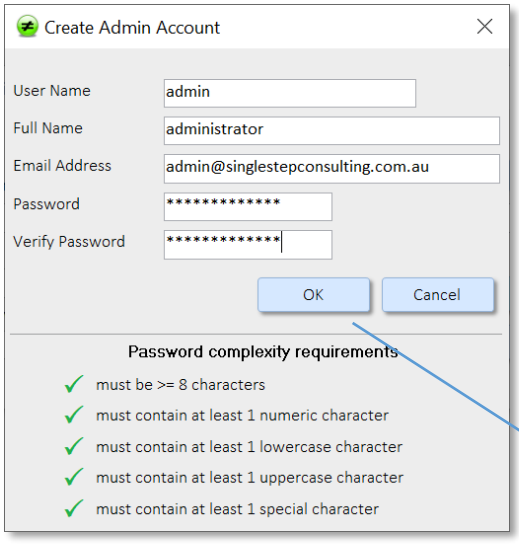
A Windows dialog will appear where you can create, navigate to and select a directory to store your project directory.

Note: Each ReQuon 3 backend database stores the reconciliation capture definitions and captured metrics for one Project.

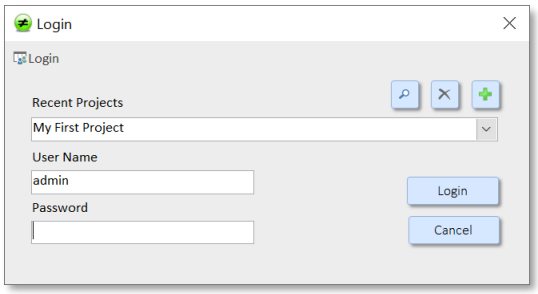
In the following example, a directory "C:\ReQuon\Projects\My First Project" has been created and selected. Click OK to continue.



A new empty project database will be created in the selected directory and then the user will be prompted to create the administrator account for the new project database.



Enter the relevant details and click OK to create the admin account and open the new project database. You will be presented with the ReQuon login screen, with the newly created project database pre-selected in the list of "Recent Projects". Enter the admin account User Name and Password and click the Login button to continue.



3.4.2.2 On a Network File Server

Prior to creating a new Ms Access based project database on a network file server, perform the following steps:

- Create a drive mapping on the end-user's PC pointing to an appropriate directory on the file server. E.g., R: drive
- Create a suitable directory structure to store future ReQuon project databases under the mapped driver. An example directory structure would be:

```
R:\ReQuon\Project Dbs\
```

- Ensure the end-user as Read, Write and Delete permissions on the lowest level folder in the directory structure and that the permissions are inherited to any subsequently created sub-directories.

Once this initial setup has been completed, the creation process for a new Ms Access based project database on a network file server is identical to that for a local database.

3.4.3 SQL Server Backend Installation

Currently under development.

4 Application Overview

4.1 Navigation

The main application window has four (4) main areas:

1. Title Bar
2. Drop-down menus
3. Main document window
4. Status / message bar

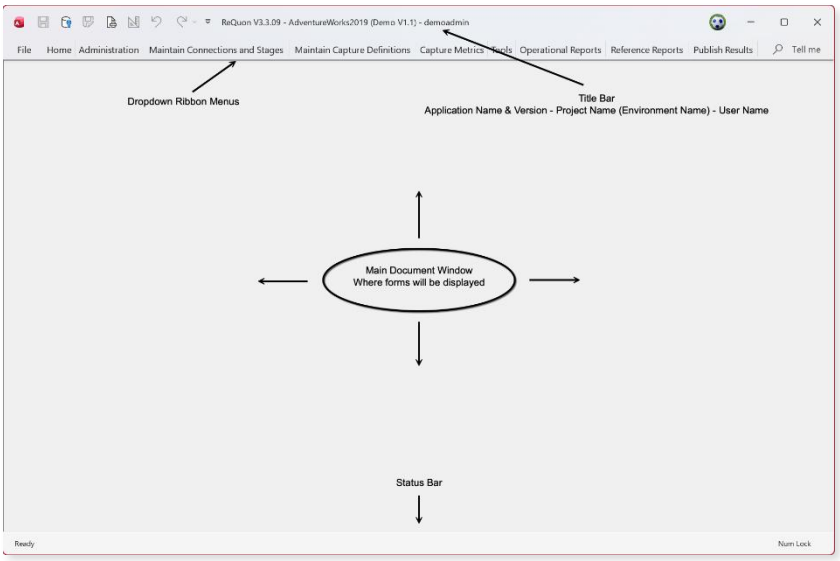


Figure 6 - Overall Application Window

Starting with the Administration Menu, the organization of the drop-down menus, working from left to right, generally represents the chronological order in which the functionality within the application will be exercised.

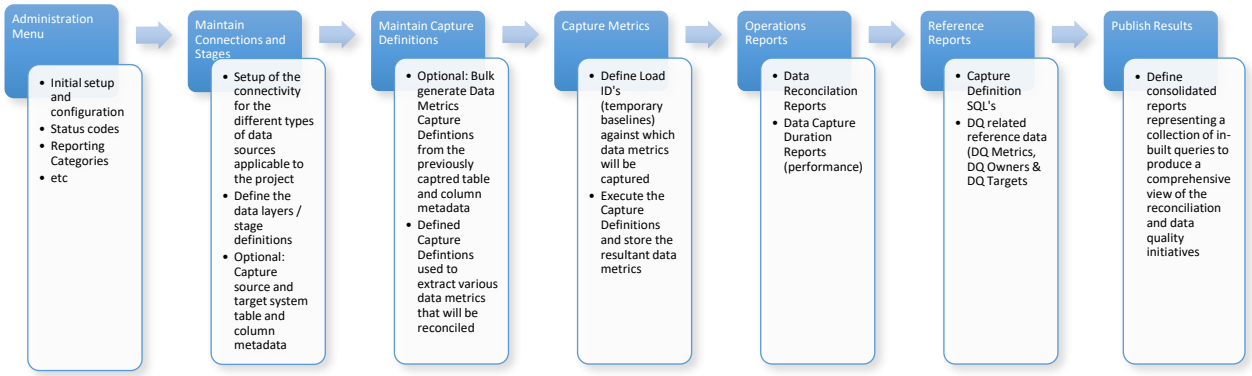
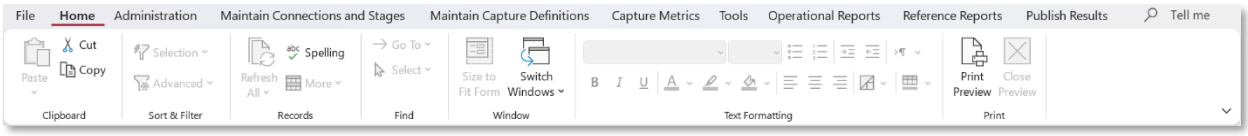


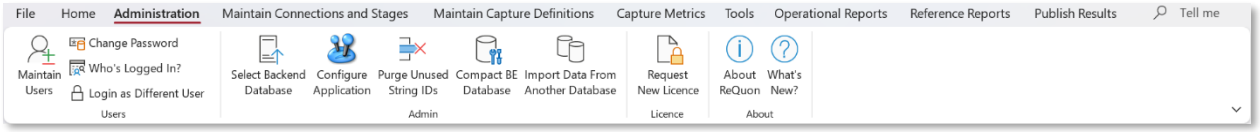
Figure 7 – Menu Ordering vs Activity Ordering

4.1.1 Home Menu

The home menu provides access to basic functionality that is shared across the entire application such as Copy and Paste, Finding and Filtering of data, etc.

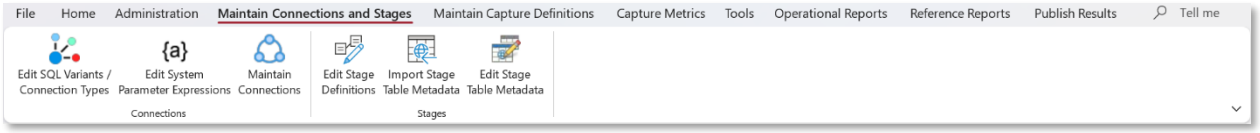


4.1.2 Administration Menu



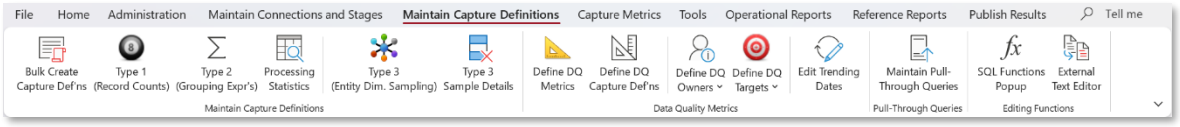
Menu Group	Menu Item	Notes
Users	Maintain Users	• Add, edit and change users and user roles. • Reset passwords
	Change Password	
	Who's Logged In?	View who is logged into either the front-end application or backend database
	Login as Different User	Displays the login screen where either: • a different user can be selected for the current project database, and/or • a different user and project database
Admin	Select Backend Database	Used to select an existing project database and make it the current database ready for a subsequent login.
	Configure Application	Provides access to application / project specific configuration settings and reference data.
	Purge Unused String IDs	ReQuon maps textual values (such as Group By values returned in Capture results, entity samples values, etc.) to String IDs to reduce space across multiple load id's. This function will delete any String ID's that are not being referenced so that the backend database size can be reduced – refer Compact BE Database below.
	Compact BE Database	**For Ms Access backend databases only**, this function will run the standard Compact and Repair function to reduce the size of the backend database and attempt to repair any data corruption, if detected.
	Import Data from Another Database	• Allows the user to import data from another ReQuon project database (Ms Access only). • Useful for sharing reference data between projects as well as reusing Capture Definitions for projects involving similar source and/or target systems.
Licence	Request New Licence	Request a new ReQuon licence
About	About ReQuon	Provides application version, licence details and access to support library details
	What's New?	Display recent release notes

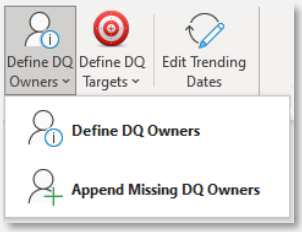
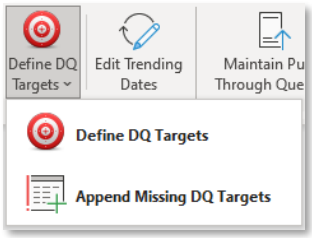
4.1.3 Maintain Connections and Stages Menu



Menu Group	Menu Item	Notes
Connections	Edit SQL Variants / Connection Types	• A **SQL Variant** recognizes that different databases (e.g., Ms Access vs SQL Server vs Oracle, etc.) contain variations in supported SQL syntax. ReQuon comes with several pre-configured SQL Variants, however the user can configure a set of key syntax settings to support other variations. • Connection Types represent the different Connection String syntaxes that can be used to connect to a database via ODBC / OLEDB connections. For example, SQL Server supports Standard Security (SS User and Password), Trusted Connections, Azure Standard Security, etc.
	Edit System Parameter Expressions	• Different SQL Variants also support different syntax for key functions such as date and time related calculations that are used in several predefined substitution parameters.
	Maintain Connections	• This is where the actual DNS-Less connection strings for each source and target database are defined. • If different client computers have different ODBC drivers for the same database, it is possible to define a different connection string for each client computer for the same "logical" connection name.
Stages	Edit Stage Definitions	• Used to define the persistent data **Stages** of interest to the reconciliation exercise. • Note: a single "logical" connection can relate to 1 or more **Stages.**
	Import Stage Table Metadata	• Used to review a list or tables or views from a particular "Stage", select the items of interest and import the column metadata (names, data types, sizes, etc.) into the ReQuon project repository. • The imported metadata can be used to bulk generate Capture Definition skeletons as well as provide lookups the users maintaining the capture definitions.
	Edit Stage Table Metadata	• Provides the ability to manually maintain the imported metadata. This is particularly useful for Excel and text files where the correct data types may not be retrieved via the applicable driver.

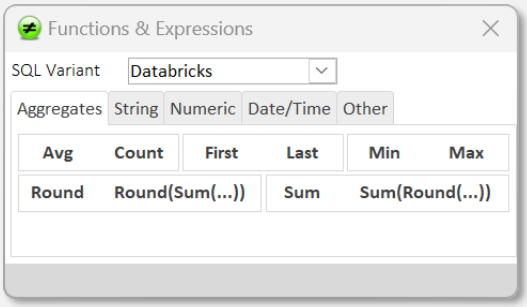
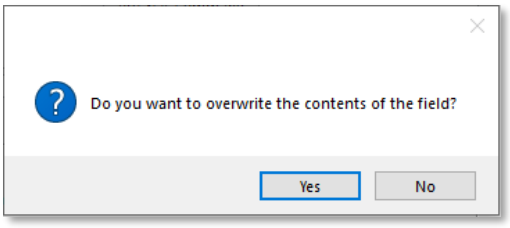
4.1.4 Maintain Capture Definitions Menu



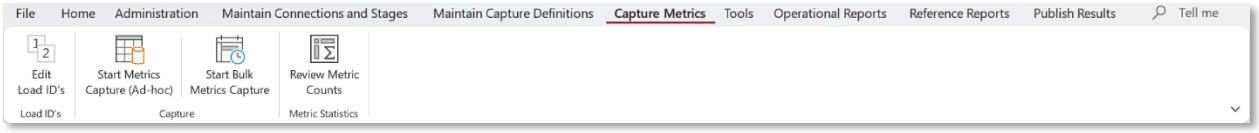
Menu Group	Menu Item	Notes
Maintain Capture Definitions	Bulk Create Capture Def'ns	• Used to bulk generate Capture Definition skeletons from the imported Stage Table metadata. The skeleton definitions can then be allocated to the Reconciliation Engineers to build out the capture definitions as required.
	Type 1 (Record Counts) Type 2 (Grouping Expr's) Processing Statistics Type 3 (Entity Dim. Sampling)	• Used to maintain the relevant type of Capture Definition
	Type 3 Sample Details	• Used to import or edit a list of Entity Id's by Entity Type. • Define whether the Entity Id should be treated as Case Sensitive or Case Insensitive. • Individual sample values can be assigned an optional "Sample Set" label. This can be useful for selecting a small sample subset for development / testing purposes, for example.
Data Quality Metrics	Define DQ Metrics	• Used to define the set of DQ metrics that need to be measured and tracked over time. Generally, these will represent the key data quality measures that need to be achieved before a final go-live data migration is performed.
	Define DQ Capture Def'ns	• Used to maintain the Data Quality Capture Definitions used to measure and collect the DQ Metrics
	Define DQ Owners 	• **Define DQ Owners:** Accountability for the DQ Metrics needs to be assigned to an DQ Owner within the Business. The DQ Owners will be the recipients of the DQ Reports and DQ Extracts containing details of the data that needs to be remediated. • **Append Missing DQ Owners:** Will create DQ Owners that are found in the DQ Metrics captured that haven't been previously defined. These can then be edited via the "Define DQ Owners" form.
	Define DQ Targets 	• **Define DQ Targets:** Each DQ Metric can be assigned a target value representing the acceptable number of invalid records (usually zero) that need to be achieve by a given date. • Append Missing DQ Targets: Will create DQ Targets for any combinations of DQ Metric, Owner and relevant drilldown levels (1 to 4) found in the DQ Metrics captured that haven't been previously defined. These can then be edited via the "Define DQ Targets" form.

Menu Group	Menu Item	Notes
	Edit Trending Dates	• Used to define the regular (e.g., daily, weekly, fortnightly or monthly) DQ Target measurement dates to support the reporting of the metrics over time and assess if the business is on track to meet the targets by the relevant target date.

Pull-Through Queries	Maintain Pull-Through Queries	• Pull-Through Queries are used to extract data from a **Stage** and write the data to one or more Excel files for subsequent review and analysis
		• Typically used in conjunction with the Data Quality Metrics, to extract the erroneous data requiring remediation, for sharing with the relevant DQ Owner
		• Other potential use-cases include:
		  ○ Comparing two data sets at the field level by using Row Hashes
		  ○ Preparing a set of data load files that contain size & content constraints
		  ○ Extracting source system reference data

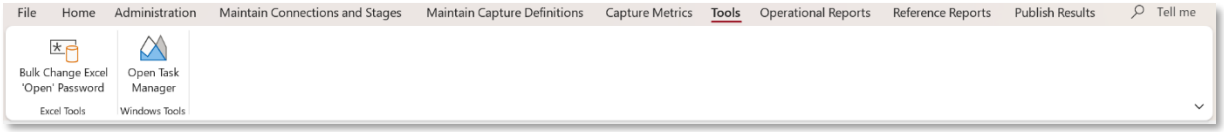
Editing Functions	SQL Functions Popup	• Displays a floating form with common SQL functions & expressions. Used to insert into relevant fields on the Capture Definition forms, using the syntax specific to the selected SQL Variant. 
	External Text Editor	• This command will copy the text from the currently selected control to the external editor (default is NotePad.exe) defined under the Application Settings menu. • If the text is edited and then the editor closed, the user is asked if they what to overwrite the contents of the current control. 

4.1.5 Capture Metrics Menu



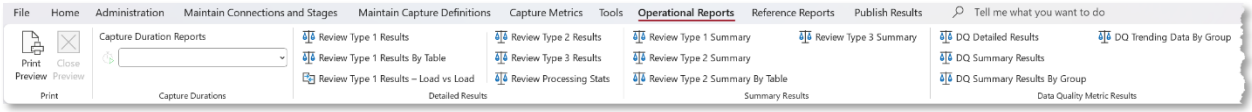
Menu Group	Menu Item	Notes
Load ID's	Edit Load ID's	• Used to define meaningful identifiers for the specific point-in-time data captures and reconciliation results. Refer to section 2.2.2 Load Id, under Key Concepts
Capture	Start Metrics Capture (Ad-hoc)	• Allows the user to execute all Capture Definitions of a specific type (E.g., Type 1 Record Counts) and/or one or more individually selected capture definitions within a type. Particularly useful while developing and testing the Capture Definitions.
	Start Bulk Metrics Capture	• Allow the user to edit and maintain a Data Capture queue or different types of Capture Types for different Stages and then execute the same. • Execution can be either immediate or at a specific date / time in the future (scheduled).
	Review Metric Counts	• This function displays the record counts of the detailed metric captures, by Capture Definition and Load ID (or Measurement Date for Data Quality related captures).

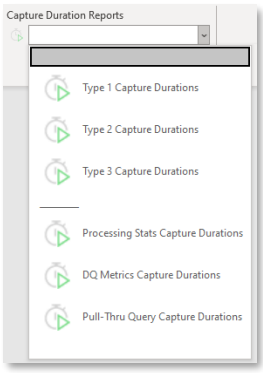
4.1.6 Tools Menu



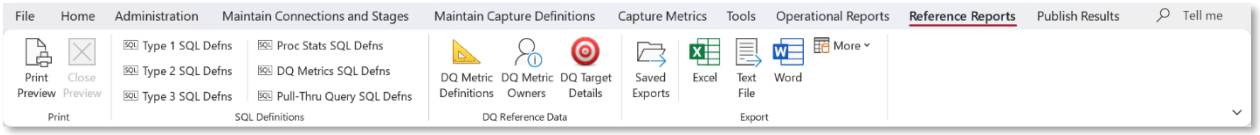
Menu Group	Menu Item	Notes
Excel Tools	Bulk Change Excel 'Open' Password	• Used to add, change or remove 'Open' passwords from a selection of Excel workbooks
Windows Tools	Open Task Manager	• Opens the Windows Task Manager. Useful to check for Ms Access or Excel processes running in the background

4.1.7 Operational Reports Menu



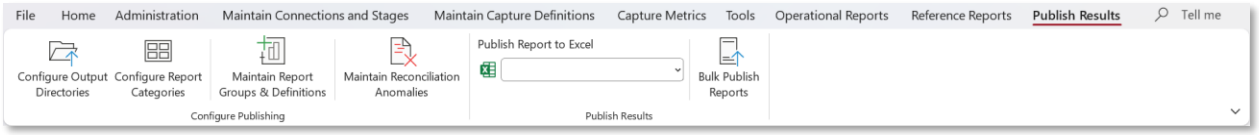
Menu Group	Menu Item	Notes
Capture Durations	Capture Duration Reports 	• Report on the amount of time each individual Capture Definition took to execute. • Useful during development and testing to identify long-running queries that may require performance tuning to reduce the time taken to perform and end-to-end reconciliation process
Detailed Results	Review Type 1 Results Review Type 2 Results Review Type 3 Results Review Processing Stats	• View reports containing the lowest level of detail for each of the different reconciliation types.
	Review Type 1 Results by Table	• Where a Stage supports Table Members, the report rolls-up the results to the Table Levels • If the Stage does not support Table Members, this report effectively provides the same results as "Review Type 1 Results"
	Review Type 1 Results – Load vs Load	• This report compares the record counts from applicable to a single Stage, from one Load ID (e.g., Dry-Run 1) to another (e.g. Dry-Run 2). • Useful to assess data volume growth over time
Summary Results	Review Type 1 Summary Review Type 2 Summary Review Type 2 Summary by Table Review Type 3 Summary	• Summarised versions of the results indicating where the reconciliation is a Pass / Fail status
Data Quality Metrics Results	DQ Detailed Results DQ Summary Results DQ Summary Results by Group DQ Trending Data by Group	• Collection of Data Quality related reports. The “By Group” versions allocate the DQ metrics to the nominated “Owner Group” responsible for addressing the issues

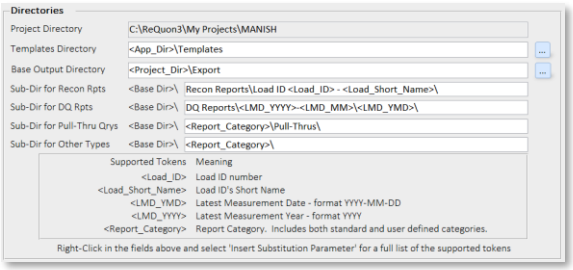
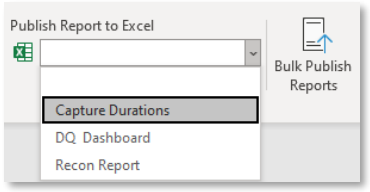
4.1.8 Reference Reports Menu



Menu Group	Menu Item	Notes
SQL Definitions	Various	• Used to report on the generated SQL Statements associated with the various Capture Definitions • Can be run for a specific Stage (or all Stages) • Can be run with or with substitution parameters values • Can be run for a specific Load ID (applicable when using parameter substitution)
DQ Reference Data	DQ Metric Definitions DQ Metric Owners DQ Target Details	• Used to report on the various data quality related reference data.

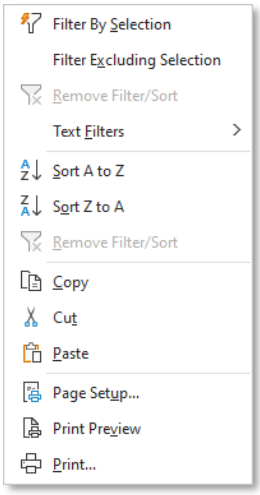
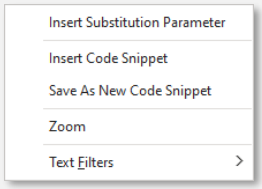
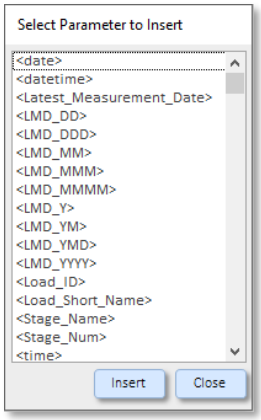
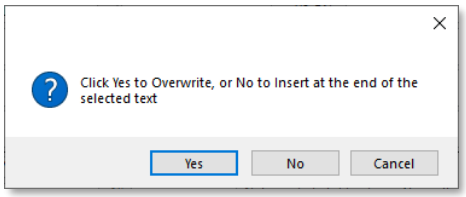
4.1.9 Publish Results Menu

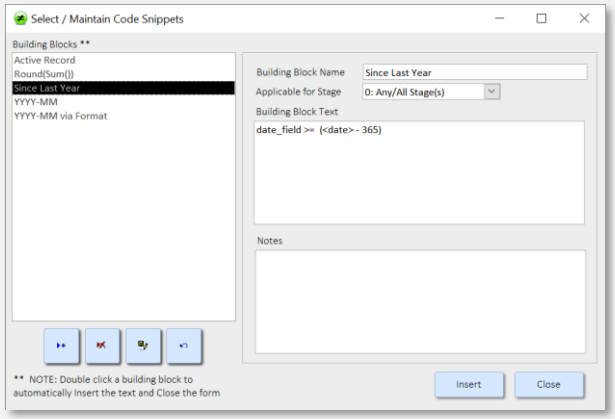
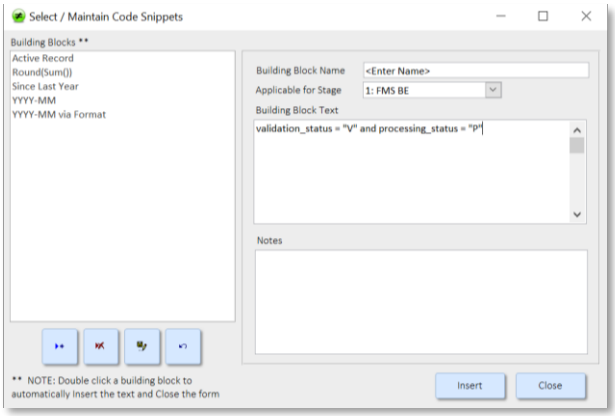


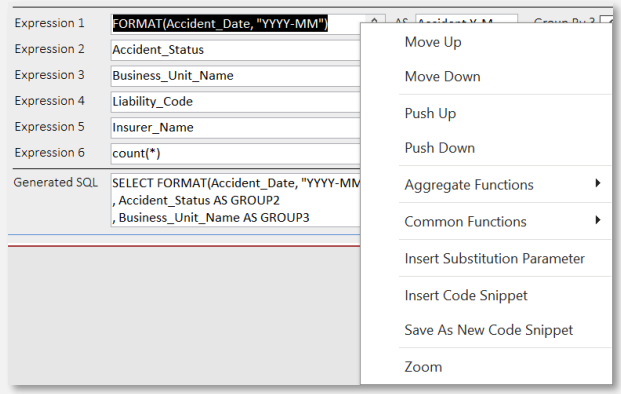
Menu Group	Menu Item	Notes
Configure Publishing	Configure Output Directories	• Configures the directory that the published reports should be exported to, dependent upon the type of extract. Supports several different substitution parameters. 
	Configure Report Categories	• Maintains the list of user-defined reporting categories that can then be associated with different extracts. • Note: Reporting Category can used as a substitution parameter in the definition of the Output Directories (see above)
	Maintain Report Groups & Definitions	• Allows the user to define a **Report Group** containing a collection of data extracts (based on pre-defined queries with user-defined parameter values) to produce a consolidated Excel based report. Typical "Report Groups" include: • Reconciliation Report • Data Quality Dashboard • Capture Durations / Performance • A predefined set of Excel templates are provided which can be customized to suit the organizational requirements
	Maintain Reconciliation Anomalies	• While not actually a report, this form is used to document the reconciliation anomalies identified in the reconciliation results, including any root cause analysis and planned resolutions. • This is usually the last activity required before the final Data Reconciliation Report can be published.
Publish Results	Publish Report to Excel Bulk Publish Reports 	• This dropdown presents the list of Report Groups (as previously defined above) for execution and saving to file for subsequent distribution to the interested parties • The Bulk Publish Reports option allows the user to select multiple reports to be generated and saved to a common folder for later retrieval.

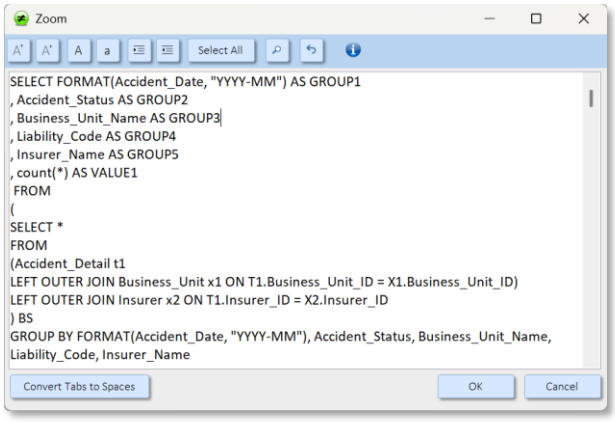
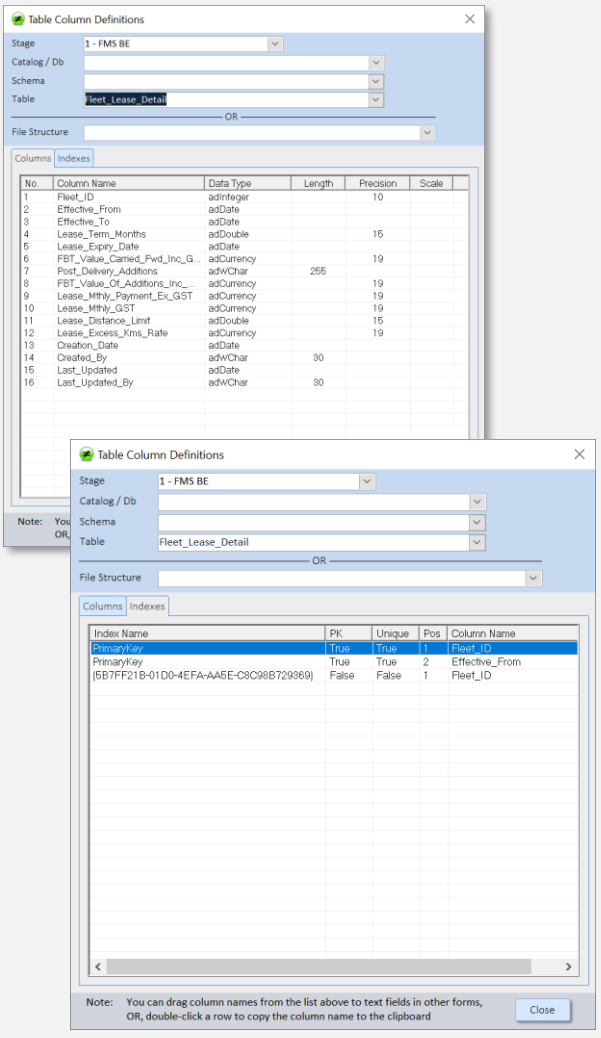
4.1.10 Shortcut Menus and Miscellaneous Popups

In addition to the editing and maintenance forms, there are several supporting shortcut menus and popup forms available throughout the application. Shortcut menus are accessed by right-mouse-clicking in any form field.

Shortcut Menu / Popup Name	Screenshot
Standard Shortcut Menu • Most common fields will display the following shortcut menu providing access to an array of filtering, record sorting and copy/paste functions.	
Insert Shortcut Menu Provides access to several other forms / popups.	
Insert Substitution Parameter • Allows the user to select one of the pre-defined substitution parameters and either replace the entire field or append the parameter at the end of the field.	 

Shortcut Menu / Popup Name	Screenshot
Insert Code Snippet Code snippets are re-usable pieces of SQL that may be applicable to multiple capture definitions. A typical example would be **Staging Source** table validation and/or status flags that are used to filter the valid records for transformation to the downstream **Staging Target** stage. • Allows the user to select and insert a previously defined code snippet either replacing the entire field or appending the snippet at the end of the field. • A code snippet can itself contain substitution parameters e.g., **<date>** in the example refers to the current date substitution parameter.	
Save as New Code Snippet • Allows the user to select a snippet of SQL from an existing field and save it as a new re-usable code snippet. • The selected text is copied into the Build Block Text field awaiting the user to assign a meaningful name to the snippet and specify which **Stage** the snippet is valid in.	

Expression Reordering
 	
• This shortcut menu is available for the aggregate expression fields on the Type 2 and Type 3 Capture Definition forms.
• This provides the following reordering commands:
  ○ **Move Up & Move Down**: These commands move the currently selected Expression row up or down while swapping position with the row above or below as applicable.
  ○ **Push Up & Push Down**: If there are empty expressions in the set of Expression rows, these commands push all the Expressions rows above or below the selected row while maintaining the relative positions of the expressions.
	

Shortcut Menu / Popup Name	Screenshot
Zoom • Opens a popup edit form containing the contents of the currently selected field • Available on long text fields • Can be accessed via the shortcut menu or by pressing Shift-F2 in an eligible field.	
Table Column Definitions (Stage table metadata) • Accessed via the information icon on the various Capture Definition forms (generally found next to the Table Name field) • Includes tabs for both Column and Index metadata • Column names can be dragged from the popup into the main form, either replacing the whole field, or if the cursor has been positioned in the target field, it will insert into the nominated position within the existing text.	

5 Detailed Instructions

5.1 Initial Setup and Configuration

After logging into the system as an administrator, there are several initial setup and configuration tasks that should be performed. These can be accessed via the "Configure Application" option under the Administration Menu (refer section 4.1.2). These tasks would generally be performed in the order (top to bottom) of the configuration groups listed in the left-hand side of the form, starting with "General Settings".

5.1.1 General Settings

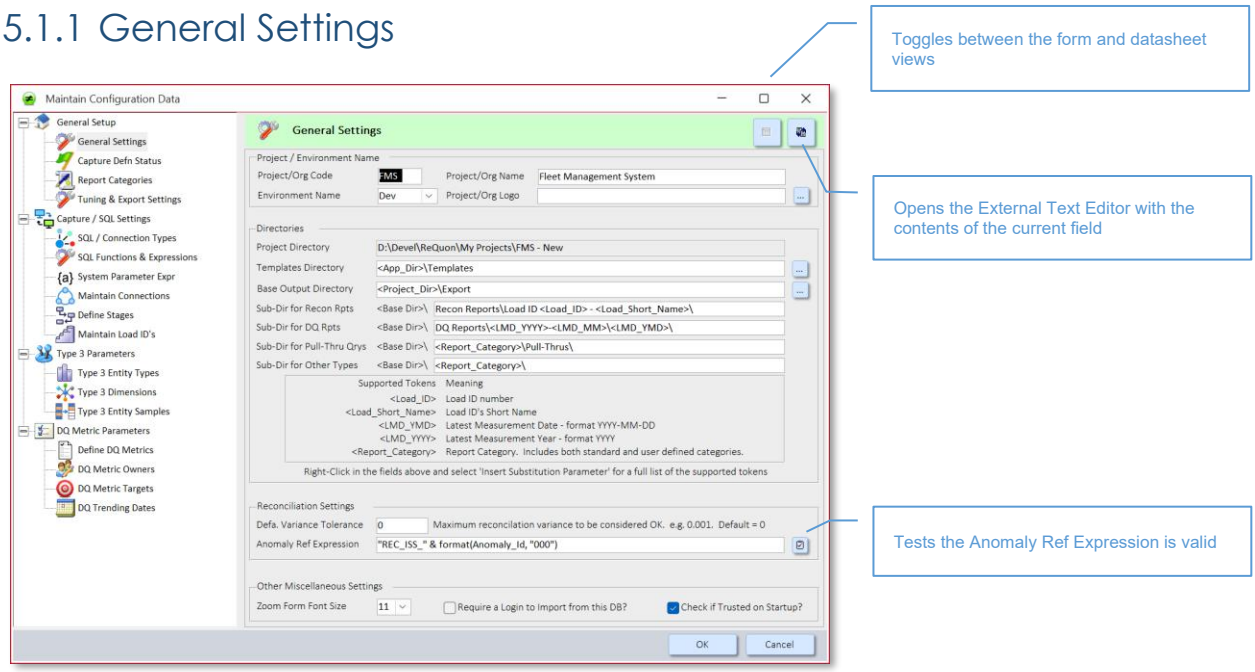
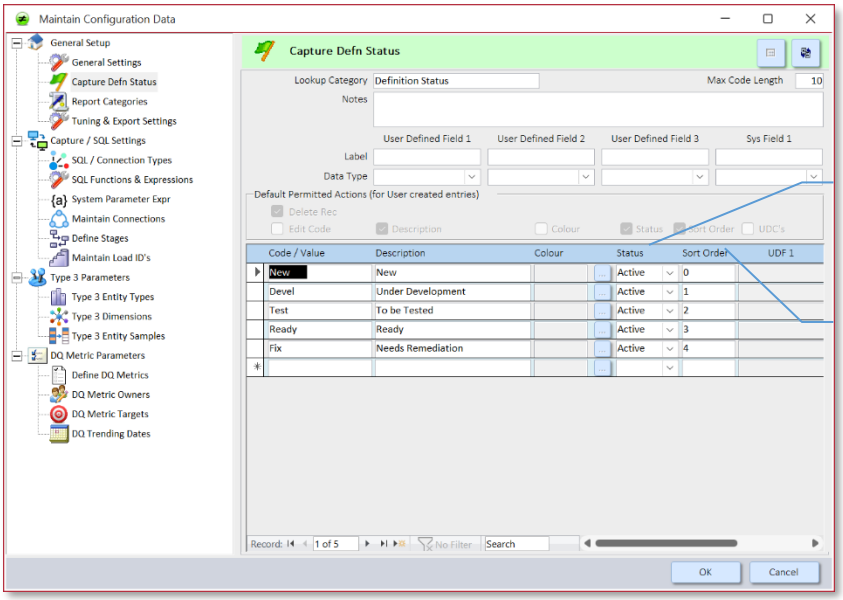


Figure 8 - General Settings

5.1.2 Capture Definition Status

Used to track the development and testing state of the Capture Definitions

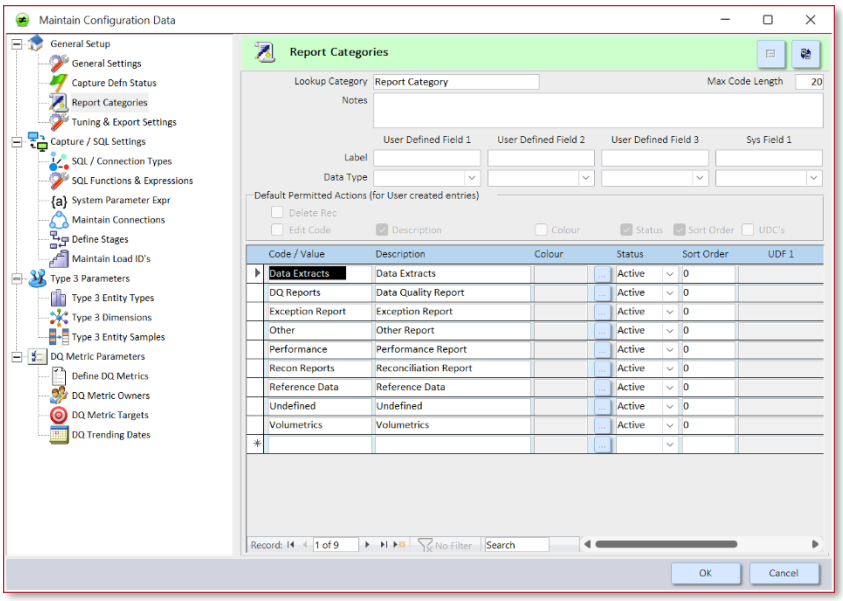


Inactive values will not be displayed in the dropdown lists on other forms

Order of the options as displayed in the dropdown lists

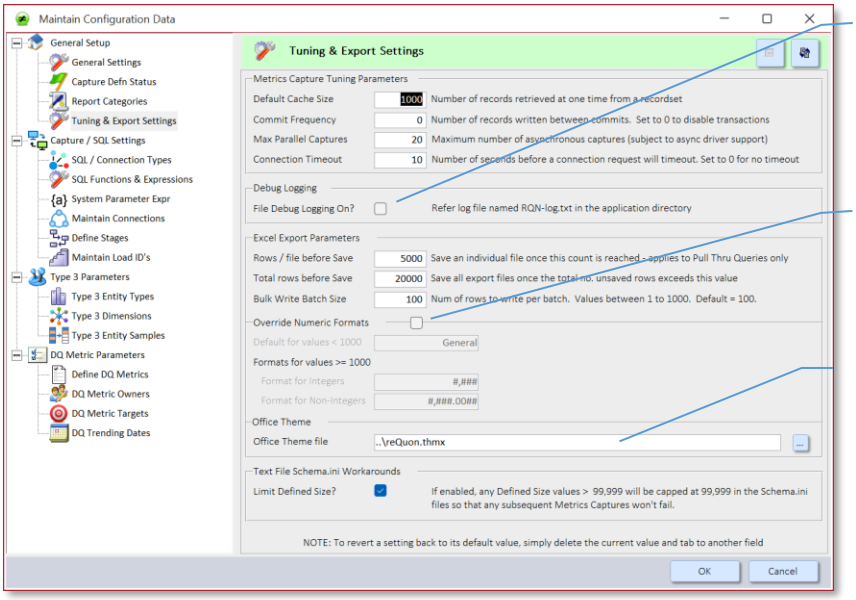
5.1.3 Report Categories

Used to categorise different types of report and provide substitution parameters that can be used to define the target export directory.



5.1.4 Tuning & Export Settings

It is generally recommended to leave the Tuning Parameters set to the default values.

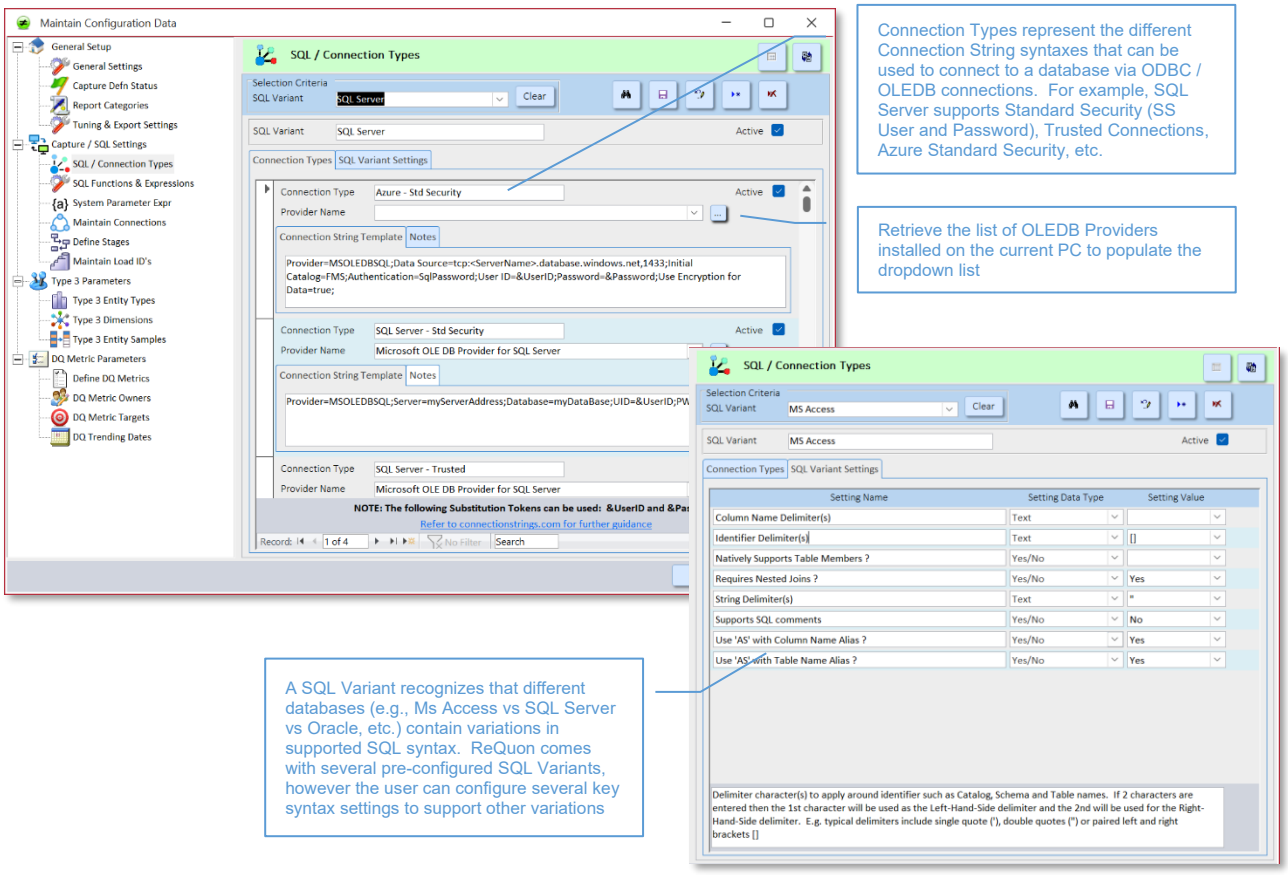


Enabling debug logging will slow down capture performance. Should be left off unless instructed by support when diagnosing an error condition

When exporting to Excel, enabling this option will override the default numeric formats.
Note: Using an Excel Template can provide more control i.e., at an individual field level

Specifies an Ms Office Theme to be applied to Excel export files, even where no template has been provided.

5.1.5 SQL Variants / Connection Types

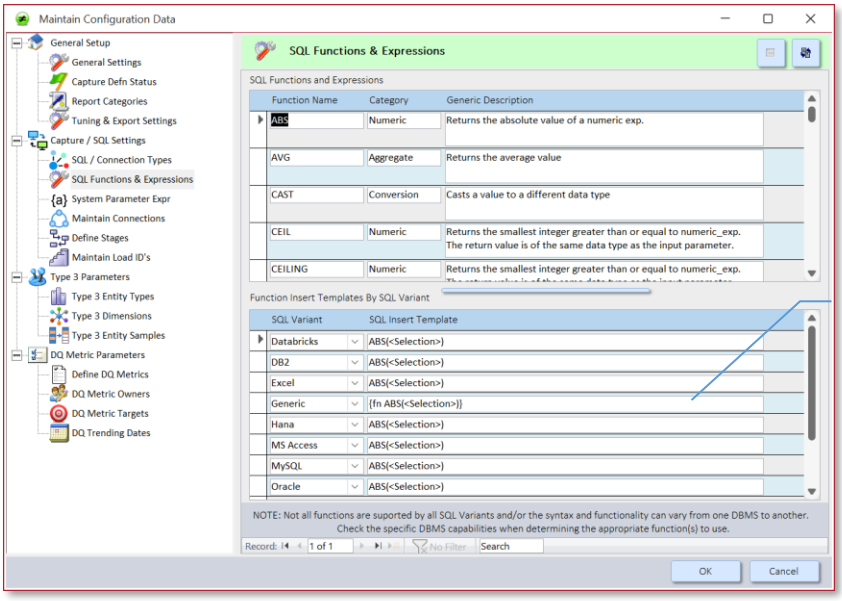


Connection Types represent the different Connection String syntaxes that can be used to connect to a database via ODBC / OLEDB connections. For example, SQL Server supports Standard Security (SS User and Password), Trusted Connections, Azure Standard Security, etc.

Retrieve the list of OLEDB Providers installed on the current PC to populate the dropdown list

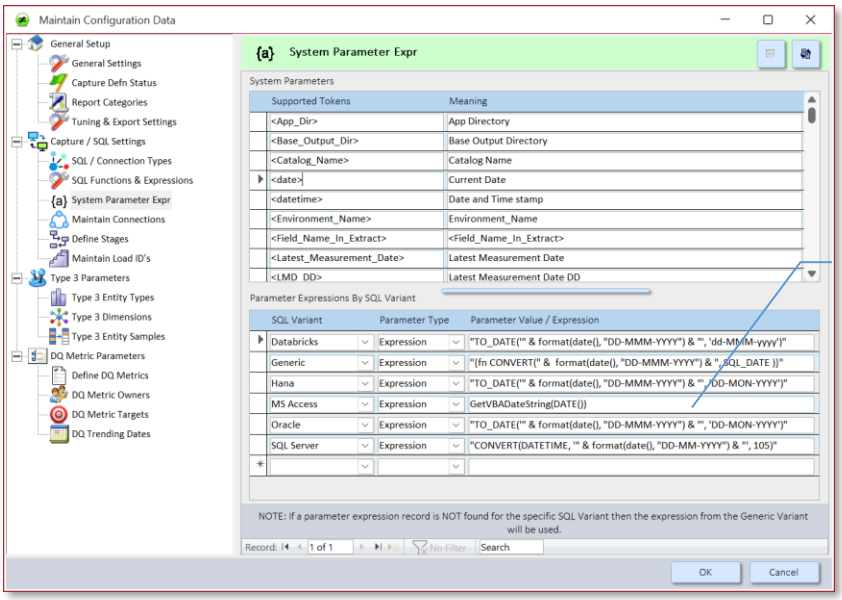
A SQL Variant recognizes that different databases (e.g., Ms Access vs SQL Server vs Oracle, etc.) contain variations in supported SQL syntax. ReQuon comes with several pre-configured SQL Variants, however the user can configure several key syntax settings to support other variations

5.1.6 SQL Functions & Expression



Where applicable, the Generic variant uses the ODBC Scalar Function referred to in Section Appendix B.

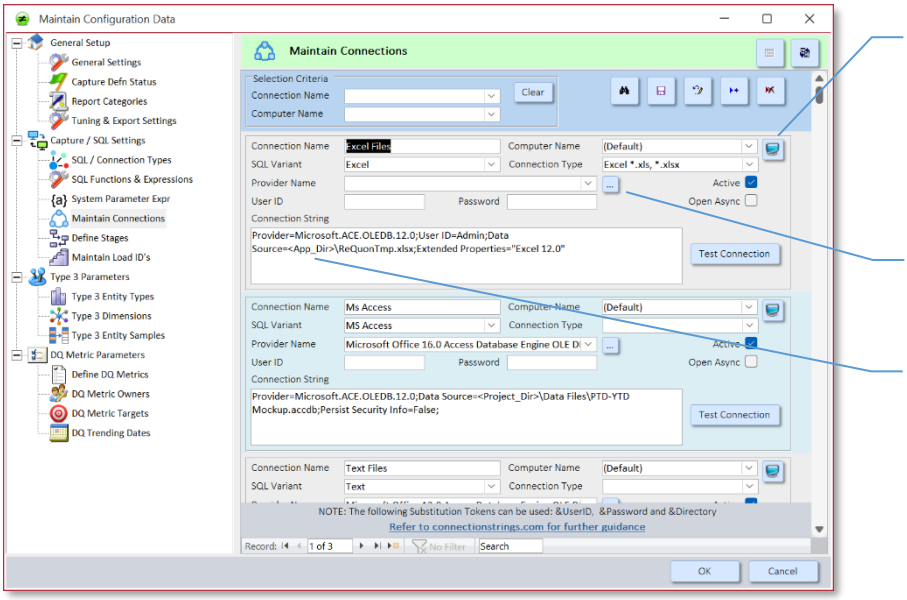
5.1.7 System Parameter Expressions



Different SQL Variants also support different syntax for key functions such as date and time related calculations that are used in several predefined substitution parameters.

5.1.8 Maintain Connections

This is where the actual DNS-Less connection strings for each source and target database are defined.



If different client computers have different ODBC drivers for the same database, it is possible to define a different connection string for each client computer for the same "logical" connection name.

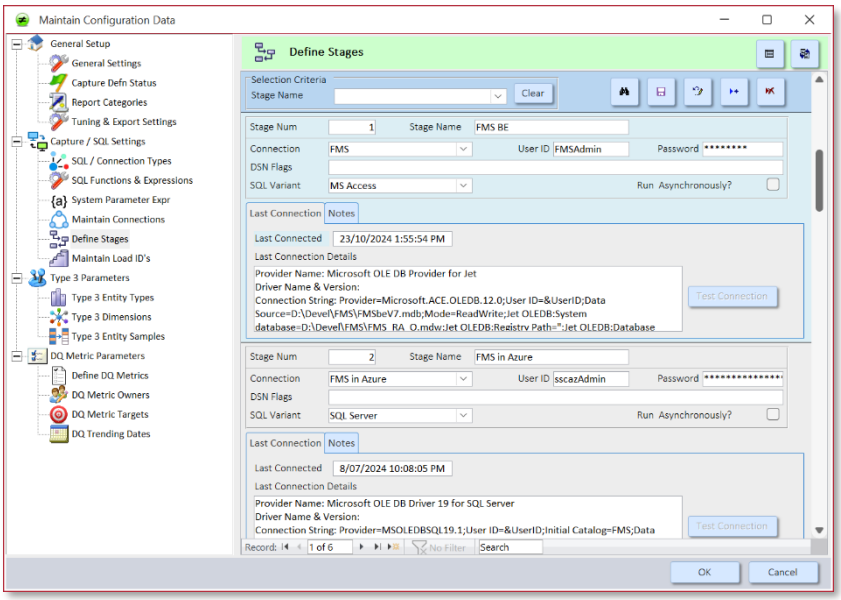
This button retrieves the current computer name.

Retrieve the list of OLEDB Providers installed on the current PC to populate the dropdown list

Several system substitution parameters (e.g. <App_Dir> and <Project_Dir>) can be included in the Connection String. They can be accessed via the right-mouse-click menu

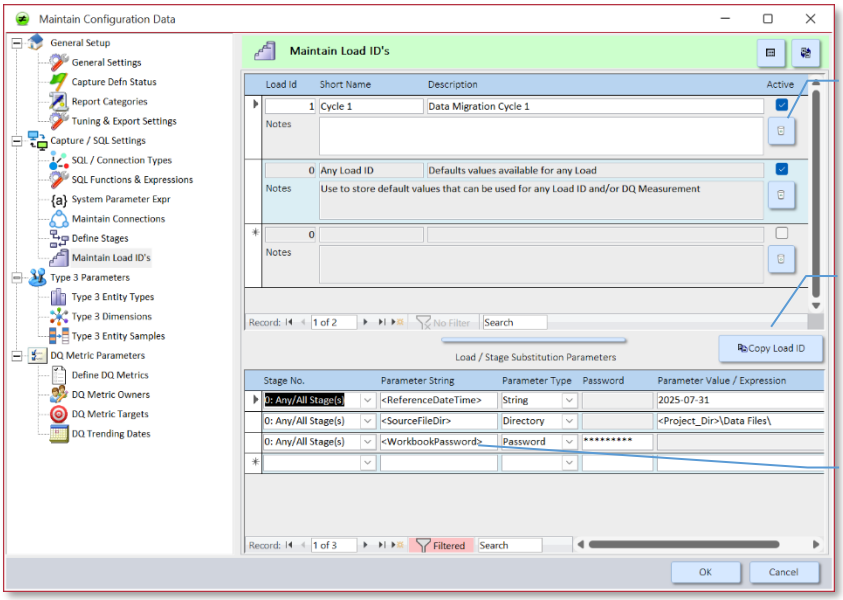
5.1.9 Define Stages

This is where we define the **Stages** of interest for reconciliation purposes and associate them with one of the previously defined **SQL Variants, Connections** and associated login details.



5.1.10 Maintain Load ID's

Used to define meaningful identifiers for the specific point-in-time data captures and reconciliation results. Refer to section 2.2.2 Load Id, under Key Concepts.



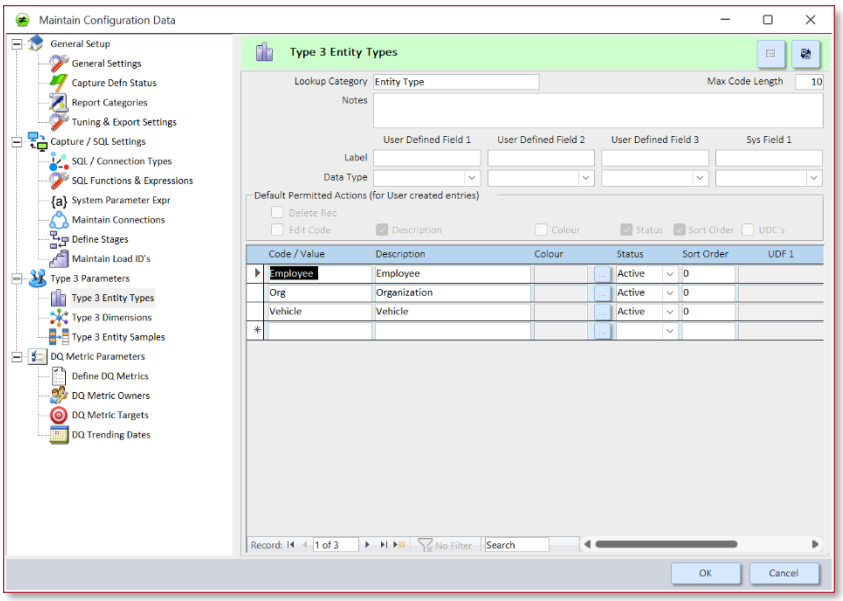
Used to purge all results associated with the selected Load Id. Deleting the Load ID record will also result in the purging of the results.

Copy Load ID will create a new Load ID and copy all the associated Substitution Parameters.

The Parameter String can use any string, with or without leading and/or trailing delimiters, as long as it won't get confused with the surrounding text / SQL.

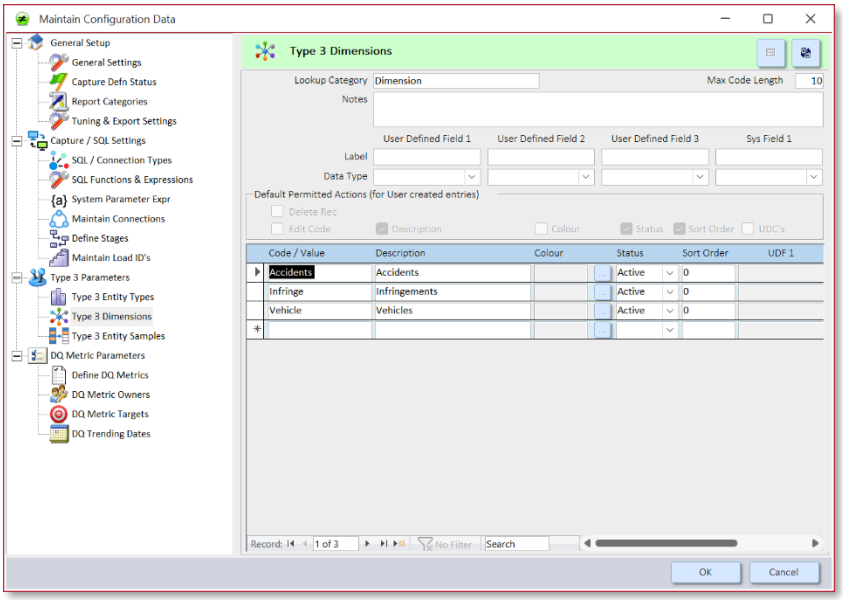
5.1.11 Type 3 Entity Types

This form is used to define the various business **Entities** of interest from a reconciliation perspective to be used within the Type 3 Entity Dimension Sampling Capture Definitions.



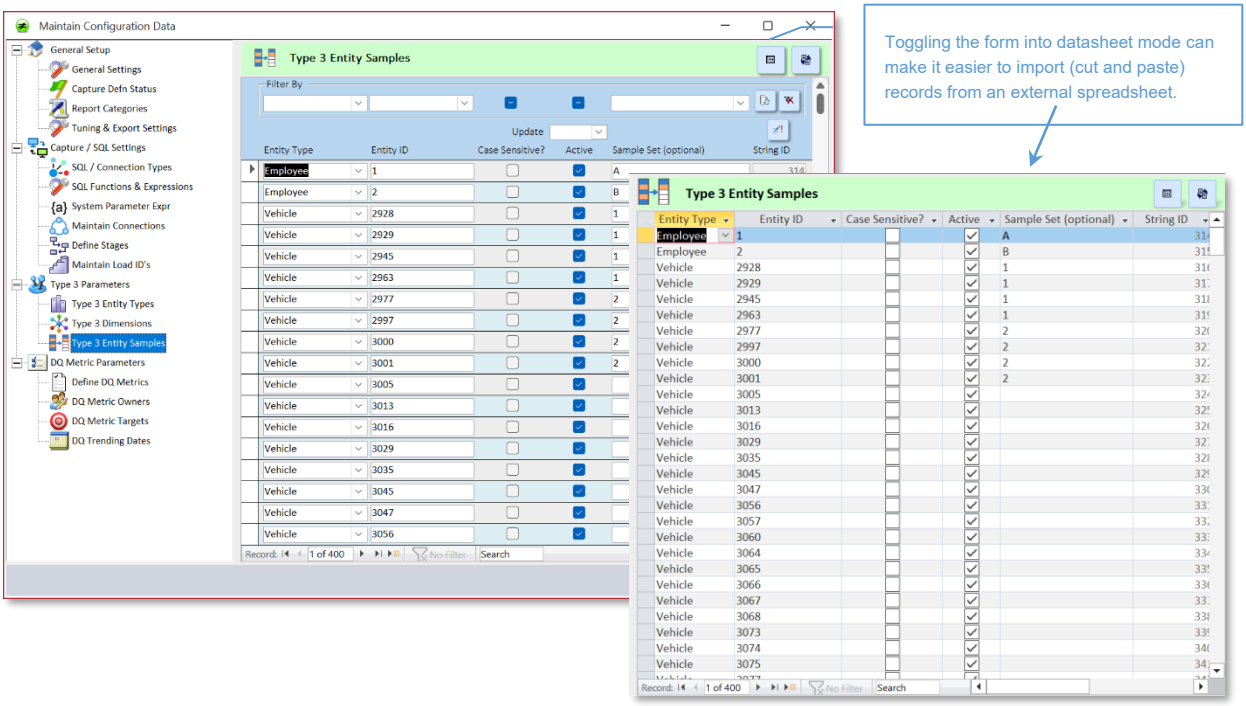
5.1.12 Type 3 Dimensions

This form is used to define the various **Dimensions** by which the business **Entities** will be measured within the Type 3 Entity Dimension Sampling Capture Definitions.



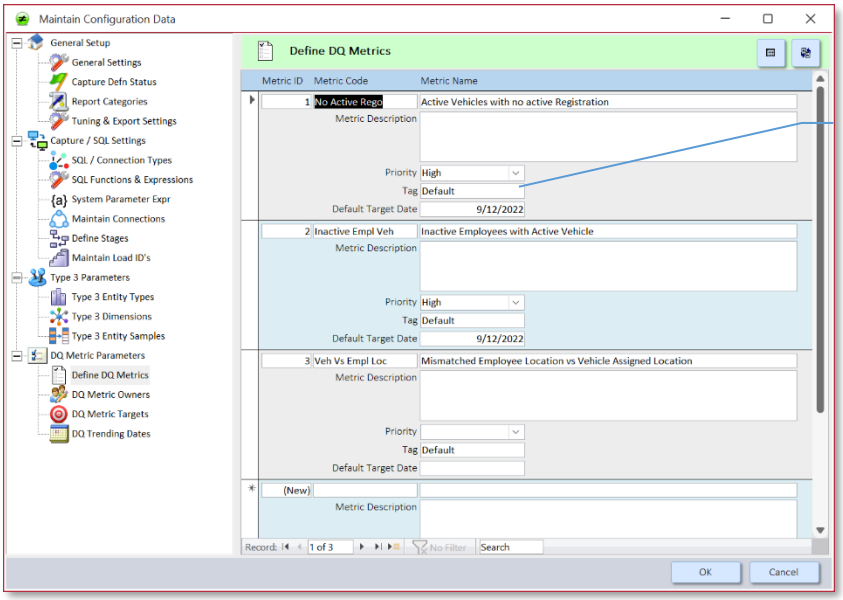
5.1.13 Type 3 Entity Samples

This form is used to define the **Entity Id's** sample values that the Type 3 Capture Definitions will be executed against and subsequently reconciled across its' various **Dimensions**.



5.1.14 Define DQ Metrics

This is used to define the key Data Quality Metrics that are important to monitor and track in the lead-up to the final data migration. The Default Target Date is the date by which the data quality target levels should be achieved.

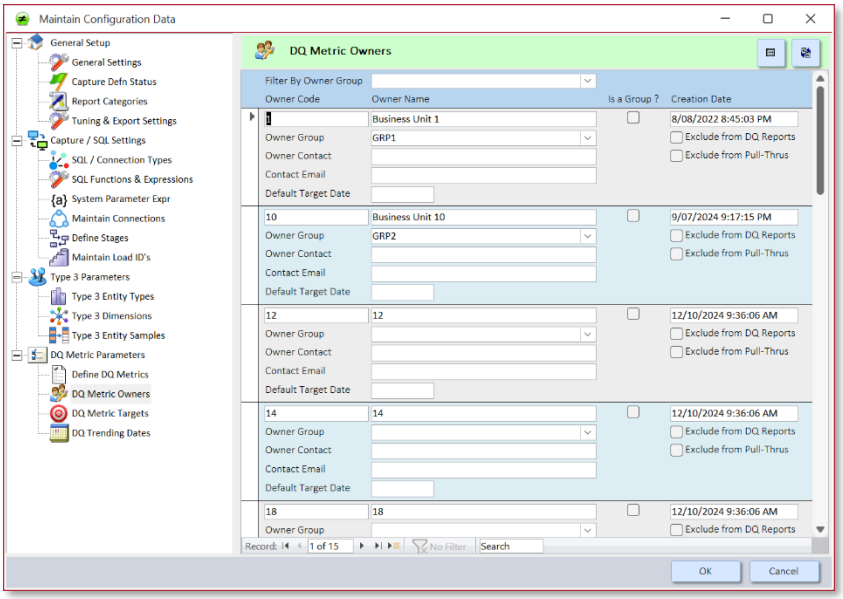


The Tag can used in the "DQ Trending Data By Group" reporting query to filter the set of metrics included in a particular report.

This can be useful where you need to produce different DQ reports for different parts of the business, such as Accounts Payables versus Accounts Receivables.

5.1.15 DQ Metric Owners (and Owner Groups)

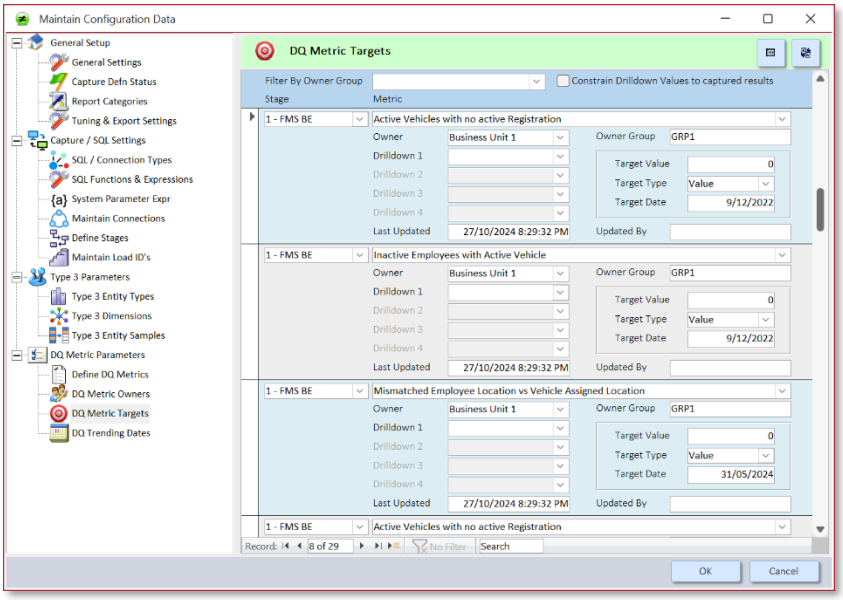
This form is used to define and maintain the Data Quality Owners (and Owner Groups) that will be accountable for achieving the Data Quality Targets by the required Target Date. The table below explains each of the fields in more detail.



Field	Description
Owner Code	The Owner Codes (that are not Owner Groups) should correspond to values contained within the DQ Metrics Capture results. An Owner Code that IS an Owner Group may or may not exist with the DQ Metrics Capture results. A typical example of a data field from which Owner Codes would be sourced would be an Organization Code or a Business Unit ID i.e., where the organization and/or business unit takes responsibility for remediating any identified data quality issues.
Owner Name	Self-explanatory. Used in DQ Reports
Is a Group ?	This field indicates if the Owner Code is also an Owner Group. An Owner Group is a grouping of one or more Owner Codes and can be used to roll-up results for reporting purposes OR for taking accountability for the DQ issue remediation.
Exclude from DQ Reports	Excludes the applicable Owner Codes from the Data Quality reports
Exclude from Pull-Thrus	Excludes the applicable Owner Codes from the Data Quality extracts defined within the Pull-Through queries.

5.1.16 DQ Metric Targets

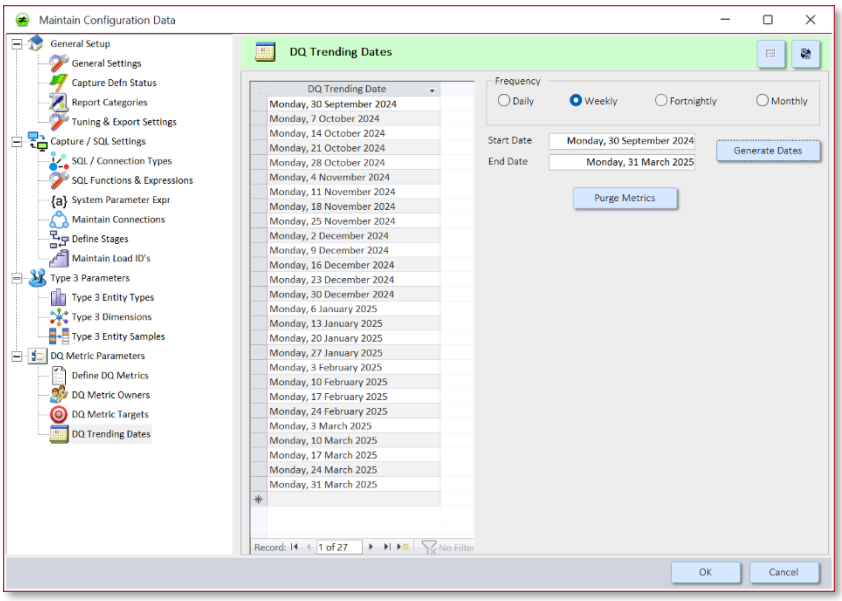
This form is used to define the Data Quality Target levels applicable to a particular DQ Metric that need to be achieved by the nominated Target Date.



Field	Description
Drilldown 1 to 4	If required, targets can be set up to 4 drilldown levels below the Owner level. This will depend upon whether additional Group-By Expressions were defined within the corresponding DQ Metrics Capture Definition.
Target Value	The Target Value represents the number of records with a DQ issue and will generally set to zero, however, the business may accept a small quality of issues to remain unresolved for the final data migration.
Target Date	The date by which the Target Value needs to be achieved by. E.g., The target date may be set to the week before the final Dry Run or Dress Rehearsal to ensure the data migration is successful.

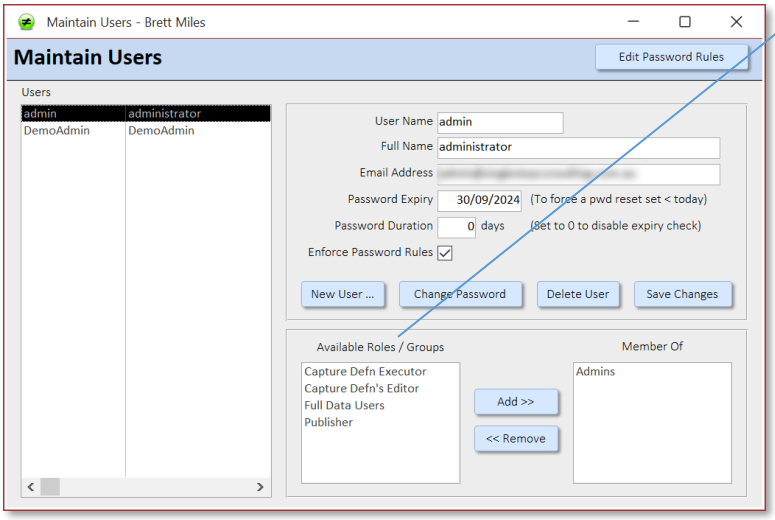
5.1.17 DQ Trending Dates

This form is used to define the DQ reporting dates that the DQ Metrics captures will be executed and reported against, showing the trend of the data quality improvement initiatives over time and forecasting whether the Target Value will be achieved by the Target Date based on the average clearance rate.

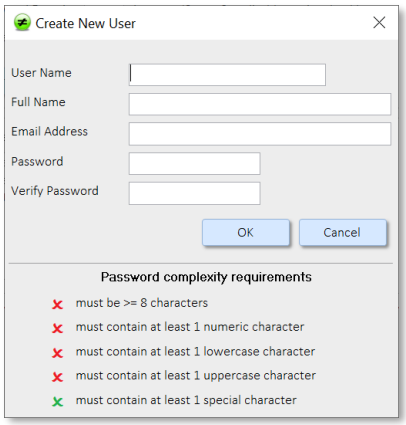


5.2 Adding Users

While the initial administrator account was created during the project database creation process, we can use the "Maintain Users" form (under the Administration Menu) to define additional users and assign the applicable Roles.



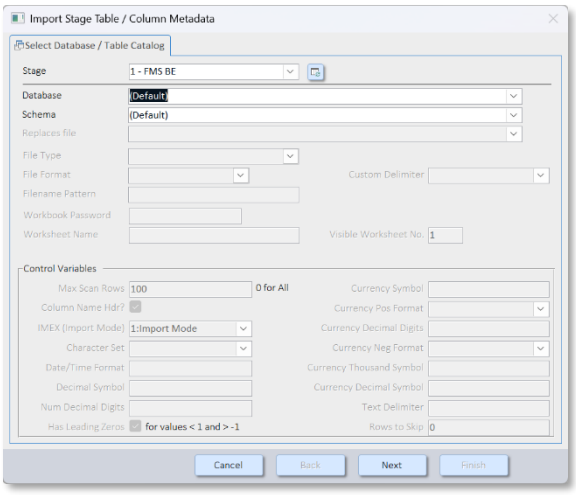
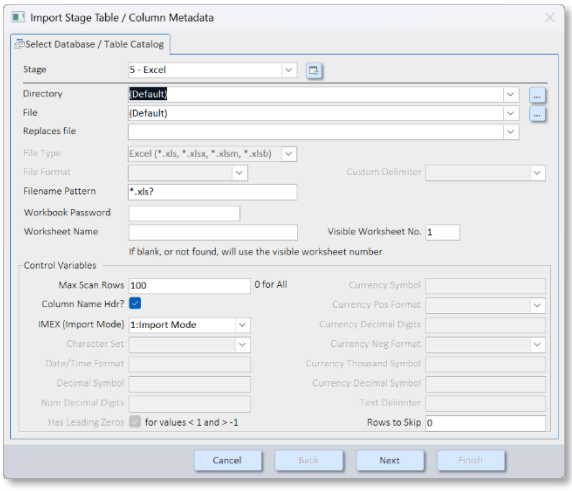
After creating a new user, select the applicable role(s) to be assigned to the user and click the "Add >>" button.



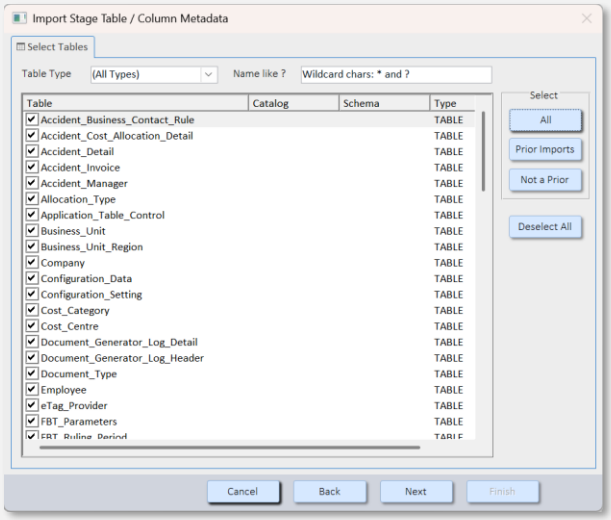
5.3 Importing Stage Table Metadata

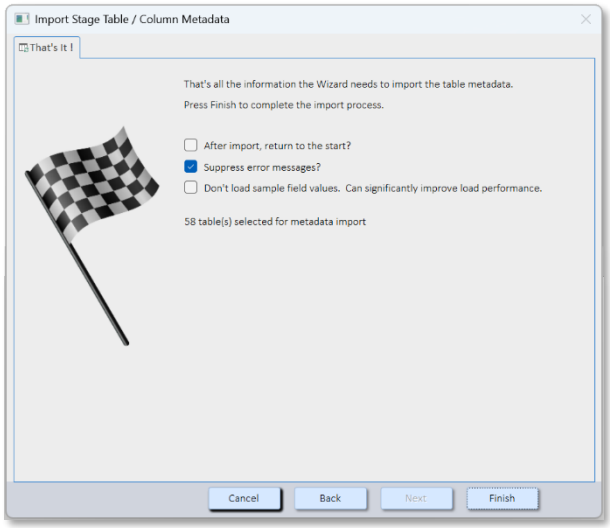
While optional, importing the table, column and index metadata from the applicable Stage databases / files can assist in fast-tracking the Capture Definition process (by used bulk creation) and reduce development times by reducing table and field level naming errors.

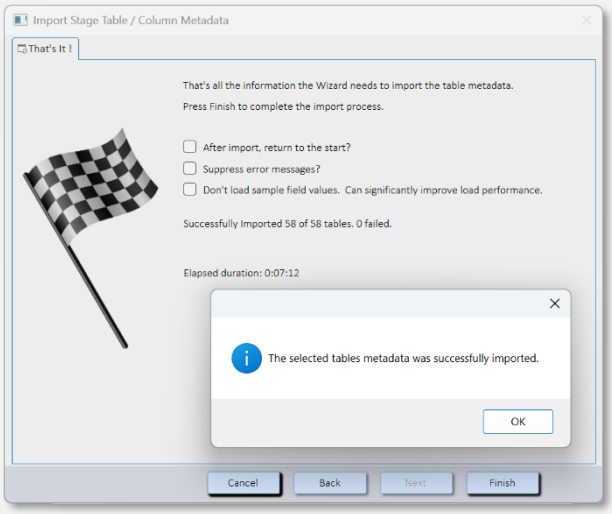
If the table metadata changes after the initial import was performed, it can simply be re-imported to retrieve the latest up-to-date information.

No.	Instructions	Screenshot
1	• Select the applicable Stage, Database and Schema from which the Table meta data is to be imported and Click Next. Note: Not all data sources will have a Database or a Schema. For example, an Ms Access database is standalone and does not contain multiple internal databases or Schema. In contrast, Oracle or SQL Server based sources do support multiple internal databases and schema.	Non-file-based data Stage  File-based (Text / Excel) data Stage 

No.	Instructions	Screenshot

2	• Depending upon the type of data source, there may be several "Table Types" available for import. Use the Table Type dropdown to filter the list of objects by Type.
	• You can filter the list by entering a pattern in the 'Name like ?' textbox.
	• Select one or more tables from the list or click:
	  ○ Select "All" to select all objects in the list
	  ○ "Prior Imports" to select any objects that have previously had their metadata imported.
	  ○ "Not a Prior" to select any objects that have not previously had their metadata imported.
	• Click Next to continue
		

3	• If you want to import Table Metadata from more Stages, enable the "After import, return to the start?" checkbox. • When ready, click "Finish" to being the import process.	

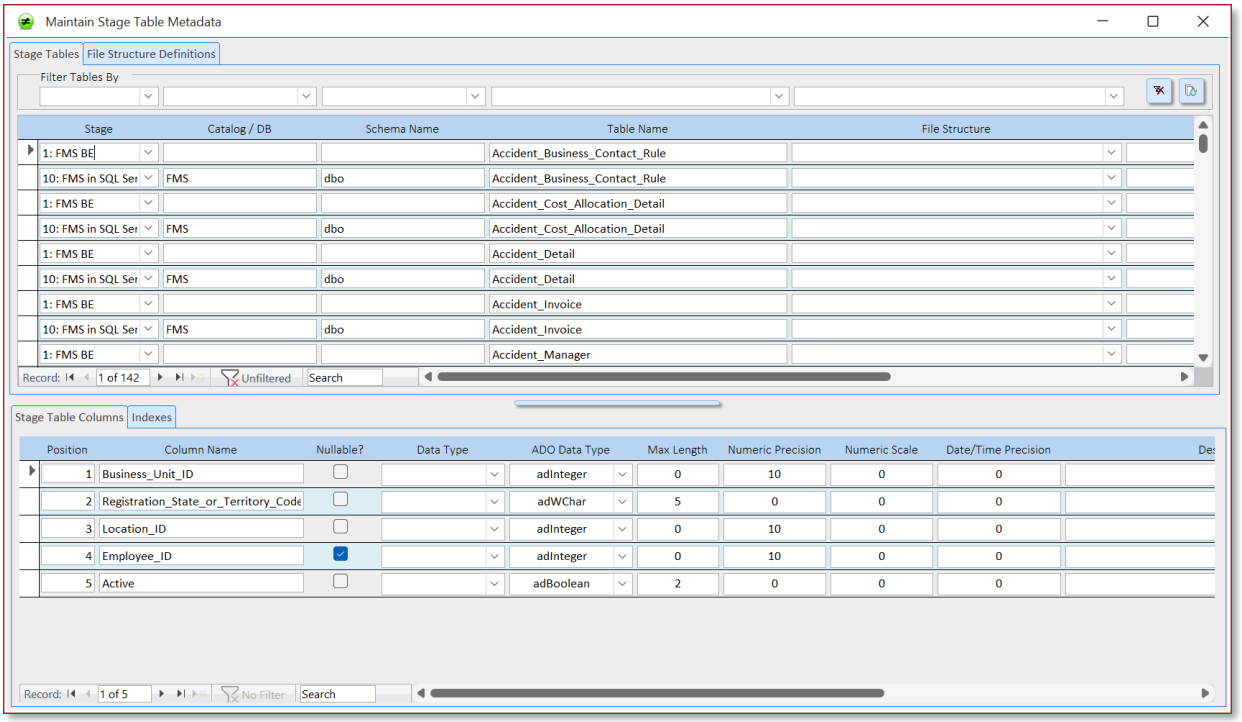
No.	Instructions	Screenshot
4	• The form will show the progressive number of tables imported as well as the Elapsed and Estimated time remaining to complete the import. • At the end of the import process, you should receive a popup message advising that the import was successful.	

5.4 Maintaining Stage Table Metadata

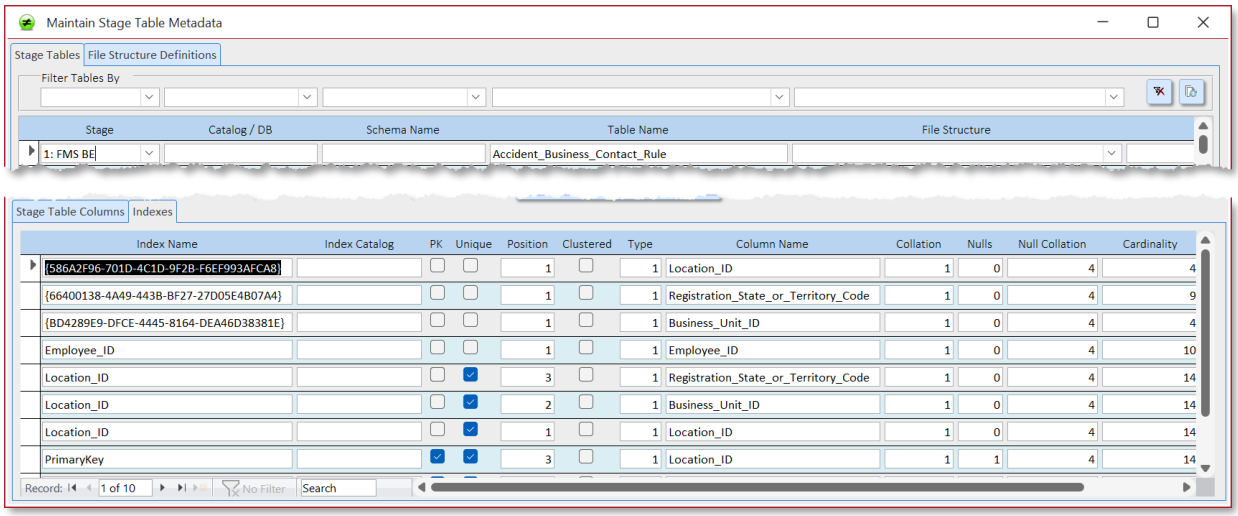
Manually maintaining the Stage Table Metadata is particularly useful for:

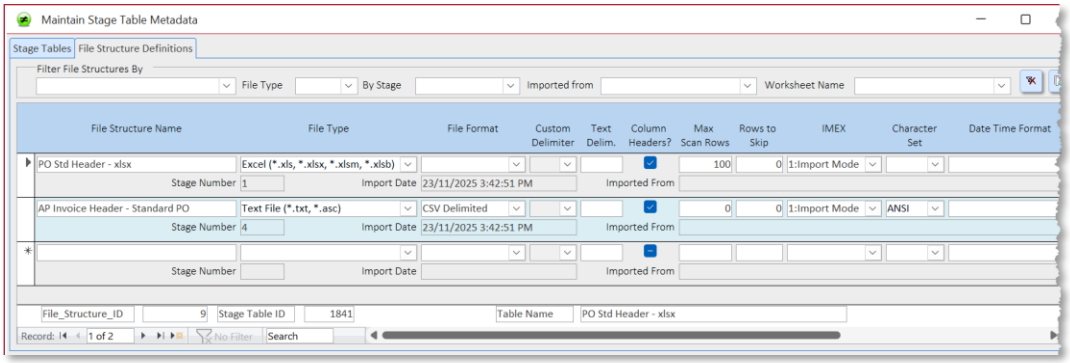
- file-based data sources (i.e., Excel and text files) where the correct data types may not be retrieved via the applicable driver,
- Minor adjustments to previously imported metadata

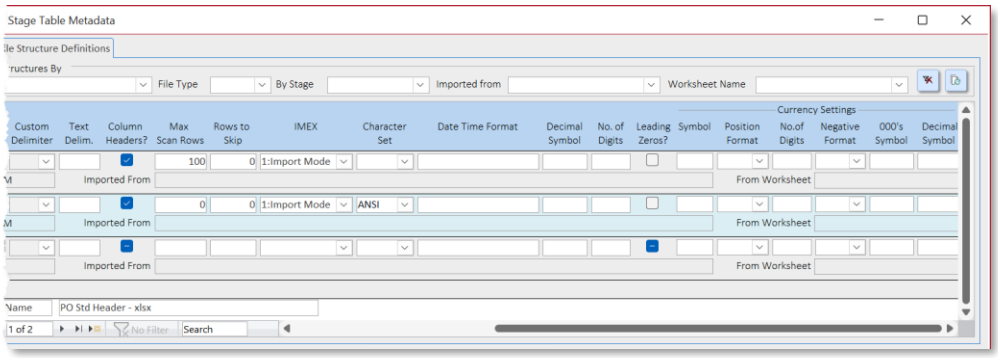
Column view



Index view
File Structure Definitions view







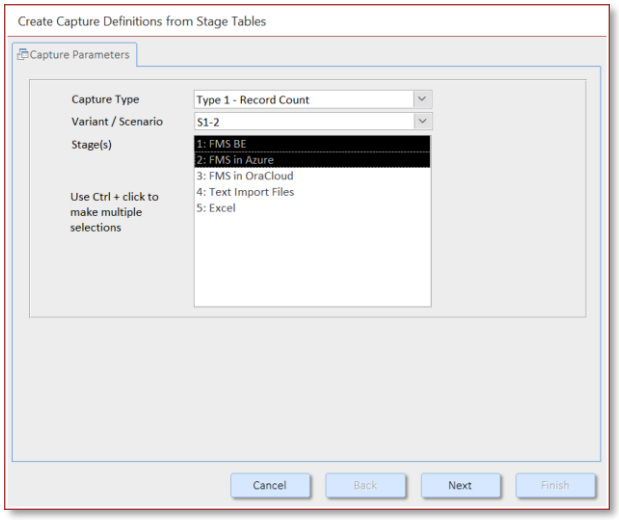
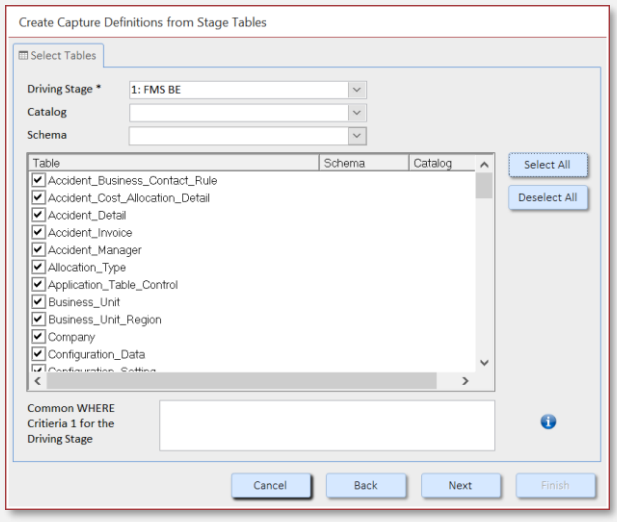
The File Structure Definitions view shows the parameters used to read the Text or Excel files and import the column metadata (as per the Column View above). Additional information includes when the import occurred, the file that the metadata was imported from, and in the case of Excel files, the Worksheet Name that contained the data that was analysed.

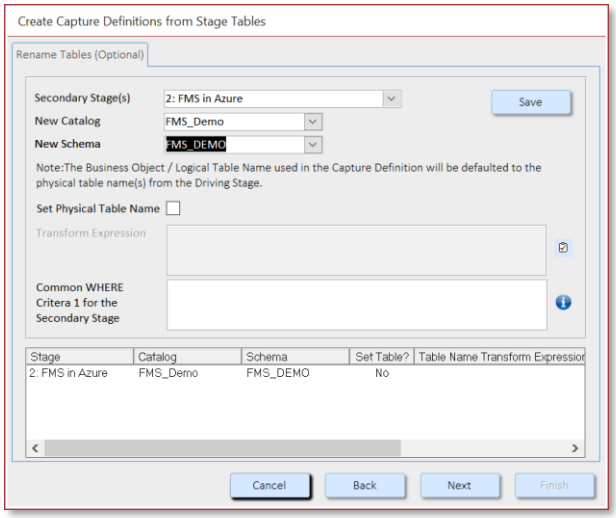
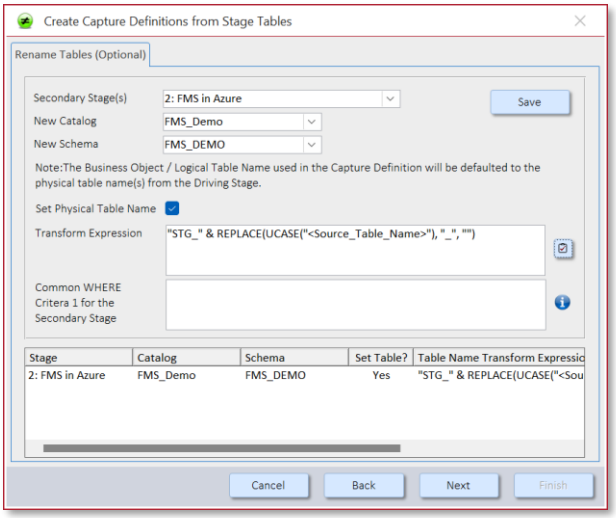
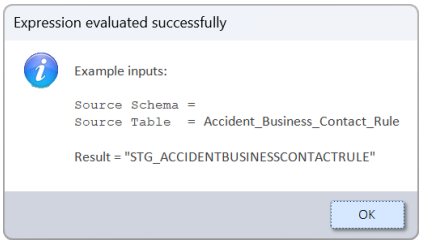
Note: If you have multiple source files of the same structure you only want /need to import the metadata from ONE of the source files. The resulting "File Structure" can be associated with the other files, usually by doing a UNION / UNION ALL query in the relevant Capture Definition. For more details, refer to Section 5.7 "Working with File-Based Data Sources / Stages".

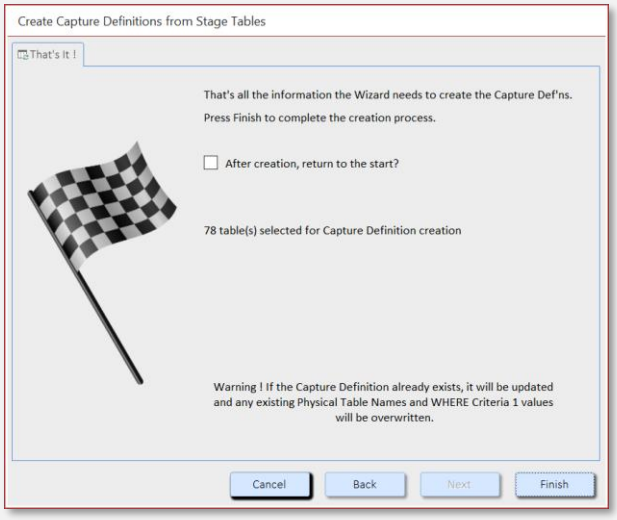
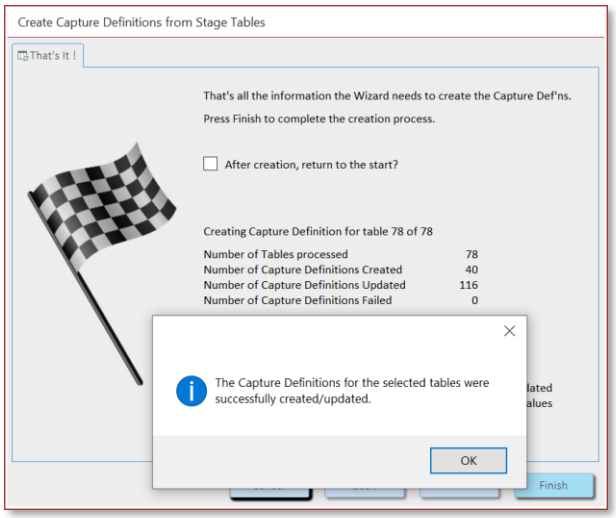
5.5 Bulk Generating Capture Definitions

Skeleton Capture Definitions can be bulk created from imported Table Metadata for the following types of Capture Definitions:

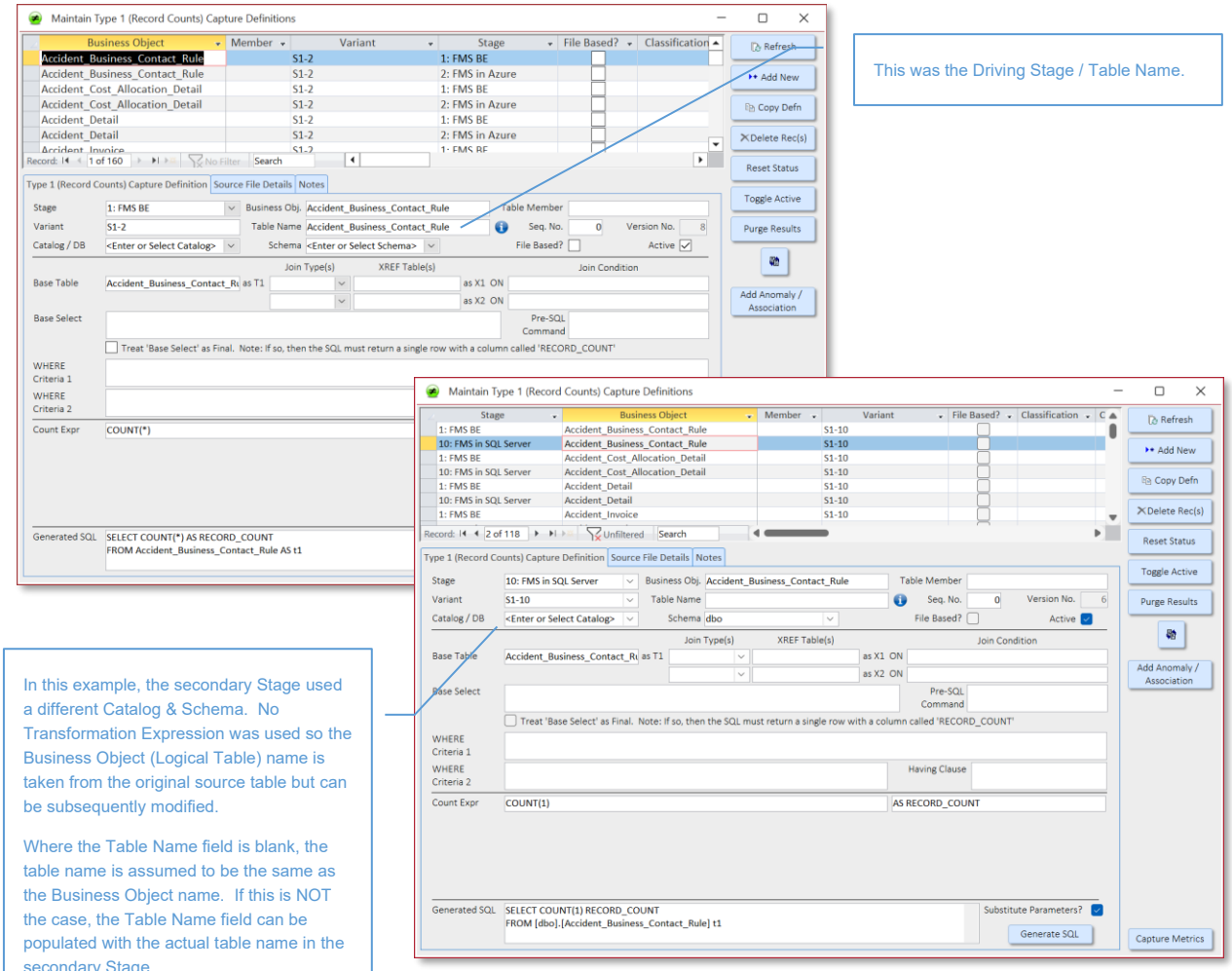
- Type 1 – Record Counts
- Type 2 – Group-By Aggregates, and
- Processing Statistics

No.	Instructions	Screenshot
1	• Select the relevant Capture Type. • Select an existing Variant, or type in a new Variant label. • Select one or more Stages for which to bulk create the skeleton Capture Definitions. • Click Next	
2	• Select the Driving Stage plus the Catalog and Schema (if applicable) to display the list of available tables. • Select the desired tables for which Capture Definitions should be generated. • Select Next An optional common WHERE condition can be defined that will be added to every capture definition on the Driving Stage side. Clicking the  icon will open the “Table Column & Index Metadata” form, which can assist in identify the correct name of a common column to use in the WHERE condition.	

No.	Instructions	Screenshot
3	For each secondary Stage: • Select the Stage • If the table in the Secondary Stage will be in a different Schema to the table from the Driving Stage, enter the New Schema name • If the table name in the secondary Stage is different to name in the Driving Stage, AND if the secondary stage table name can be derived from any combination of: o the source schema name, o the source table name o a static prefix and/or suffix o the application of common transformation functions as such replacing / removing certain characters, changing the case (upper or lower) etc., then this can be accomplished ticking the “Set Physical Table Name” checkbox and adding an appropriate Transformation Expression. The expression can be validated by clicking on the icon. Note: Right-clicking in the Transformation Expression text box will display a popup menu below. Insert Substitution Parameter Insert Code Snippet Save As New Code Snippet Common Functions Zoom The popup menu provides access to the substitution parameters <Source_Schema_Name> and <Source_Table_Name>, as well as Common VBA functions, however, ANY appropriate VBA function can be used. The Transformation Expression can be saved as a “Code Snippet” for future re-use. • Click the "Save" button to save the updates. • Repeat the above process for all secondary Stages. • Once all the updates have been saved, click Next.	  As an example, the following Transformation Expression, "STG_" & REPLACE(UCASE("<Source_Table_Name>"), "_", "") … will take the name of the source table, convert it to uppercase (using the UCASE function), remove any underscores from the name (using the REPLACE function) and then prefix the target table name with the static text “STG_”. 

No.	Instructions	Screenshot
4	• If you want to generate Capture Definitions for more Stages, enable the "After creation, return to the start?" checkbox. • When ready, click "Finish" to being the generation process.	
5	• The form will show the progressive number of tables processed. • At the end of the import process, you should receive a popup message advising that the generation process was successful.	

After the process has completed, the results can be viewed in the relevant Capture Definition maintenance form. In the example above, here are the results for one of the source tables:



This was the Driving Stage / Table Name.

In this example, the secondary Stage used a different Catalog & Schema. No Transformation Expression was used so the Business Object (Logical Table) name is taken from the original source table but can be subsequently modified.

Where the Table Name field is blank, the table name is assumed to be the same as the Business Object name. If this is NOT the case, the Table Name field can be populated with the actual table name in the secondary Stage.

5.6 Creating / Maintaining Capture Definitions

5.6.1 Structure of a Capture Definitions form

While each type of Capture Definition form has its own specific features based on the reconciliation type, they all share the same basic structure. In this manual, we will explore the Type 2 – Group-By Aggregates form and then note any key points of difference for the other types.

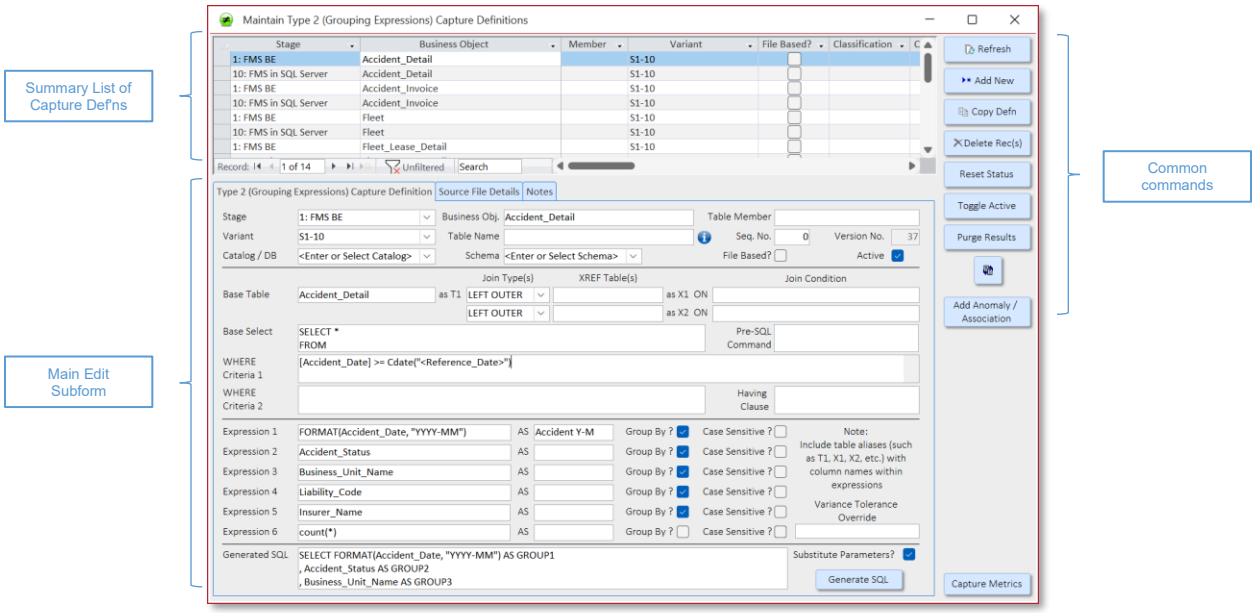


Figure 9 - Anatomy of a Maintain Capture Definitions form

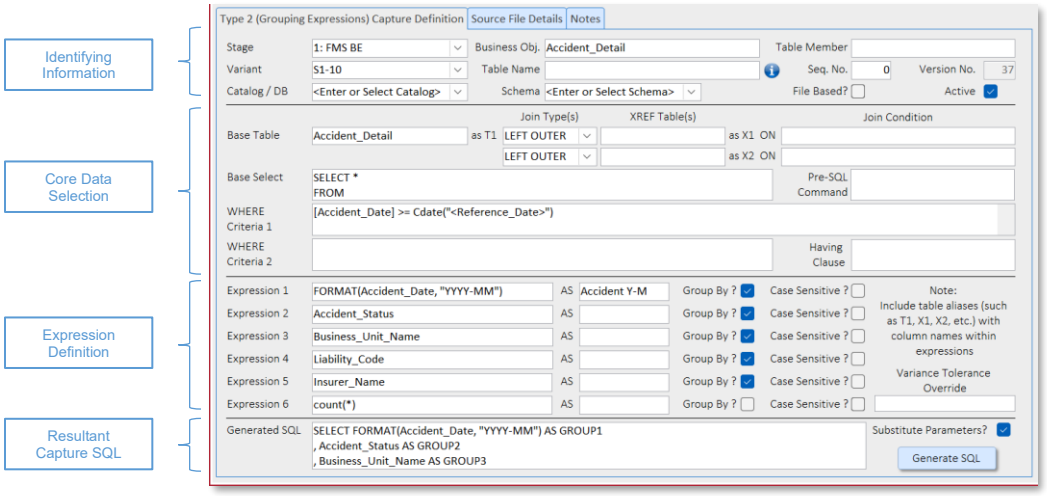
Summary List of Capture Definitions

This section of the form provides a list of the existing Capture Definitions. The list can be sorted and/or filtered as required and provides details such as:

- Definition Status – custom status to indicate the development status of the definition. E.g., New, Under Development, To be Tested, Ready, etc.
- Active Status – only capture definitions flagged as Active will be executed.
- Details of the most recent (active Load ID) execution including:
  - processing status e.g., Null or U = Unprocessed, F = Failed, P = Processed, M = Mixed (applicable to Type 3 capture results)
  - Most recent error message, if applicable
  - When the execution started
  - Record Count

Main Edit Subform

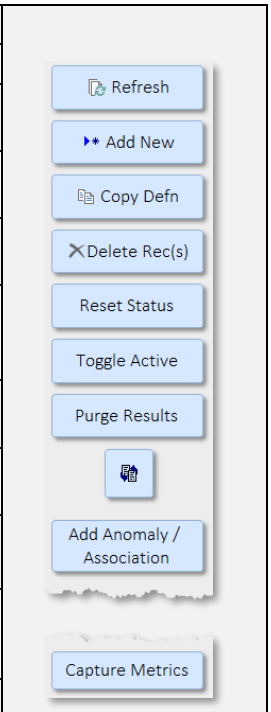
This is where the core details of the capture definitions are entered / defined and is comprised of four (4) key sections.



1. **Identifying information:** This section contains the fields that uniquely identify the capture definition, namely Capture Type, Stage, Business Object, Table Member and Variant. Also includes details of the table Schema and physical Table Name as well as the version number of the definition.
2. **Core Data Selection:** This section defines the core SQL selection statement together with the relevant filtering conditions. For simply SQL statements comprising 1 to 3 tables, the Join Types, Cross-Reference (XRef) Table(s) and Join Condition fields can be used to generate the core data selection SQL. Alternatively, for complex extraction logic, the Base Select field can store any custom-made SQL.
3. **Expression Definition:** This section contains details for the various group-by and aggregate expressions i.e., the metrics to be captured.
4. **Resultant Capture SQL:** This section displays the SQL generated from the information in the 2 previous sections. The "Substitute Parameters?" checkbox can be enabled to view the SQL after any parameter values have been substituted in the SQL.

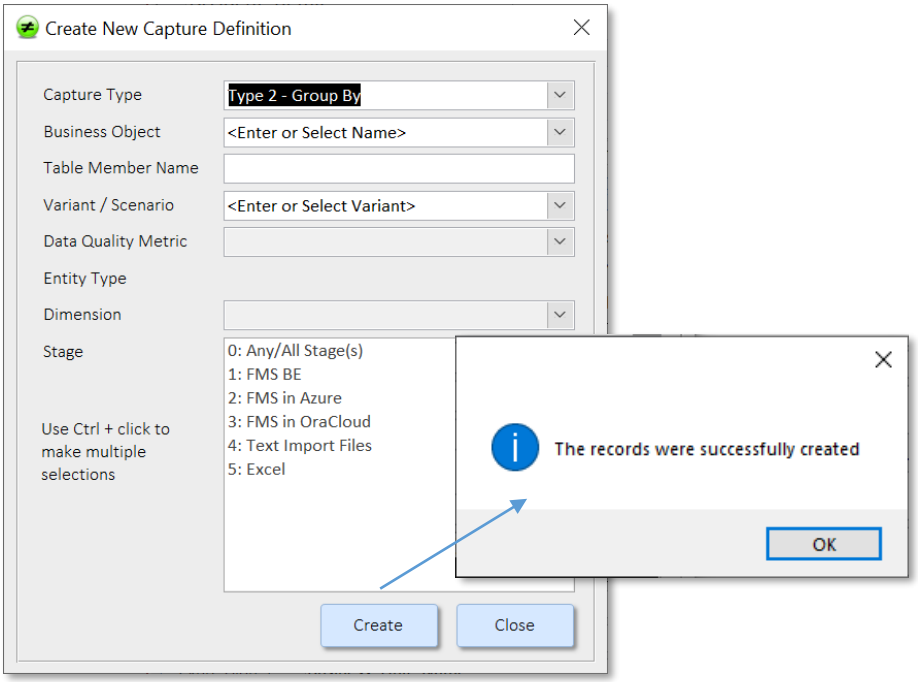
Common Commands

Button	Function
Refresh	Re-queries the form to reflect the latest data
Add New	Launches a popup form to create a new Capture Definition across one or more Stages
Copy Defn	User to clone an existing capture definition allowing re-use for similar capture definitions. E.g., similar expressions across different Stages.
Delete Rec(s)	Allows the deletion of any capture definitions selected from the Summary Information List
Reset Status	Clicking this button will set the processing status of any records in the Summary List to a state of Unprocessed (null or U). If the Summary List has been filtered, only the filter records will be updated
Toggle Active	This will toggle the Active Status of the records in the Summary List i.e., if the status was Active, it will be set to Inactive, and vice versa
Purge Results	This will delete any previously captured details results (applicable to Type 2, Type 3, Processing Statistics, DQ Metrics).
Edit field in Text Editor	Copies the contents of the currently selected field into the external text editor.
Add Anomaly / Association	Launches a form to create a new reconciliation anomaly record (or select a previously defined one) and associates it with the current Capture Definition.
Capture Metrics	Launches the Start Metrics Capture (Ad-hoc) form with the currently selected Capture Definition selected



5.6.2 Adding a New Capture Definition

The form is used to create a new Capture Definition across one or more Stages. Once they have been created, they can be edited to add the core selection SQL and metric expressions.



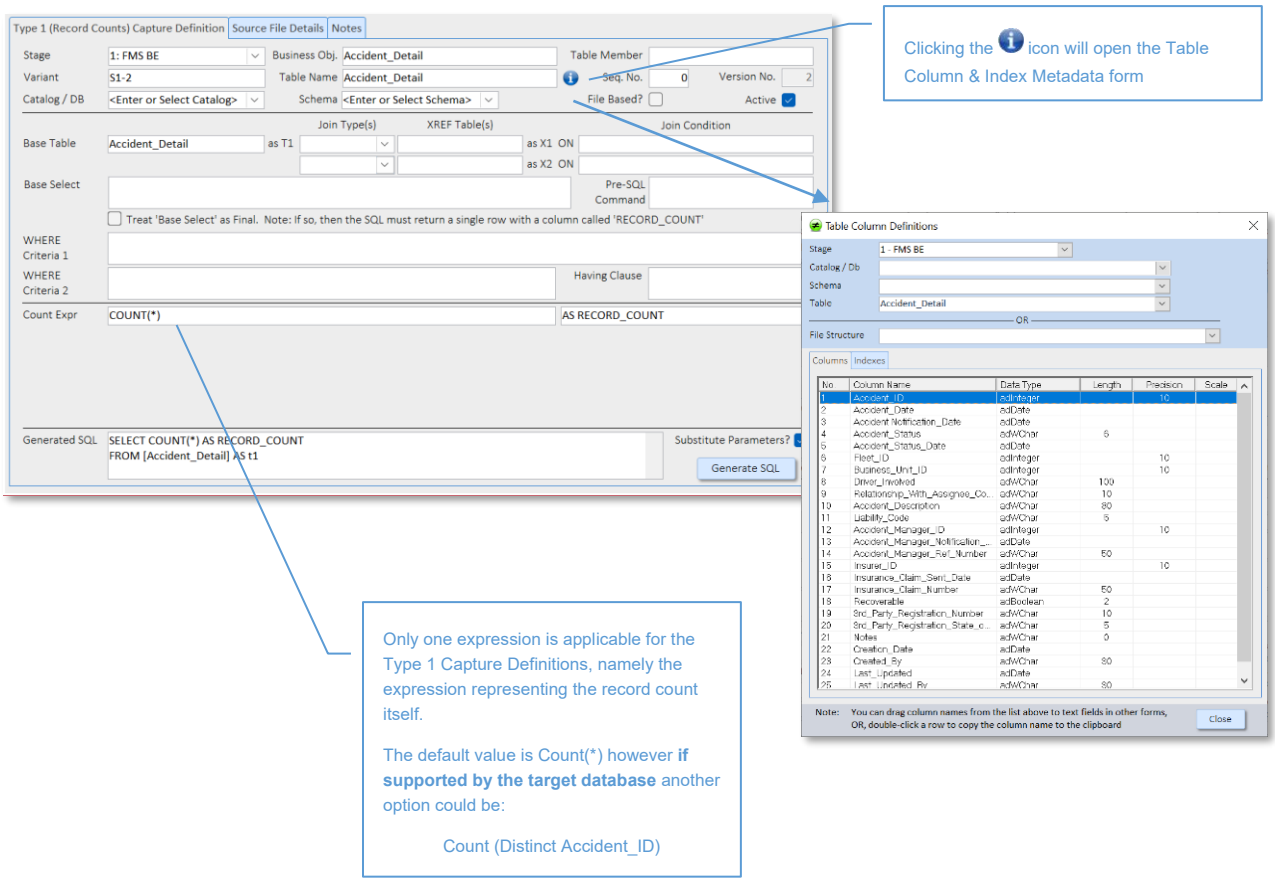
5.6.3 Defining the Capture Definition details

The essence of a Capture Definition is the data selection logic together with the capture metrics in the form of Grouping and Aggregate expressions.

The following sub-sections will illustrate the various aspects of defining Capture Definitions, using examples across each of the different types, highlighting both common and type-specific elements.

5.6.3.1 Type 1 – Record Counts

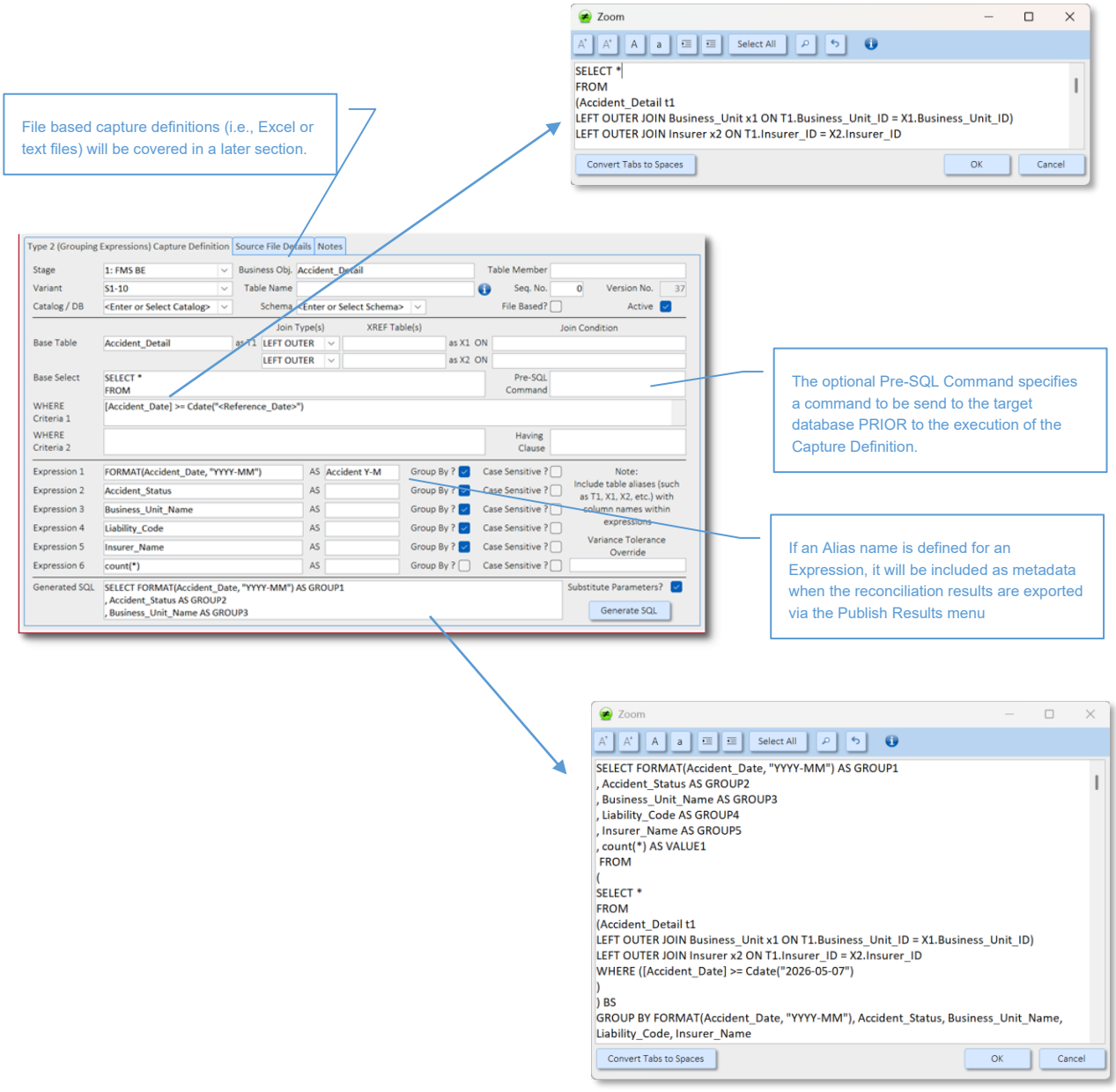
Type 1 (Record Counts) are the most straightforward of the various capture types. In its most basic form, to capture the record count from a single table with no WHERE conditions, the only input parameters required to generate the SQL statement, are the 'Table Name' and the 'Count Expr' field. If the Table Name is the same as the Business Object Name, then the 'Table Name' field can be left blank.



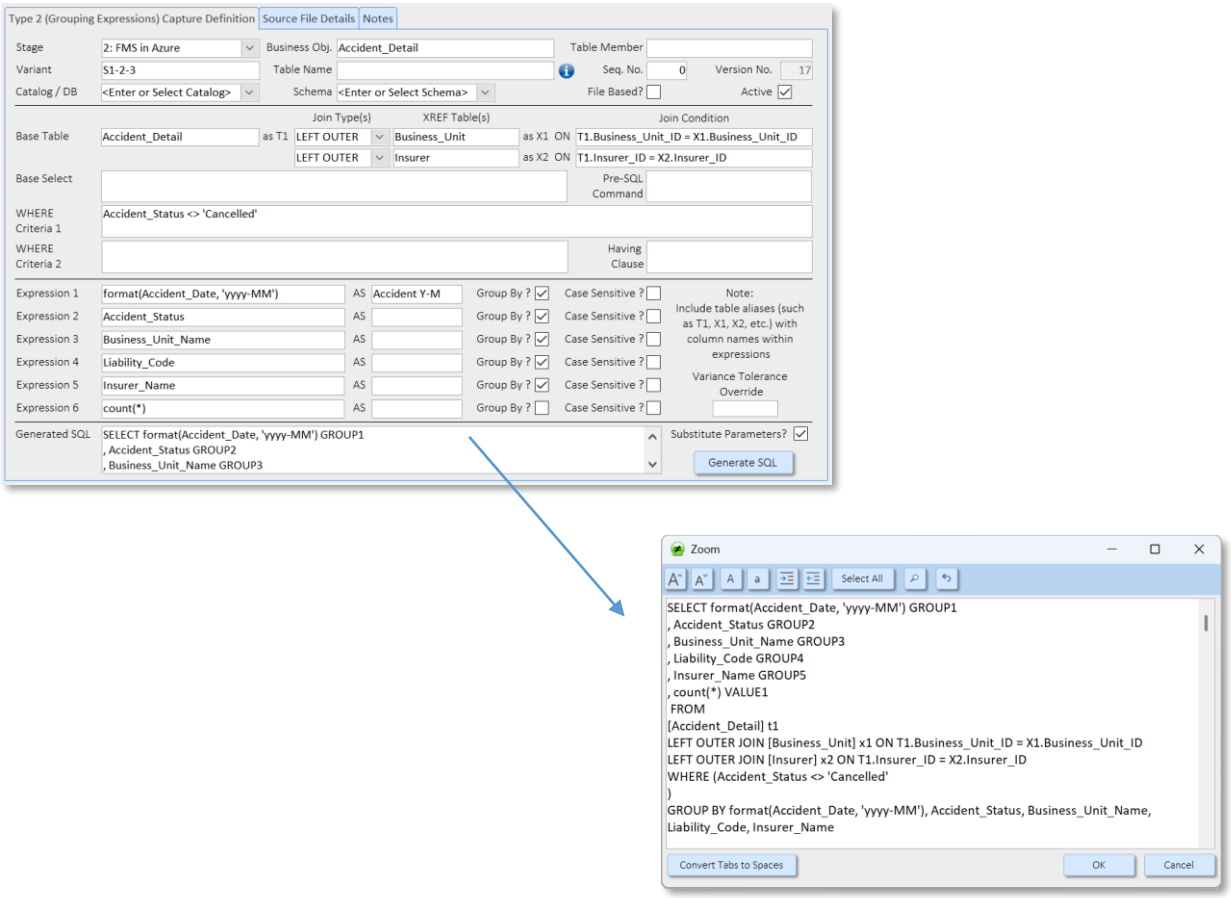
5.6.3.2 Type 2 – Grouping Expressions

Example 1. In this example, the "Base Select" field has been used to define the core SQL select statement, together with the "WHERE Criteria 1" field.

> While the WHERE criteria can be included in the Base Select statement, its' inclusion in the dedicated `WHERE Criteria n` fields help improve visibility without having to Zoom (via a double-mouse-click) into the Base Select field.



Example 2. In this example, the Join Type(s), Xref Table(s) and Join Condition fields are used to define the core SQL select statement, together with the "WHERE Criteria 1" field.

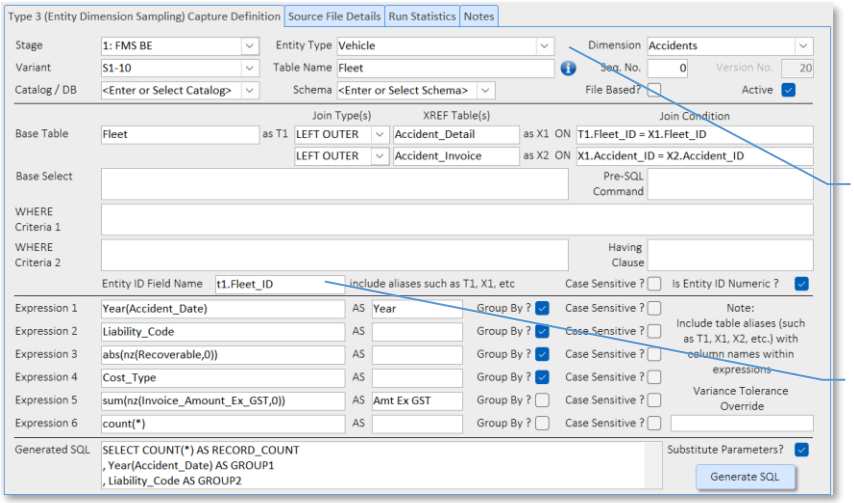


5.6.3.3 Type PS – Processing Statistics

Processing Statistics are described in section 2.2.4 and are equivalent to a one-sided Type 2 Grouping Expression. i.e. they aren't used for comparing data across 2 different stages but rather gathering statistics from a single stage.

5.6.3.4 Type 3 – Entity Dimension Sampling

For Type 3 Capture Definitions, very similar to Type 2 definitions, however the identifying field 'Business Object' has been replaced with the combination of 'Entity Type' and 'Dimension'. In addition, we need to specify which field represents the Entity ID and whether it is numeric (or text based).



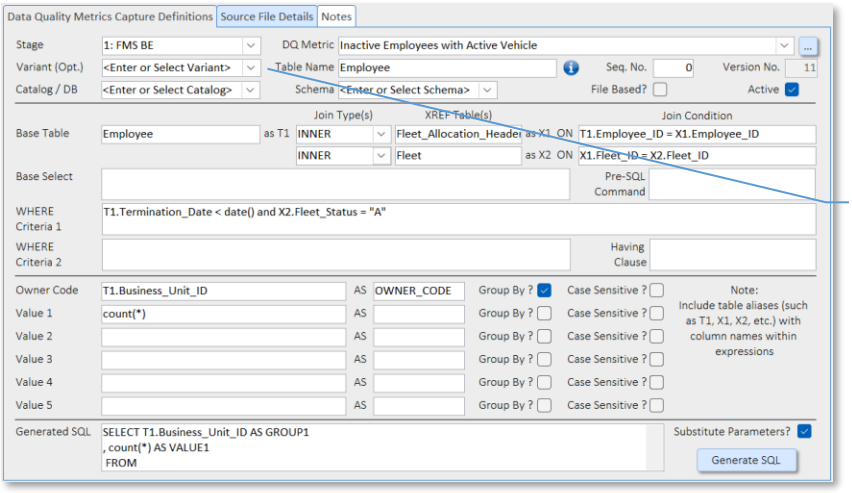
For Type 3 Capture definitions the Business Object field has been replaced with the combination of Entity Type and Dimension

Type 3 Capture definitions require the field corresponding to the Entity Id to be specified

5.6.3.5 Data Quality Capture Definitions

Unlike other types of Capture Definitions, the unique identifying fields for a Data Quality Capture Definition are the combination of Stage and the previously defined DQ Metric (Name) – refer section 5.1.14.

In addition, what would normally be Expression 1 in a Type 2 definition is explicitly dedicated to deriving a value used to represent the accountable DQ Owner.

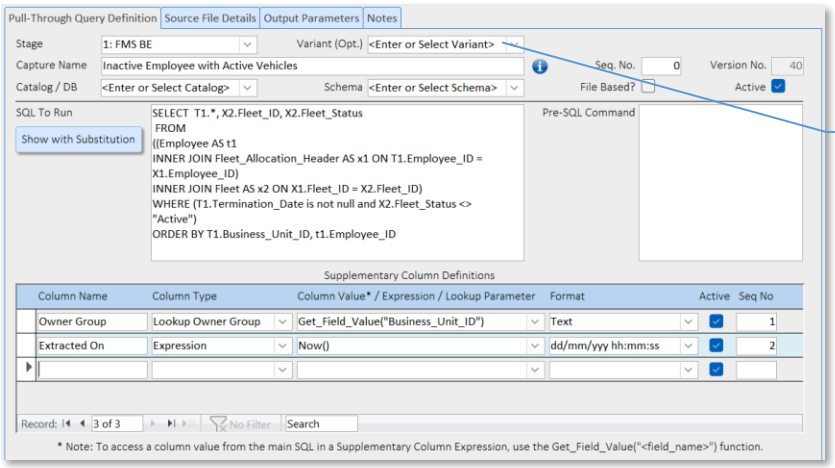


The Variant field is optionl for DQ Capture Definitions. It is provided to support filtering and segregation of Capture Query definitions.

5.6.3.6 Pull-Through Queries

Pull-Through Queries are significantly different from all other types of Capture Definitions. They are used to extract data from a **Stage** and write the data to one or more Excel files for subsequent review and analysis.

The unique identifier fields for a Pull-Through Query definition are the combination of Stage and Capture Name (free text).



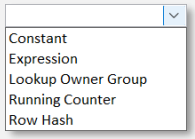
The Variant field is optional for Pull-Through Queries. It is provided to support filtering and segregation of Capture Query definitions.

Field	Description
SQL to Run	The final SQL to be executed against the Stage data source. Can include Substitution Parameters as required.
Pre-SQL Command	The optional Pre-SQL Command specifies a command to be send to the target database PRIOR to the execution of the Capture Definition.

Supplementary Column Definitions

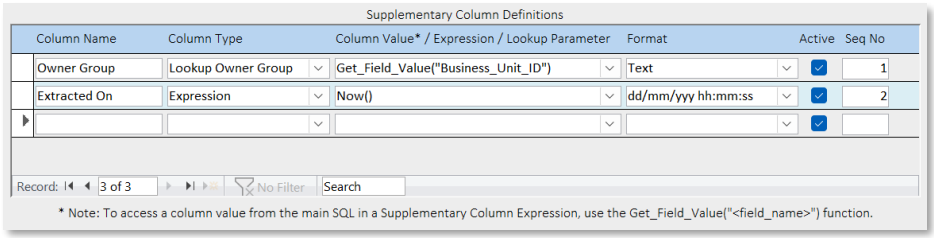
Supplementary columns represent data that is not available in the data extracted via the "SQL to Run" but which we want to include in the Excel extract itself. They will be added as additional columns after the last column from the extract itself.

Field	Description
Column Name	Self-explanatory
Column Type	The following column types are supported. • The syntax for **Expressions** is based on the Ms Access VBA expressions and inbuilt functions. In addition, the expression can use a custom function Get_Field_Value("<field_name>") to retrieve the field value for inclusion in further calculations. • The **Lookup Owner Group** type is used to lookup a field that is contained in the extract (representing an Owner Code) and then retrieving the corresponding Owner Group. Useful for subsequent filtering of the results. • The **Row Hash** type is used to calculate a hash value across all the fields returned by the 'SQL to Run' i.e. it excludes the supplementary columns. The user can select one of the supported hashing algorithms, via the downlist list presented in the 'Column Value / Expression / Lookup Parameter' field.
Column Value / Expression / Lookup Parameter	The derivation of the column value.



Field	Description
Format	The format syntax relates to the formatting codes used in Excel
Active	Only active supplementary columns will be included in the extract.
Seq No.	Controls the order that the supplementary columns will be added to the extract

Example supplementary column definitions.



For further details on Ms Access expressions and functions, refer to the following resources:

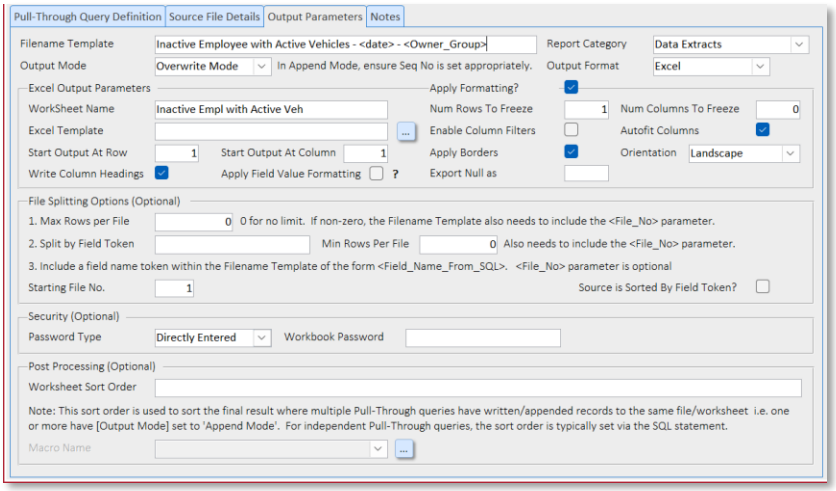
Expressions:

https://support.microsoft.com/en-us/topic/684295d0-97c2-443c-8e0a-8c6b635bec56#ID0EBBD=Creating_expressions

Access Functions (by category):

https://support.microsoft.com/en-us/office/access-functions-by-category-b8b136c3-2716-4d39-94a2-658ce330ed83

Output Parameters



Field	Description
Filename Template	Used as the template for the filename of the Excel file to be saved to. Can include Substitution Parameters in the name – e.g., <date> for the current extraction date.
Report Category	One of the user-defined categories setup during the initial configuration. Can be used as a Substitution Parameter within the Output Directory definitions to organize different types of extracts into different folders.

Field	Description
Output Mode	If set to 'Overwrite Mode' then the existing file will be overwritten. If set to 'Append Mode', then a second (or subsequent) extract will append its results to the end of the existing worksheet. The first capture definition in a set would generally be set to Overwrite Mode and then any subsequent captures can be set to Append Mode. If required, you can control the processing order by specifying appropriate Seq No values. Note: The Pull-Through capture definitions are processed in the order of Seq No, Stage Number and Capture Name
Output Format	Currently only Excel is supported

Excel Output Parameters

Field	Description
Worksheet Name	If not specified, Sheet 1 will be used
Excel Template	A pre-defined template that contains various formatting options to be used for the extract. Note: To create a template for a particular extract, generate an extract without a template, remove the data and format the header row (usually row 1) and the first data row (usually row 2) and then save it as an Excel Template file (*.xltx or *.xltm)
Max Rows per File	When used in conjunction with the <Record_Number_Band> (to be included in the Filename Template), setting this to a value > 0 will limit the number of rows added to an individual output file to the specified value, before diverting any remaining records to subsequent files.
Start Output at Row	For extracts without an Excel Template, this will usually be row 1. Where a template it used it is usually row 2 and we disable "Output Column Headings"
Start Out at Column	Usually set to column 1
Write Column Headings	If enabled, the first exported row will contain the column headings prior to the actual data. Usually disabled when using a template that has customized the headings
Apply Field Value Formatting	Enabling this setting activates an additional level of field value formatting / transformation. Examples include transforming a URL into a clickable hyperlink or displaying the description from a dropdown list rather than the raw underlying code. If the data set doesn't require the additional formatting, then disabling this can improve performance.
Apply Formatting?	Switch to quickly enable / disable the application of the cosmetic formatting options, including freezing columns / rows, enabling column filtering, autofitting columns, applying borders and setting the worksheet print orientation.
Num Rows to Freeze Num Columns to Freeze	Used to freeze the corresponding number of rows and columns in the Excel spreadsheet to assist when scrolling up and down through the sheet
Enable Column Filters	Applies column filters to each column in the table
Autofit Columns	Will automatically resize the columns to fit the data. Useful when no template is used.
Apply Borders	Applies borders to all the output data rows
Orientation	Formats the Excel worksheet as either Portrait or Landscape
Export Null as	Optional. If you want to be able to distinguish a null from a zero-length string in the output then you can specify what string to display e.g. Null or <Null> or N/A etc. Particularly useful when comparing the output of a Row Hash Value analysis

File Splitting Options

Pull-Through Queries can direct the output to a single Excel file or spread the data across multiple files based upon one of the following options:

- Option 1 – Maximum Rows per File. Once the current output file has reached the limit, it will be closed and a new output file created. This process continues until all source rows have been processed.
- Option 2 – Split by Field Token. The option will split the output into separate files based on the value of a nominated field. This option also supports accumulating more than 1 distinct token value per file, only switching to a new output file once the current file exceeds a minimum number of records.
- Option 3. This is an alternative method to split files by Field Token value but is limited to only one value per output file. This is achieved by including a substitution token representing the column name directly within the 'Filename Template' field. E.g. 'My Output File by <Department_Code>' will generate an output file for each and every distinct value of 'Department_Code' retrieved from the source.

Appendix D contains examples of various scenarios / use-cases where file-splitting can be used, together with examples on how to configure the options to achieve the desired outcome.

Field	Description
Max Rows per File	Only applies to Option 1. If set to a value > 0, a new output file will be created once the existing file contains the relevant number of records. To ensure each output file has a unique name, the substitution parameter <File_No> needs to be included in the 'Filename Template'.
Split by Field Token	Only applies to Option 2. When set to the name of a field / column from the 'SQL to Run', the exporter will monitor the **sorted** source data for changes in the value of the field from one record to the next. If the field value has changed AND the number of records in the current output file is >= the value of the 'Min Rows Per File' parameter, then the current file will be closed and a new output created. To ensure each output file has a unique name, the substitution parameter <File_No> needs to be included in the 'Filename Template' Note: When using this option, the 'SQL to Run' must ensure that the source is sorted by the relevant field and the confirmed by checking the corresponding checkbox 'Source is Sorted by Field Token?'
Min Rows Per File	Only applies to Option 2. Refer to usage above.
Starting File No.	Applies to both Options 1 and 2. Sets the starting file number of the next run. Can be used to support a level of restartability for long running Option 2 scenarios. E.g. if the exporting fails partway through the process, and a number of files (say 1 to 50) have been successfully generated, then by modifying the 'SQL to Run' to filter Field Token values greater than the values contained in last file (e.g. # 50) and then setting the 'Starting File No.' to 51, allows the process to continue from where it failed.

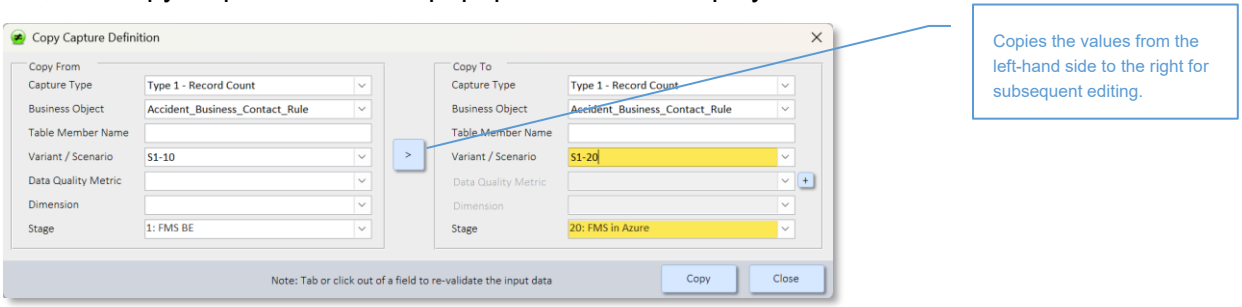
Field	Description
Source is Sorted by Field Token?	The applies to Option 2 (mandatory) and Option 3 (optional). While Option 3 can support unsorted source data, it requires the exporter to keep all the output files open until the process completes. If there are a large number of distinct values in the <Field_Name_From_SQL> this can create a significant drain on memory, as well as a performance hit due to file switching. If the data is sorted, the exporter only keeps one file open at a time, thus minimising memory consumption and improving performance by eliminating the unnecessary file switching.

Security and Post-Processing options

Field	Description
Password Type	Either 'Directly Entered' or 'Parameterised. If using Parameterised, the password can be defined via a 'Load / Stage Substitution Parameter' – refer section 5.1.10.
Workbook Password	Optional: Can be used to password protect the output Excel workbook.
Worksheet Sort Order	If populated (in the format of a SQL Order By clause), the worksheets will be sorted accordingly. Would typically only be used in the **last** 'Append Mode' Pull-Through query within a set. For single / individual Pull-Through queries, the sort order is typically set via the 'SQL to Run' statement itself.
Macro Name	An optional macro to run after the data has been written to the Workbook / Worksheet. An example where this can be used is with supplied '**Compare Records.xlsm**' Excel template. This template includes a macro called '**ApplyConditionalFormatting**' that highlights where cells from one row match the values in the previous row. If the output includes a supplementary column called 'Row Hash' it will highlight where the hash value is a duplicate (meaning that 2 rows match across all the non-supplementary output fields).

5.6.4 Copying an existing Capture Definition

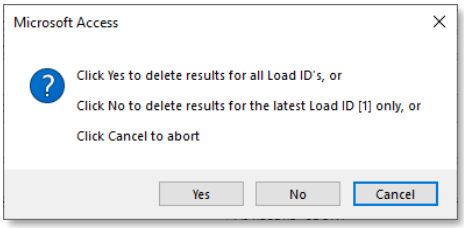
After clicking on the Capture Defn button, the user will be prompted to continue. If they click Yes, the Copy Capture Definition popup form will be displayed.



At least one of the identifying fields needs to be different before the Copy button is enabled. Clicking Copy will copy all the details from the original definition into the new definition.

5.6.5 Purging Capture Results

While not necessary, if for whatever reason you want to delete old capture results (for an individual Capture Definition) then this can be achieved by clicking on the Purge Results button. The user will be presented with the following dialog box for confirmation.

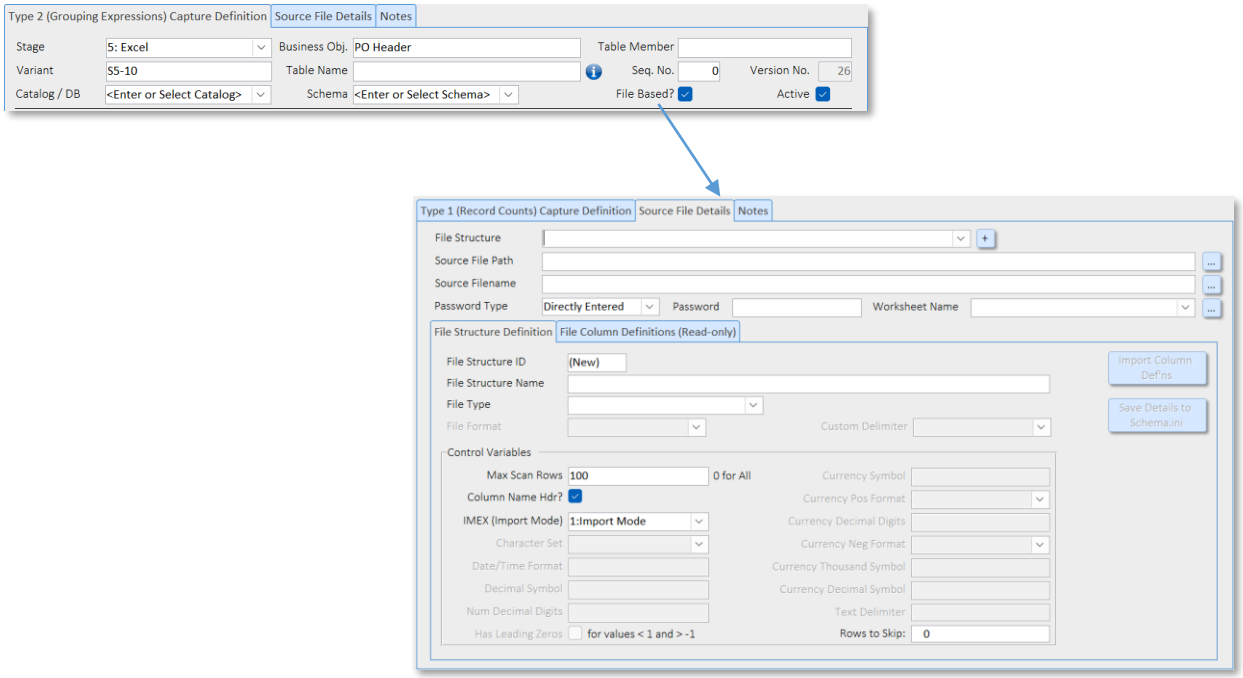


If the reason for purging old results is to reduce the database size of an Ms Access backend, the purging will need to be followed by running the "Compact BE Database" command under the Administration menu. Refer to Appendix C – “General Housekeeping”.

5.7 Working with File-Based Data Sources / Stages

When working with file-based data sources (i.e., Excel spreadsheets or flat text files – delimited, fixed width, etc.) there are several special considerations and constraints that need to be catered for.

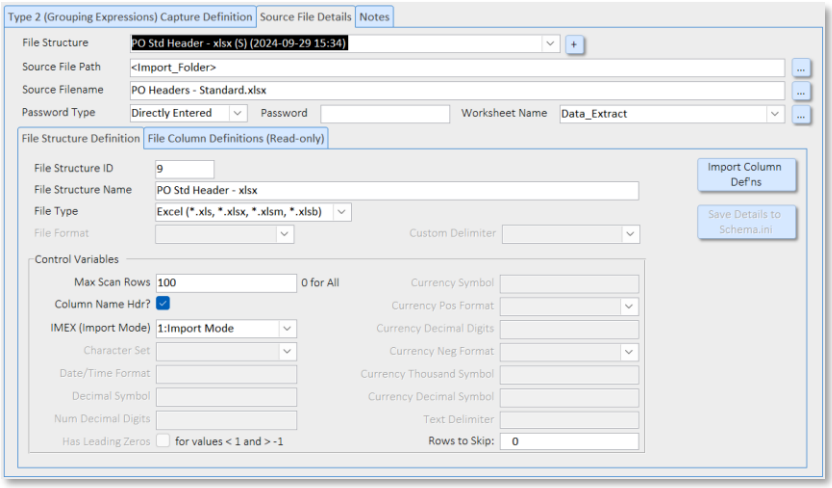
Setting the "File Based?" checkbox to true enables fields on the "Source File Details" tab.



Field	Description
File Structure	The dropdown presents the list of previously defined structures. One file structure may apply to multiple input files. To add a new structure definition, click on the + button next to the field. The dropdown includes the File Structure Name, the Stage it is associated with (optional) and the date the metadata was last imported.
Source File Path	Specifies the directory where the source file for the current Capture Definition is located. Can incorporate Substitution Parameters.
Source File Name	The name of the source file for the current Capture Definition
Password Type	Either 'Directly Entered' or 'Parameterised. If using Parameterised, the password can be defined via a 'Load / Stage Substitution Parameter' – refer section 5.1.10.
Password	Optional: Can be used to password protect the output Excel workbook.
Worksheet Name	The name of the sheet within the Excel file that we will be reading from. Clicking on the ... button will read the available worksheets within the file and populate the dropdown list for selection.

5.7.1 Excel based data sources

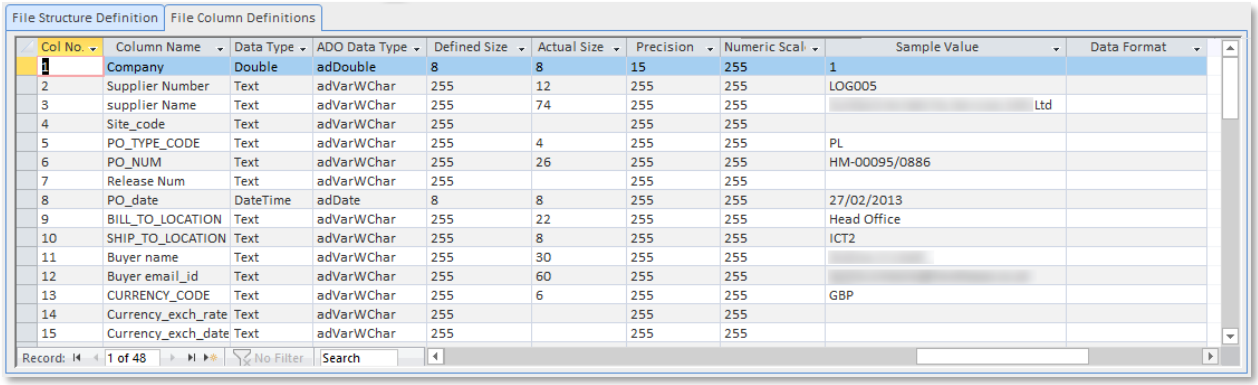
Step 1 – Specify the Control Variables



Field	Description
File Structure ID	Automatically assigned – not editable
File Structure Name	A descriptive name to refer to the file structure
File Type	Select "Excel"
Max Scan Rows	The number of rows of the data set that will be scanned to derive the data types for each of the columns
Column Name Hdr?	Set to true if the file contains column headers
IMEX (Import Mode)	Setting IMEX=1 tells the driver to use Import mode. This tells the driver to always read "intermixed" (numbers, dates, strings etc.) data columns as text.

Step 2 – Import the Column Definitions

After setting the Control Variables, clicking on the "Import Column Def'ns" button will scan the import file to read and derive the column definitions. These can then be manually revised if and as required.



Step 3 – Configuring the Capture Definition

Specifying the Capture Definition for an Excel-based data source is very similar to any other data source, however, there are some points of difference to be aware of.

1. The core structure of a SQL Statement for one or more Excel files is of the form: -

   ```
   SELECT * FROM [<WorksheetName>$] {optional AS <Alias>}
   {optional IN '<Enter_Full_Path_and_Filename>' 'Excel 12.0;'}
   ```

   In the following example, the first SELECT statement doesn't require the optional file details because they will be derived from the File Structure Definition details.

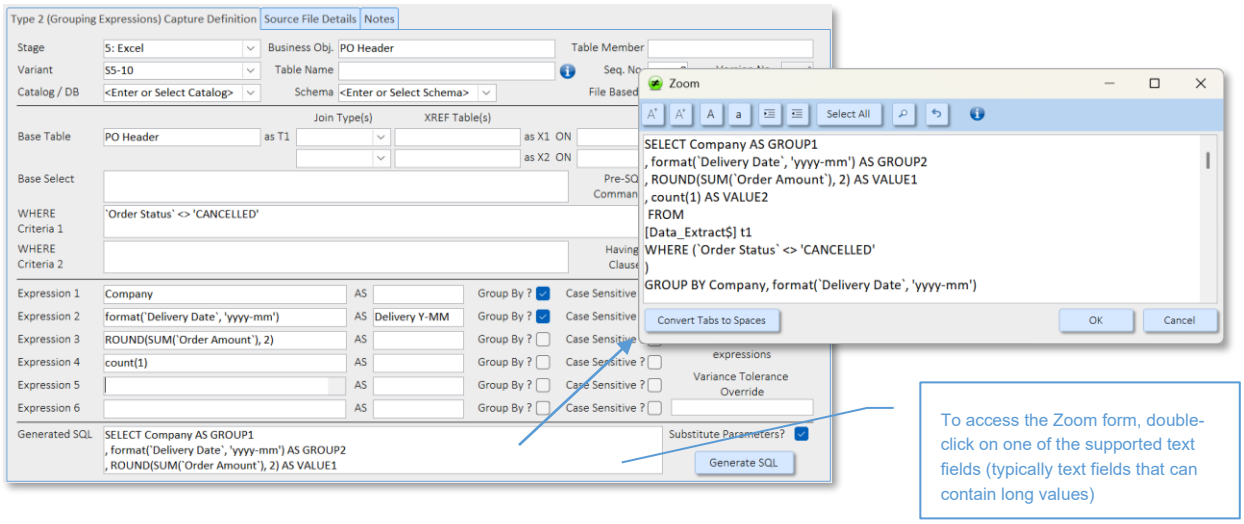
   #### Example 1 – Union Query

   ```
   SELECT * FROM [Sheet1$] AS t1
   UNION ALL
   SELECT * FROM [Data Sheet$] AS t2 IN 'D:\ImportDir\Customer.xlsx'
   'Excel 12.0;'
   ```

   #### Example 2 – Join Query with Where Condition(s)

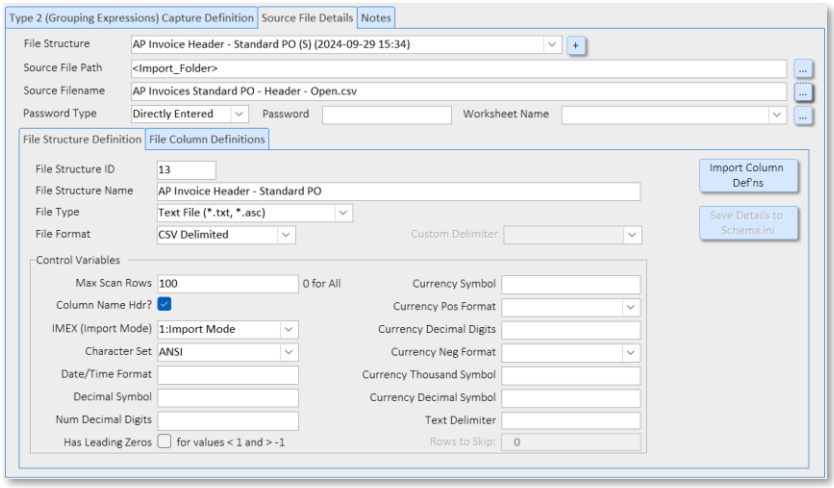
   ```
   SELECT * FROM [Sheet1$] AS t1
   INNER JOIN
   (SELECT * FROM [Data Sheet$] IN 'D:\ImportDir\Customer.xlsx' 'Excel
   12.0;') as t2
   ON t1.<join_field_name1> = t2.<join_field_Name2>
   WHERE <where_conditions>
   ```

2. The list of available scalar functions available for use in the SQL expressions depends upon the ODBC driver being used. Assuming you use the Provider=Microsoft.ACE.OLEDB.12.0 for the Excel Connection Type, the list of functions is similar to connecting to a Ms Access database. Check your specific driver / provider for details.



5.7.2 Text file data sources

Step 1 – Specify the Control Variables



Field	Description
File Structure ID	Automatically assigned – not editable
File Structure Name	A descriptive name to refer to the file structure
File Type	Select "Text File"
Format	Supported formats include: • CSV Delimited • Custom Delimited • Fixed Length • Tab Delimited
Custom Delimiter	Only applicable when the Format is set to Custom Delimited. Supports any single character except a double quotation mark (").

Control Parameters

Full details on the control parameters can be found online at https://docs.microsoft.com/en-us/office/client-developer/access/desktop-database-reference/initializing-the-text-data-source-driver#customizing-the-schemaini-file-for-text-and-html-data. ReQuon will automatically create a suitable schema.ini file based upon the information captured in the form.

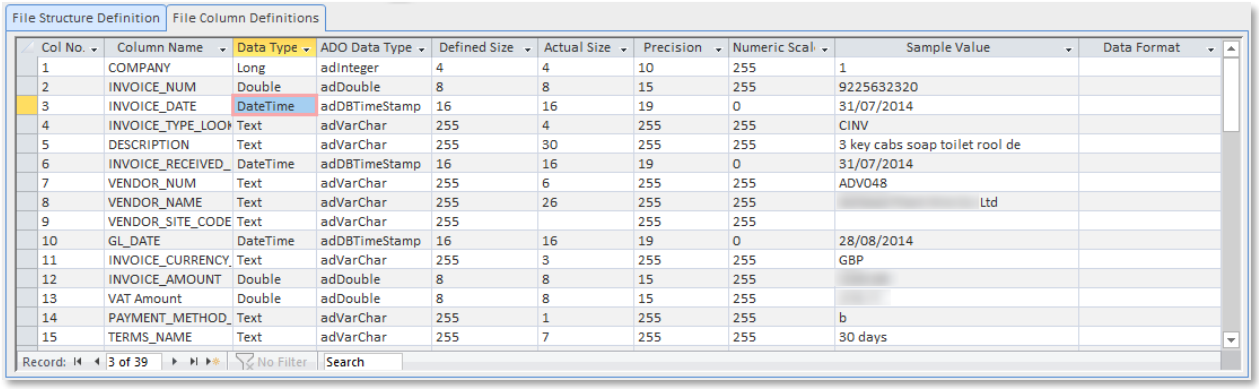
The table below provides a summary of the Control Variables.

Field	Description
Max Scan Rows	Indicates the number of rows to be scanned when guessing the column data types. If this is set to 0, the entire file is searched
Column Name Hdr?	Set to true if the file contains column headers
IMEX (Import Mode)	Setting IMEX=1 tells the driver to use Import mode. This tells the driver to always read "intermixed" (numbers, dates, strings etc.) data columns as text.
Character Set	Can be set to OEM, ANSI, UNICODE, or the decimal number of a valid code page, and indicates the character set of the source file.
Date Time Format	Can be set to a format string indicating dates and times. This entry should be specified if all date/time fields in the import/export are handled with the same format. All the Microsoft Jet database engine formats except AM and PM are supported. In the absence of a format string, the Windows Control Panel short date picture and time options are used.
Decimal Symbol	Can be set to any single character that is used to separate the integer from the fractional part of a number. If this entry is absent, the default value in the Windows Control Panel is used.
Num Decimal Digits	Indicates the number of decimal digits in the fractional portion of a number. If this entry is absent, the default value in the Windows Control Panel is used.
Has Leading Zeros	Specifies whether a decimal value less than 1 and greater than –1 should contain leading zeros; this value can either be False (no leading zeros) or True.
Currency Symbol	Indicates the currency symbol to be used for currency values in the text file. Examples include the dollar sign ($) and Dm. If this entry is absent, the default value in the Windows Control Panel is used.
Currency Pos Format	Can be set to any of the following values: • Currency symbol prefix with no separation ($1) • Currency symbol suffix with no separation (1$) • Currency symbol prefix with one character separation ($ 1) • Currency symbol suffix with one character separation (1 $) If this entry is absent, the default value in the Windows Control Panel is used.
Currency Decimal Digits	Specifies the number of digits used for the fractional part of a currency amount. If this entry is absent, the default value in the Windows Control Panel is used.
Currency Neg Format	Can be one of the following values: ($1), –$1, $–1, $1–, (1$), –1$, 1–$, 1$–, –1 $, –$ 1, 1 $–, $ 1–, $ –1, 1– $, ($ 1), (1 $) The dollar sign is shown for purposes of this example, but it should be replaced with the appropriate CurrencySymbol value in the actual program. If this entry is absent, the default value in the Windows Control Panel is used.
Currency Thousand Symbol	Indicates the single-character symbol to be used for separating currency values by thousands in the text file. If this entry is absent, the default value in the Windows Control Panel is used.
Currency Decimal Symbol	Can be set to any single character that is used to separate the whole from the fractional part of a currency amount. If this entry is absent, the default value in the Windows Control Panel is used.
Text Delimiter	Can be set to any single character that is used to delimit strings that contain any of the other special characters. E.g. "abc","xyz,pqr","hij". If this entry is not present the default delimiter is a double quote.
Rows to Skip	Used to skip a number of rows at the top of an Excel spreadsheet

Step 2 – Import (or manually define) the Column Definitions

After setting the Control Variables, if the File Format is **not set** to "Fixed Length", then the "Import Column Def'ns" button will be enabled and can be used to import column definitions. If the File Format **is set** to "Fixed Length" however, the user will need to manually enter the column definitions into the "File Column Definitions" tab.

In the imported column definitions below (from a csv file), some potential manual adjustments to the definitions might include changing the Data Type and ADO Data Type from a text related data type to Date or Timestamp related data type if the import didn't recognize the correct type.


File Structure Definition | File Column Definitions

Col No.	Column Name	Data Type	ADO Data Type	Defined Size	Actual Size	Precision	Numeric Scal	Sample Value	Data Format
1	COMPANY	Long	adInteger	4	4	10	255	1	
2	INVOICE_NUM	Double	adDouble	8	8	15	255	9225632320	
3	INVOICE_DATE	DateTime	adDBTimeStamp	16	16	19	0	31/07/2014	
4	INVOICE_TYPE_LOOK	Text	adVarChar	255	4	255	255	CINV	
5	DESCRIPTION	Text	adVarChar	255	30	255	255	3 key cabs soap toilet rool de	
6	INVOICE_RECEIVED_	DateTime	adDBTimeStamp	16	16	19	0	31/07/2014	
7	VENDOR_NUM	Text	adVarChar	255	6	255	255	ADV048	
8	VENDOR_NAME	Text	adVarChar	255	26	255	255	[redacted] Ltd	
9	VENDOR_SITE_CODE	Text	adVarChar	255		255	255		
10	GL_DATE	DateTime	adDBTimeStamp	16	16	19	0	28/08/2014	
11	INVOICE_CURRENCY_	Text	adVarChar	255	3	255	255	GBP	
12	INVOICE_AMOUNT	Double	adDouble	8	8	15	255	[redacted]	
13	VAT Amount	Double	adDouble	8	8	15	255	[redacted]	
14	PAYMENT_METHOD_	Text	adVarChar	255	1	255	255	b	
15	TERMS_NAME	Text	adVarChar	255	7	255	255	30 days	

Record: 3 of 39 | No Filter | Search

Step 3 – Save to the File Structure Definition to a Schema.ini file

Clicking the "Save Details to Schema.ini: button allows the File Structure Definition, including any associated File Column Definitions, to be written to a schema.ini file in the same directory as the file. The Schema.ini file is used by the Text Importer to determine how to interpret / convert the source file. It also provides the ability to save the same file structure details to a list of multiple files (selected in a subsequent step) where they have the same structure (as determined by the user)

This is useful where a UNION / UNION ALL, of multiple files will be used as the "Base Select" for a single capture definition, where only one file can be specified in "Source Filename" field.

Step 4 – Configuring the Capture Definition

Specifying the Capture Definition for a Text-file-based data source is very similar to any other data source, however, there are some points of difference to be aware of.

1. Assuming the use of the Microsoft Office 12.0 Access Database Engine OLE DB Provider in text mode, the core structure of a SQL Statement for one or more text files is of the form:

   ```
   SELECT * FROM `<Enter_Text_Filename_Without_Path>` {optional AS
   <Alias>}
   ```

   Note: If the filename doesn't contain spaces, the slanted quotes ` are not required.

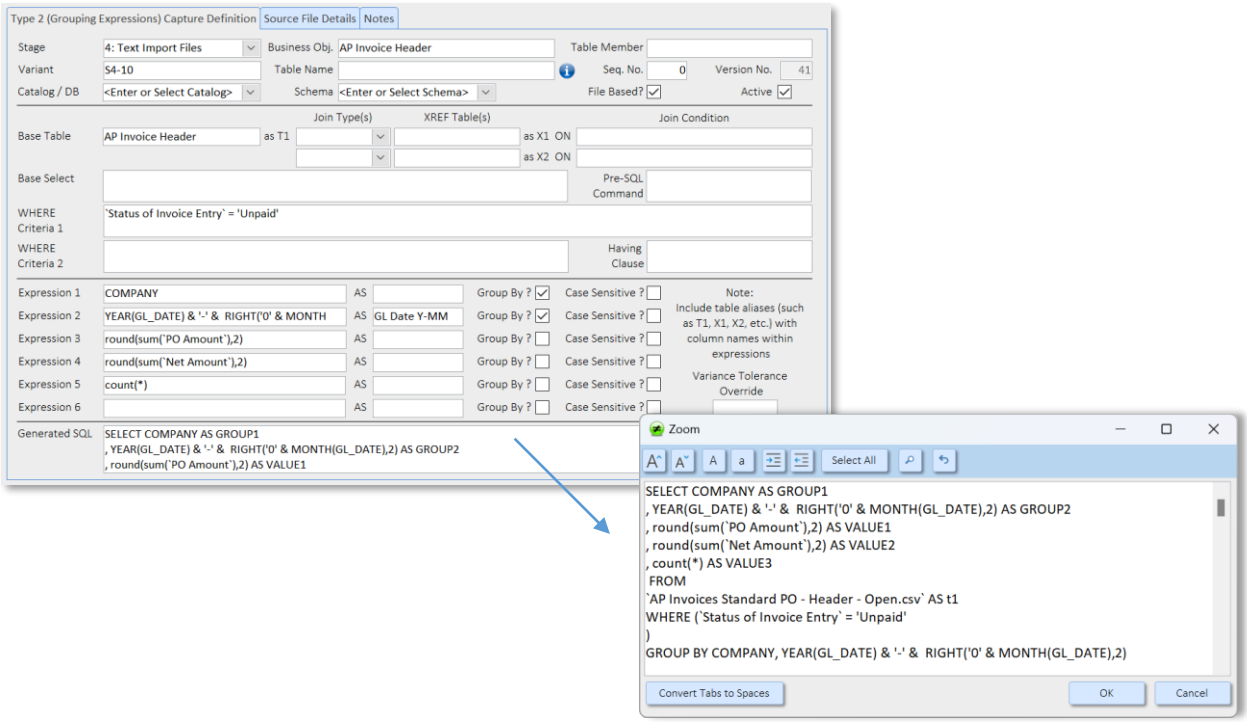
   Example 1. A Union of three files

   ```
   SELECT * FROM `InvoiceHeaders_PRD 2024-01.csv`
   UNION ALL
   SELECT * FROM `InvoiceHeaders_PRD 2024-02.csv`
   ```

2. The list of available scalar functions available for use in the SQL expressions depends upon the ODBC driver being used. Assuming you use the Driver=Microsoft Access Text Driver, the list of functions is similar to connecting to an Ms Access database, however, we have found some deviations. Check your specific driver / provider for details.

Warning: the format function available in Ms Access does not error when processing Text files but returns null when trying to format a date field to 'yyyy-mm'. In this scenario, the same result can be achieved with the follow expression:

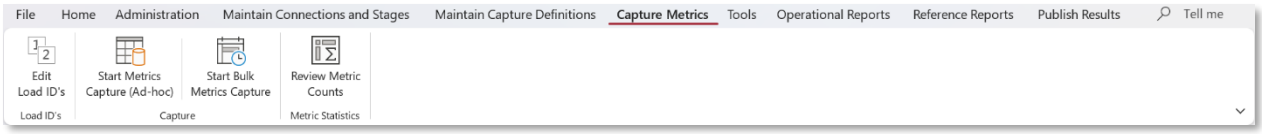
YEAR(<datefield>) & '-' & RIGHT('0' & MONTH(<datefield>),2)



5.8 Executing the Capture Definitions (Metrics Capture)

ReQuon provides to methods to execute the Capture Definitions and retrieve and store the results for subsequent analysis:

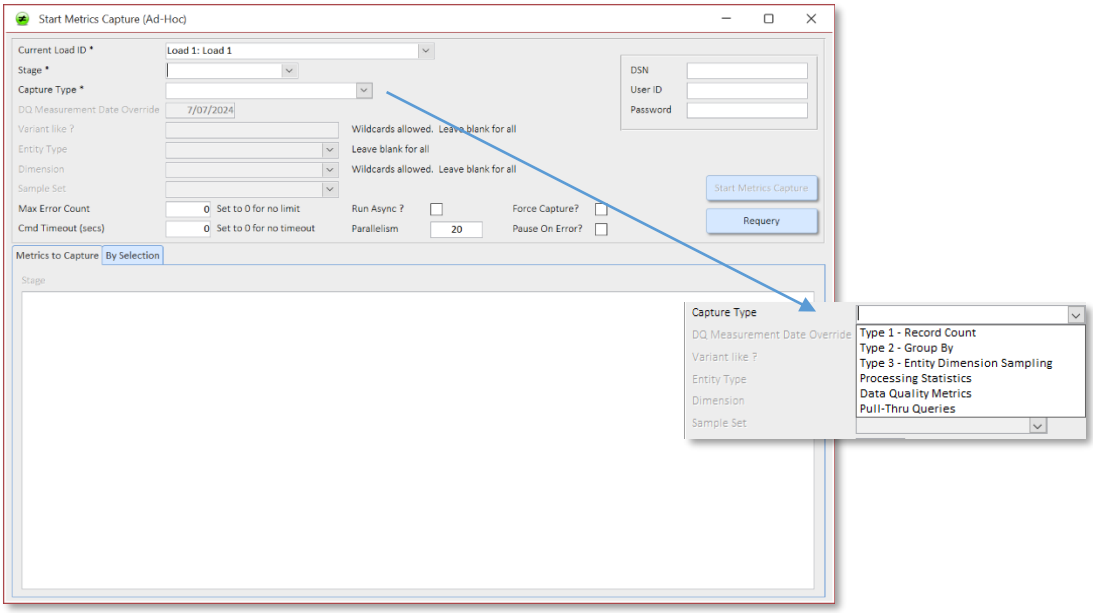
- Ad-hoc Metric Capture, and
- Scheduled / Queued Capture (Bulk Metrics Capture)



- It also allows the user to review metric count statistics to identify opportunities to purge old results

5.8.1 Ad-hoc Metrics Capture

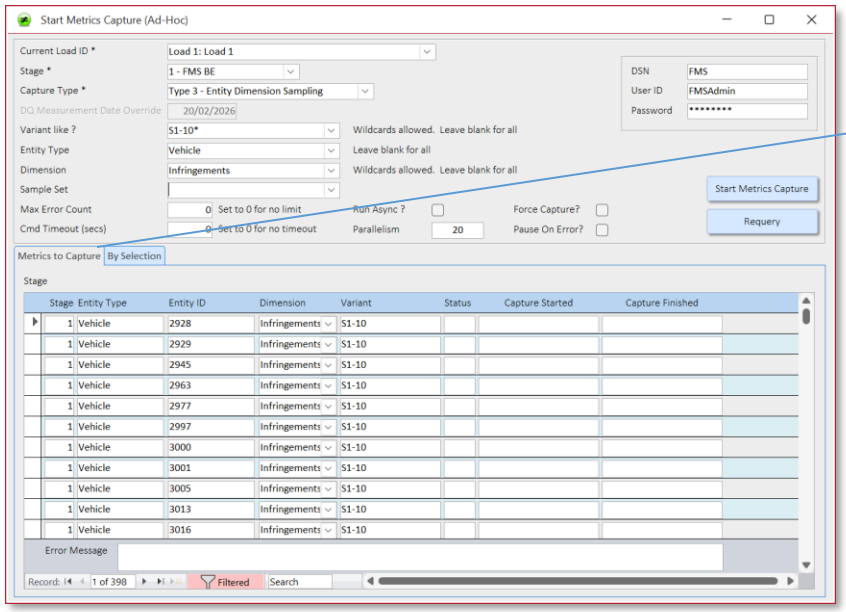
Allows the user to execute all Capture Definitions of a specific type (E.g., Type 1 Record Counts) and/or one or more individually selected capture definitions within a type. Particularly useful while developing and testing the Capture Definitions.



Field	Description
Current Load ID	The currently active Load ID – not editable. Usually there should only be one active Load ID at any point in time. If there are multiple active Load IDs however, the Load ID with the highest number will be used.
Stage	Dropdown to select the Stage from which metrics should be captured
Capture Type	Dropdown to select the type of Capture Definition(s) to be executed. Once this has been selected, several other fields on the form will be enabled / disabled depending upon their relevant to the selected type.

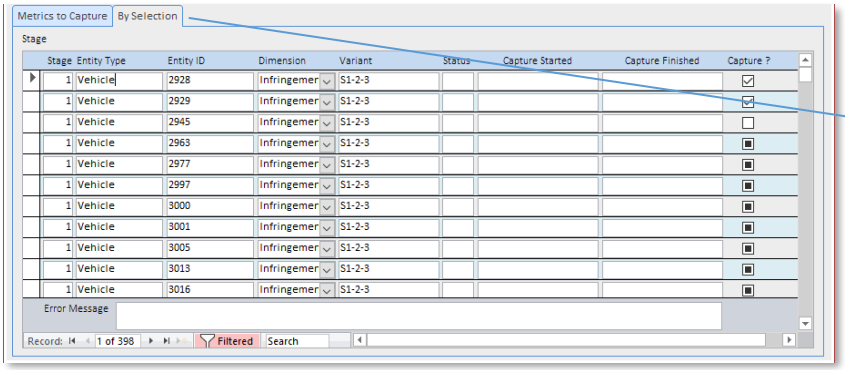
Field	Description
DQ Measurement Date Override	Only applicable for Data Quality Metrics capture type. Defaults to today's date but can be overridden if desired (generally if the capture for a pre-planned DQ Trending Date was missed but you want the results to be listed against that date anyway)
Variant Like?	Freeform text field. Allows the user to specify a specify Variant label or enter a string with one or more wildcard characters (e.g., "*" for any number of characters or "_" for exactly one character) For example, a string of "S1-10" would only pick any Capture Definitions with exactly that Variant label whereas a value of "S1-10*" would pick-up any Capture Definition with a label starting with S1-10.
Entity Type	Dropdown to filter the Capture Definitions by Entity Type. Only applicable to the Type 3 – Entity Dimension Sampling capture type.
Dimension	Dropdown to filter the Capture Definitions by Entity Dimension. Only applicable to the Type 3 – Entity Dimension Sampling capture type.
Sample Set	Dropdown to filter the list of sample Entity ID's to be used when executing the Capture Definition. Only applicable to the Type 3 – Entity Dimension Sampling capture type.
Max Error Count	If the capture process encounters this quality of errors, it will abort. Set it to 0 for no limit.
Cmd Timeout (secs)	If an individual capture definition (SQL) takes longer than this number of seconds to execute, the individual capture will be aborted. Set it to 0 for no timeout.
Run Async	If the ODBC / OLE provider supports asynchronous execution AND this checkbox is enabled then the SQL statements will be run in asynchronous mode, allowing several SQL statements to be executing concurrently.
Parallelism	If the capture SQLs are running in asynchronous mode, then this sets the upper limit on how many SQL's can be running concurrently.
Force Capture?	If the "Force Capture?" checkbox is enabled, all Capture Definitions, regardless of their processing status, will be selected for capture. If it is disabled, only Unprocessed or previously Failed capture definitions will be selected for execution.
Pause On Error?	If the "Pause On Error?" checkbox is enabled AND an error occurs during a capture process, the processing will be paused, and a popup error message will be displayed. This is useful while debugging the capture definitions. If it is disabled, the capture definition will be skipped, and processing will continue with the next definition in the list. This is useful in an unattended mode where the user wants to run several capture definitions and then review the outcome later.
DSN	DSN in the context is equivalent to the connection name. Gets pre-populated when the Stage is selected
User ID	The user ID to use for the capture executions. Gets pre-populated when the Stage is selected (if one has been defined). Not applicable for Text-based Capture Definitions / Stages.
Password	The password to use for the capture executions. Gets pre-populated when the Stage is selected (if one has been defined). Not applicable for Text-based Capture Definitions / Stages.

The screen below shows an example of the form after the various input parameters have been set.



This tab shows all Capture Definitions satisfying the selection criteria

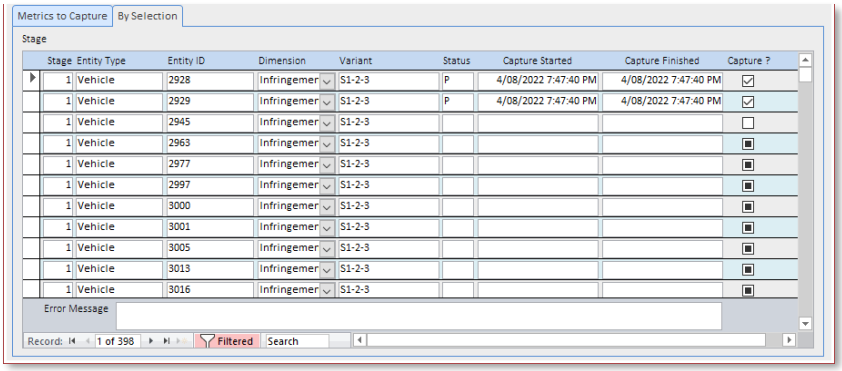
By Selection tab



This tab shows all Capture Definitions satisfying the selection criteria but provides the user with the ability to select a subset of records for processing, by setting the "Capture ?" checkbox.

Useful for testing the capture definitions during development.

After the metrics have been executed (by clicking the "Start Metrics Capture" button), the results of the execution will be shown in the form.

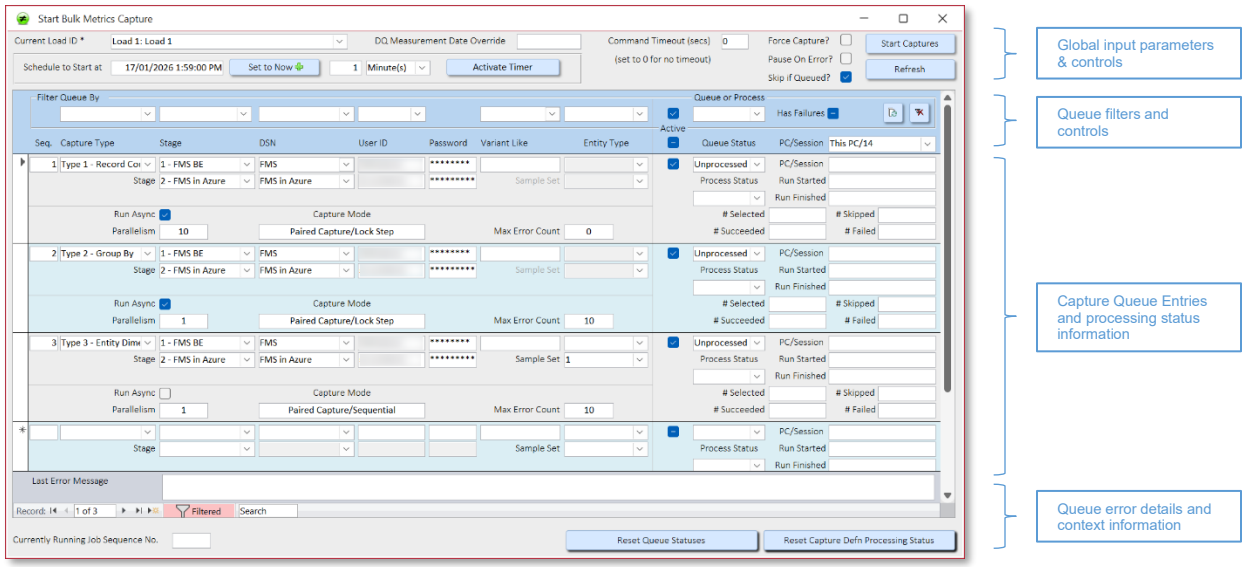


5.8.2 Scheduled / Queued Capture

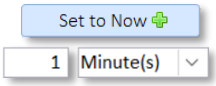
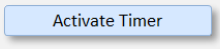
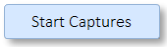
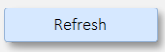
Bulk Metrics Capture form allows the user to edit and maintain a Data Capture queue for different types of Capture Types and/or different Stages and then execute the same. Execution can be either manually started or scheduled to start at a specific date / time in the future.

The Bulk Metrics Capture form has four (4) distinct sections:

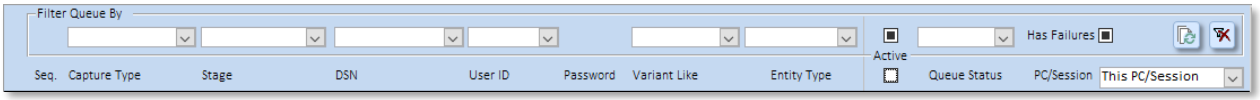
1. The top section contains the global control parameters that apply to all queue entries. It also includes the main fields and controls to schedule or manually initiate the capture processing.
2. The next section contains various fields to filter the list of queue entries. If a queue entry has been excluded by one of the filters it will NOT be captured, even it is marked as Active. This section can also be used to toggle the Active status of all unfiltered queue entries as well as filter the processing status section for a specific PC and user session.
3. The main section contains the actual queue entries specifications (left-hand side), and the processing status details (right-hand side).
4. The bottom section displays the most recent error reported during the latest execution of the selected capture queue entry, as well as the total number of queue entries included in the current filtering.



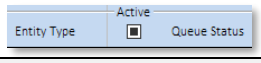
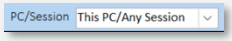
Global Input parameters and controls

Field	Description
Current Load ID	The currently active Load ID – not editable. Usually there should only be one active Load ID at any point in time. If there are multiple active Load IDs however, the Load ID with the highest number will be used.
DQ Measurement Date Override	Only applicable for Data Quality Metrics capture type. Defaults to the latest pre-defined DQ Trending date <= today. If no trending dates have been defined it will default to today's date. Can be overridden if required.
Max Error Count	If the capture process encounters this quality of errors, it will abort. Set it to 0 for no limit.
Command Timeout (secs)	If an individual capture definition (SQL) takes longer than this number of seconds to execute, the individual capture will be aborted. Set it to 0 for no timeout.
Force Capture?	If the "Force Capture?" checkbox is enabled, all Capture Definitions, regardless of their processing status, will be selected for capture. If it is disabled, only Unprocessed or previously Failed capture definitions will be selected for execution.
Pause On Error?	If the "Pause On Error?" checkbox is enabled AND an error occurs during a capture process, the processing will be paused, and a popup error message displayed.
Skip if Queued?	If multiple computers have been setup to execute the queue, the first computer to select a queue entry will set its' status to Queued. • If this checkbox is enabled (set) any subsequent computers will skip over this queue entry and pick up the next unqueued entry. • If the checkbox is disabled, any subsequent computers will still attempt to process the same queue entry but will compete for individual capture definitions within the set. **Note:** Would generally NOT be used when "Force Capture?" is enabled (as it will unnecessarily re-process the same Capture Definition multiple times)
Schedule to Start at	Allows the user to specify a scheduled time for the capture processing to start.
	Used to update the "Schedule to Start at" field by x units of time from Now. The available units of time include Minutes, Hours or Days. Note: when using Minutes as the unit, it doesn't add 60 seconds to the current time but rather sets the value to the next whole minute. E.g., if the current time was 10:17:34 (hh:mm:ss) then adding one minute will set the time to 10:18:00.
	Starts the schedule timer. Once the current time exceeds the date/time specified in the "Schedule to Start at" field, the capture processing will start for the current PC.
	Manually starts the capture processing for the current PC
	Refreshes the entire form.

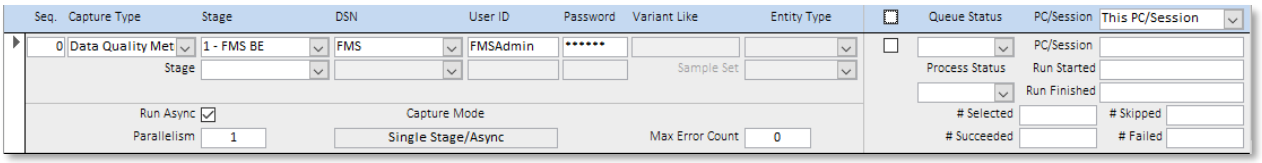
Queue filters and controls



Field	Description
Top / 1st Row	
Contains various dropdown and checkbox filters	Used to filter the queue entries by: • Capture Type • Stage • DSN / Connection • User ID • Variant Like • Entity Type

Field	Description
	• Active status (tri-state: Active, Inactive, both) • "Queue or Process" Status • Has Failures
	Filter controls. • The first button refreshes the filter dropdowns. • The second clears all the filters to display all queue entries
Bottom / 2nd Row	
	Clicking the checkbox set the Active status of any filtered queue entries, the match the status of the checkbox.
	The dropdown does NOT filter the queue entries per se but rather filters which processing statistics data to display for the queue entries. Generally, this defaults to the current computer and any session but can be used to view the processing results from different computers and/or different user sessions.

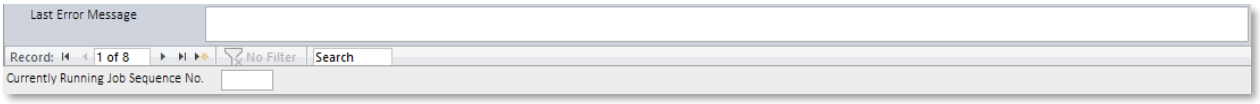
Capture Queue Entries and processing status information

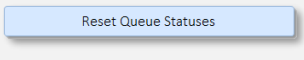
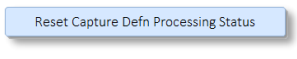


Field	Description
Core Queue Entry Specification details	
Seq No.	The sequence number in which the queue entries will be processed
Capture Type	The type of Capture Definition to be executed for the current queue entry
Stage(s)	Either one or two Stages for which the metrics are to be captured from
DSN (Connection Name)	The Connection Name to use – defaulted once the Stage is selected
User ID	User Id to use for the specific Stage
Password	The Password to use for the specific Stage
Variant Like	Optional filter to restrict the Capture Definitions to be executed as part of this queue entry. Can contain wildcards
Entity Type	Optional filter to restrict the Capture Definitions to be executed as part of this queue entry. Only applies to Type 3 Capture Type.
Sample Set	Optional filter to restrict the set of Entity ID's to be used in Type 3 Capture Definitions (only).
Queue Entry Specific Control Parameters	
Active state	Enables or disables the specific queue entry
Run Async	If the ODBC / OLE provider supports asynchronous execution AND this checkbox is enabled then the SQL statements will be run in asynchronous mode, allowing several SQL statements to be executing concurrently.
Parallelism	If the capture SQLs are running in asynchronous mode, then this sets the upper limit on how many SQL's can be running concurrently.
Capture Mode	This read-only field provides informational on the Capture Mode that will be used. The following modes are available: • Single Stage/Async • Single Stage /Sequential • Paired Capture/Lock Step: In this asynchronous mode, the capture process submits the pair of Stage SQLs for the same Capture Definition at essentially the same time. i.e. keeps them in "Lock Step". This option is useful for source / targets that are not static but are being dynamically migrated incrementally (in near real time or micro batches) • Paired Capture/Sequential
Max Error Count	The capture process will abort if this value is exceeded. Set it to 0 for no limit.
Queue Entry Processing Status fields	
Queue Status	Processing status of the overall queue entry:

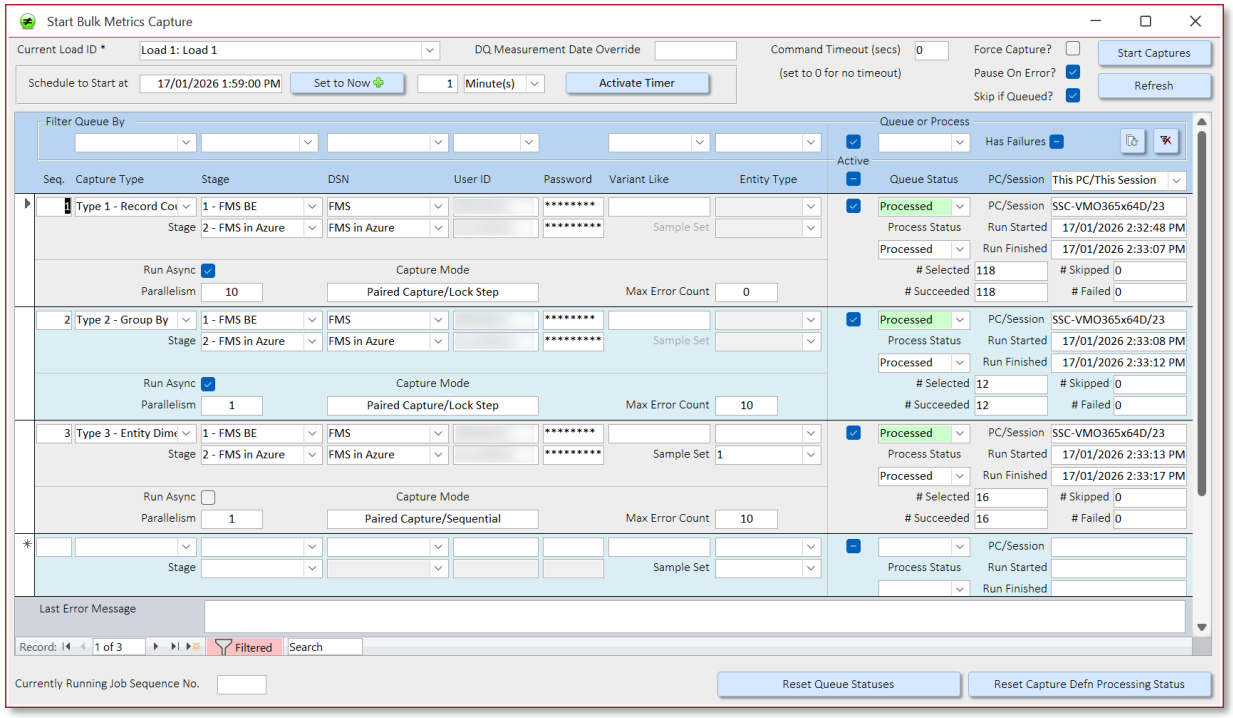
Field	Description
	• Unprocessed • Queued • In Progress • Processed • Failed • Aborted
Process Status	Status of the processing for the specific PC and Session: • Unprocessed • In Progress • Processed • Failed • Aborted • Mixed (only applies if the selected PC / Session is "All Computers" – indicates that the processing status varies across the set of PCs used)
PC/Session	Provides the context for the processing statistics: • This PC/Session • All Computers • 0 to many entries representing specific PC and Session combinations
Run Started	When the specific capture process for the selected PC and Session started
Run Finished	When the specific capture process for the selected PC and Session finished
# Selected	The number of Capture Definitions that met the Queue Entry selection criteria at the time the processing commenced for the specific PC and Session.
# Skipped	The number of Capture Definitions that were skipped by this selected PC during the applicable session – generally because another PC picked them up first.
# Succeeded	The number of Capture Definitions that were successfully processed by the selected PC during the applicable session.
# Failed	The number of Capture Definitions that failed processing by the selected PC during the applicable session.

Queue error details and context information



Field	Description
Last Error Message	If the Queue Entry selected in the main part of the form encountered an error condition, the most recent error message will be displayed here
Record x of y	The record number of the selected queue entry (x) within the filtered list of y records
Currently Running Job Sequence No.	Self-explanatory. If multiple queue entries have the same Job No., the active one is also highlighted in Yellow while it is running.
	Resets the Queue and Processing statuses back to Unprocessed. Generally, not needed unless there has been an interruption during a previous load and queue entries were left in a Queued or In Progress state.
	Resets the status of the capture definitions, back to U (for Unprocessed) related to the filtered list of capture queue entries. This is provided as a convenience to avoid the need to use “Force Capture?”

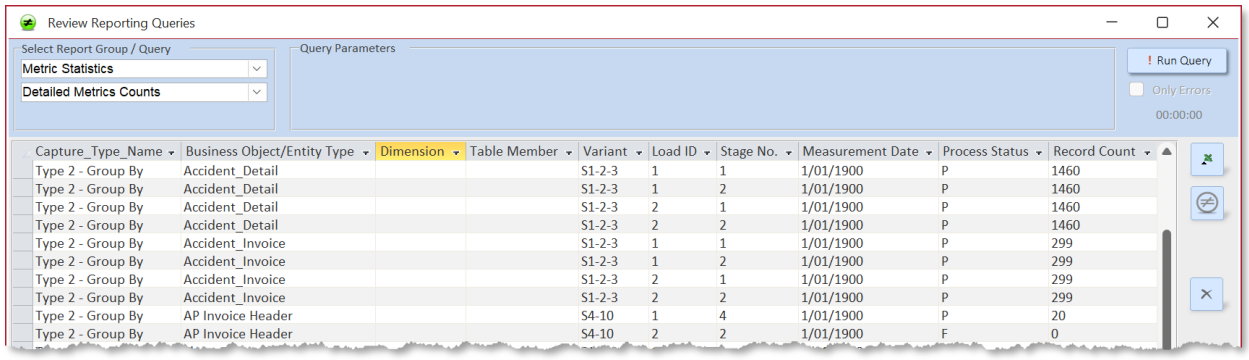
The screenshot below provides an example of what the queue looks like after the capture process has been run.



Where failures were encountered, the user would generally go back to the relevant "Maintain Capture Definitions" form to determine the cause of the issues and resolve them accordingly.

5.8.3 Review Metric Counts

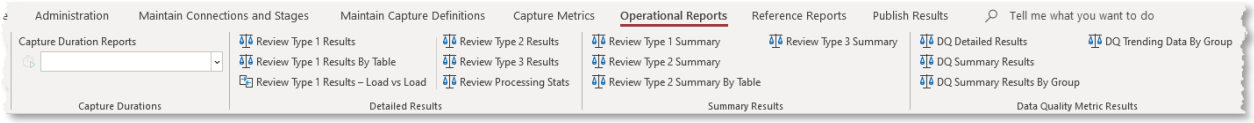
This function displays the record counts of the detailed metric captures, by Capture Definition and Load ID / Measurement Date (for Data Quality related captures). This is useful to identify opportunities to purge old results e.g. prior to a database upgrade or just to reduce disk space (by conducting a purge followed by a Compact BE Database command). Users with the appropriate permissions can purge individual results from this form using the × button.



Note: you will need to rerun the query to see the updated counts.

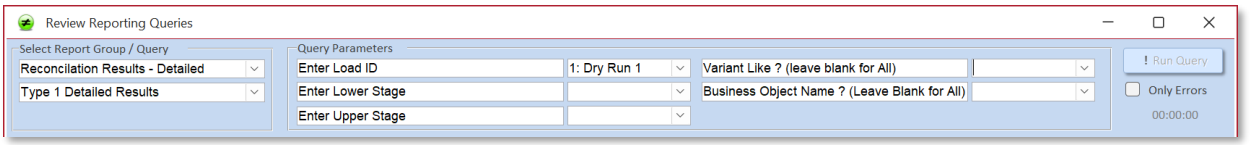
5.9 Reviewing the Capture / Reconciliation Results

Having executed the Capture Definitions and retrieving the relevant metrics, the next step is to review the results via the Operational Reports menu.

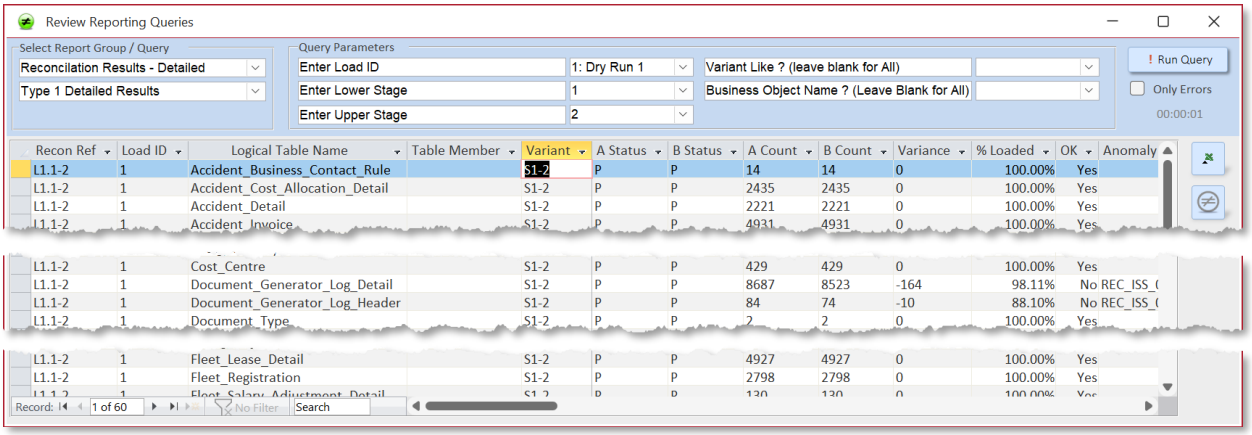


The various types of Operational Reports are described in section 4.1.6. All the reports use a common Report Viewer form that allows the user to select the required report, enter the relevant input parameters and then execute the report.

For the purposes of this guide, we will select and run Type 1 Detailed Results report (within the "Reconciliation Results – Detailed" Report Group).

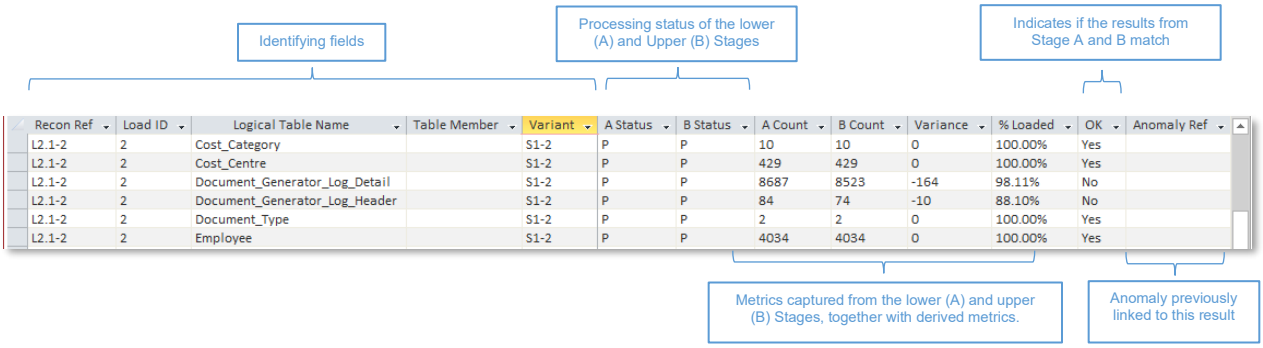


The form displays the list of Query Parameters required to run the specific report. Once all the mandatory input parameters have been entered, the "Run Query" button will be enabled. Below is an example output for a Type 1 Detailed Results report.



5.9.1 How to interpret the results

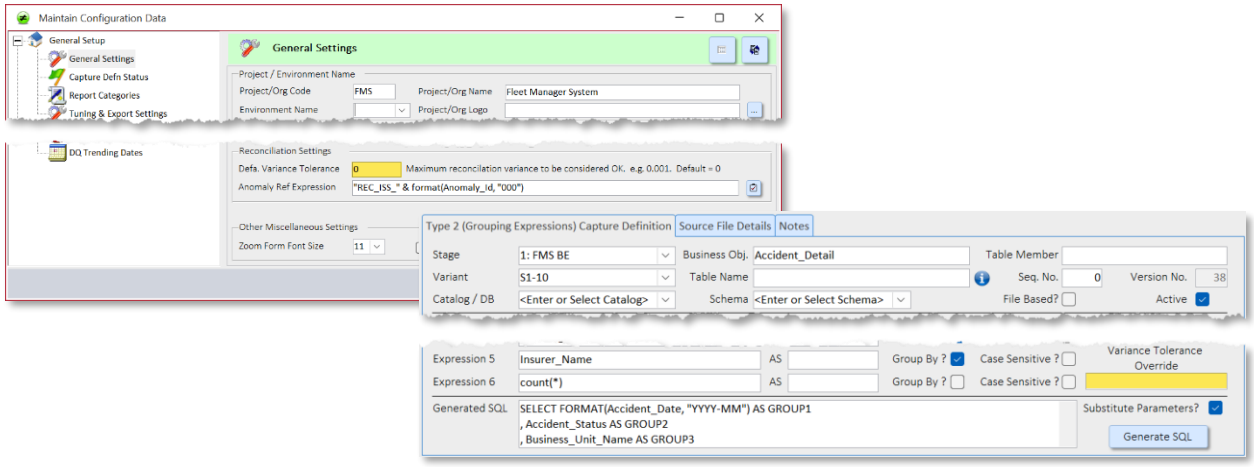
The key elements of a reconciliation report are illustrated below.



Recon Ref	Load ID	Logical Table Name	Table Member	Variant	A Status	B Status	A Count	B Count	Variance	% Loaded	OK	Anomaly Ref
L2.1-2	2	Cost_Category		S1-2	P	P	10	10	0	100.00%	Yes	
L2.1-2	2	Cost_Centre		S1-2	P	P	429	429	0	100.00%	Yes	
L2.1-2	2	Document_Generator_Log_Detail		S1-2	P	P	8687	8523	-164	98.11%	No	
L2.1-2	2	Document_Generator_Log_Header		S1-2	P	P	84	74	-10	88.10%	No	
L2.1-2	2	Document_Type		S1-2	P	P	2	2	0	100.00%	Yes	
L2.1-2	2	Employee		S1-2	P	P	4034	4034	0	100.00%	Yes	

Differences for the Type 2 and Type 3 Reconciliation results include:

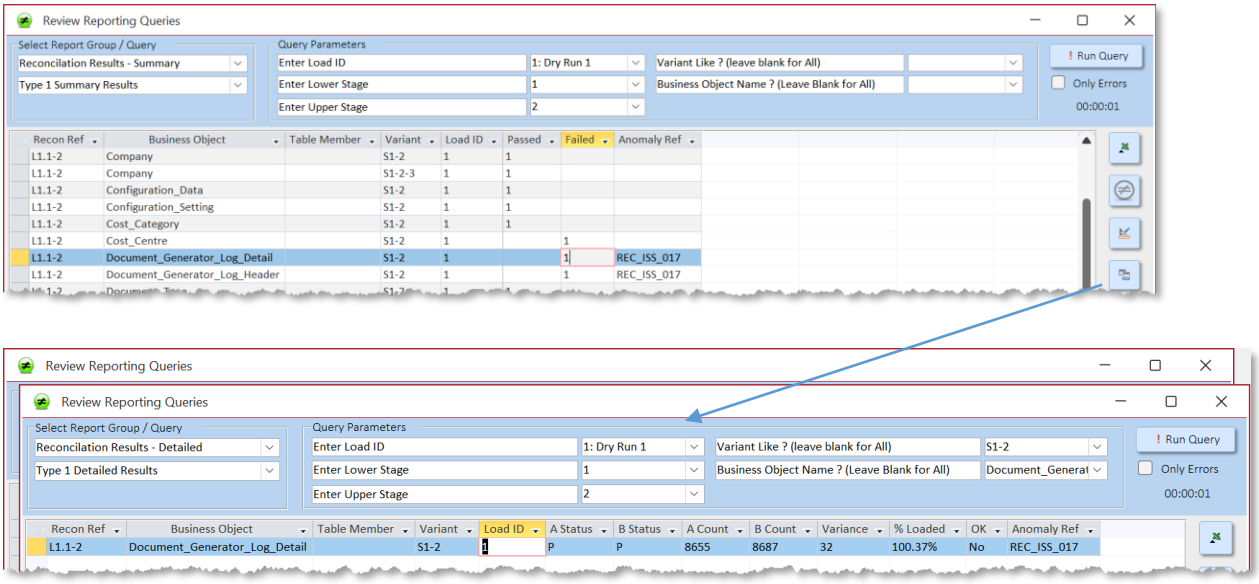
- The Identifying fields will include Group 1 to Group 5 fields, as well as Entity and Dimension elements (for Type 3 only)
- “A Value 1” to “A Value 5” (for the lower Stage) and “B Value 1” to “B Value 5” (for the upper Stage) – these contain the expression metrics not flagged as Group By fields.
- Var 1 to Var 5 – the variances between the corresponding A & B Value metrics.
- Matched field. This is a Yes/No derived column that indicates that the Identifying fields from Stage A were also found in Stage B. Where mismatches are found (i.e., Matched = No) this could be indicative of either:
    - Missing data from one of the Stages, or
    - Incorrect group by expressions in the Capture Definition in one of the Stages. e.g. if for one Stage we were grouping a date field in "yyyy-mm" format while in the other stage we are grouping by "yyyy", we will get mismatched results.
- OK field. This is a Yes/No derived column that indicates whether all Variance fields are 0 (or less than a pre-configured Variance Tolerance). The Default Variance Tolerance is defined under the Administration / Configure Application / General Settings. A Variance Tolerance Override can be defined at an individual Capture Definition and takes precedence over the default value.



5.9.2 Drilling Down from Summary Results

When reviewing any of the “Reconciliation Results – Summary” related reporting queries, the review form provides the ability to select a particular record from the results and drill-down into the corresponding detailed report.

For example, having selected the “Document_Generator_Log_Detail” record in the screenshot below and then clicking on the Drilldown button,  a second review form will be opened with the corresponding detailed results.



5.10 Recording Anomalies

After the Capture Definitions have been executed and the results analysed, reconciliation errors (anomalies) may be identified. The anomalies may be the result of errors in the data migration logic, or they may be errors in the reconciliation logic. Either way, they should be recorded, analysed and ideally resolved before the next execution.

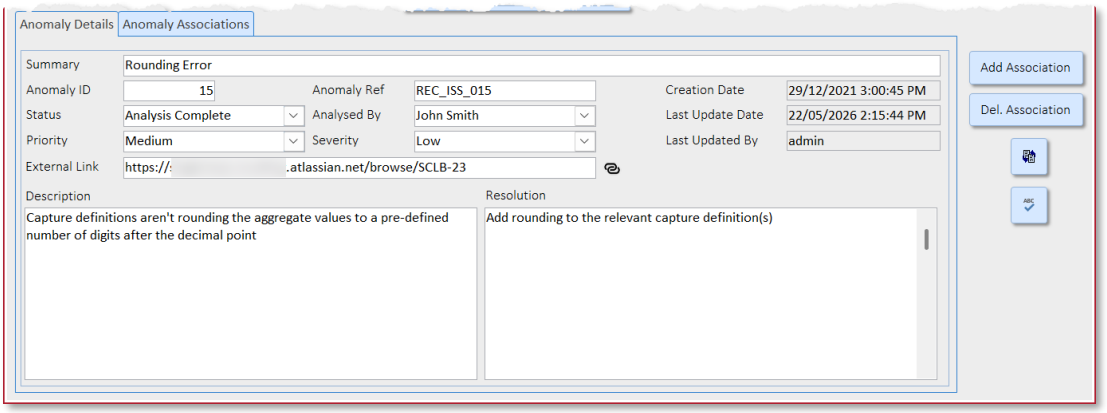
5.10.1 Anomaly Fundamentals

Before proceeding, it's worth discussing some basic definitions and concepts around anomalies, as implemented within ReQuon.

There are two (2) key components required to define an anomaly with ReQuon:

1. The Anomaly Details (header), plus
2. One or more associations between the Anomaly Details and the captured reconciliation metrics where the anomaly was detected.

As a minimum, an Anomaly Details record includes a system assigned Anomaly ID (non-editable), a system generated Anomaly Ref (configurable / editable), a Description and a Status. Over time, details such as the root cause and the resolution will be added once the analysis has been performed. If the project uses a separate defect tracking system (such as Jira) a link to the external ticket can also be recorded.



Anomaly association records link the anomaly header record to one or more metric capture results, that are impacted by the specific anomaly. The association records can be linked at different levels with the metric capture hierarchy, as illustrated in Figure 10 - Anomaly Hierarchy Levels. Table 1 - Anomaly Hierarchy Levels by Capture Types, details which Association Levels are applicable to which Capture Types.

Figure 10 - Anomaly Hierarchy Levels

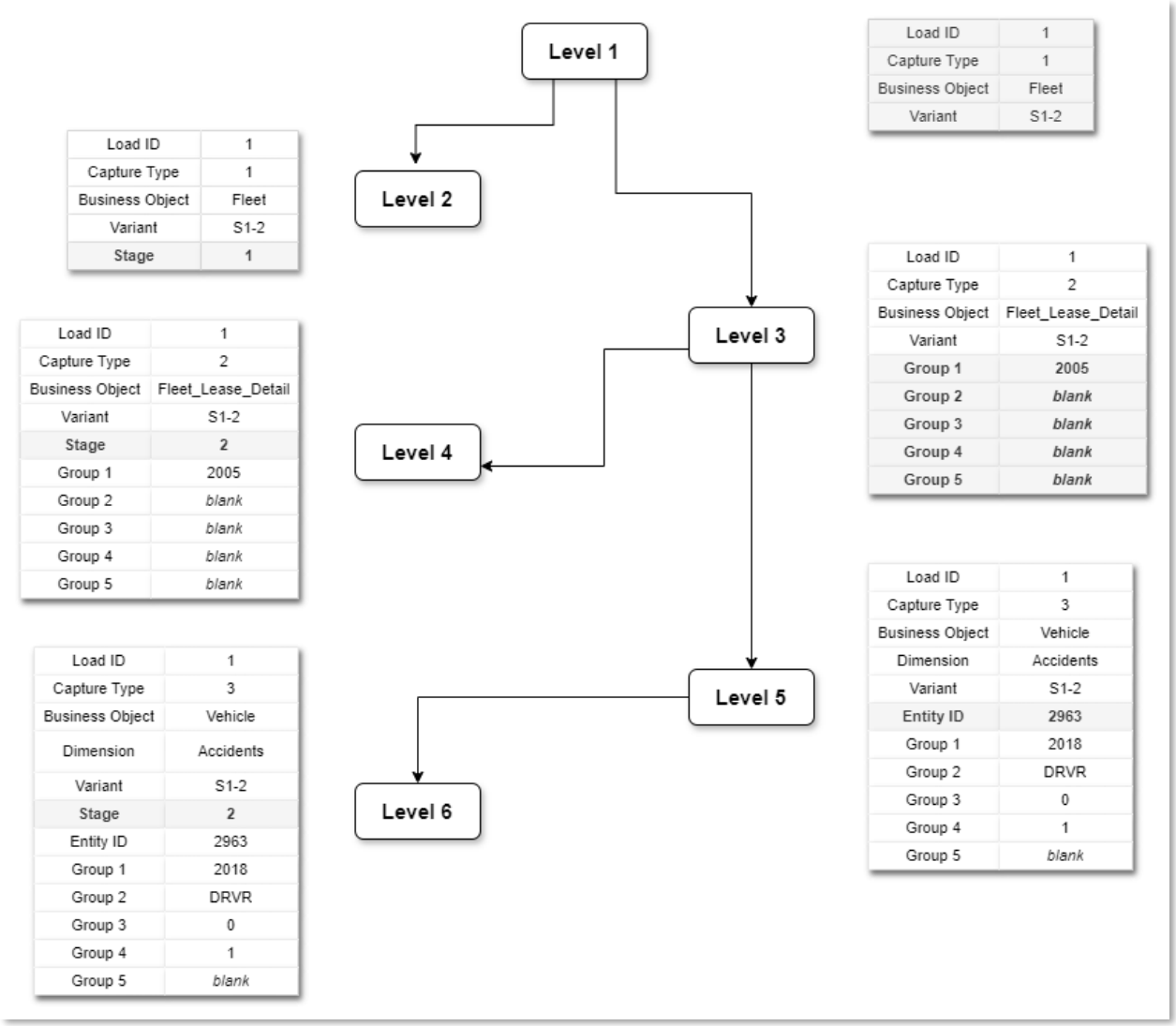


Table 1 - Anomaly Hierarchy Levels by Capture Types

Association Levels	Applicability By Capture Type			
	Type 1	Type 2	Type 3	Processing Statistics
Level 1 – Load ID & Capture Def'n	✓	✓	✓	✓
Level 2 – Load ID, Capture Def'n & Stage	✓	✓	✓	✓
Level 3 – Load ID, Capture Def'n, & Group By's	✗	✓	✓	✓
Level 4 – Load ID, Capture Def'n, Group By's & Stage	✗	✓	✓	✓
Level 5 – Load ID, Capture Def'n, Group By's & Entity ID	✗	✗	✓	✗
Level 6 – Load ID, Capture Def'n, Group By's, Entity ID & Stage	✗	✗	✓	✗

5.10.2 Different ways of creating an Anomaly and/or Association

There are several pathways for creating anomalies and their related association(s), however, they all utilize a common “New Anomaly and/or Association” wizard to guide the user through the process. The wizard is described in the section 0. The three pathways are:

1. Via the Maintain Capture Definitions forms
2. Via the Review Reporting Queries form, or
3. Via the Maintain Reconciliation Anomalies form

Note: While any anomaly, at any anomaly association level, can be created via any of the 3 pathways, there are different scenarios that may favour one method over another.

5.10.2.1 Anomalies via the Maintain Capture Definitions forms

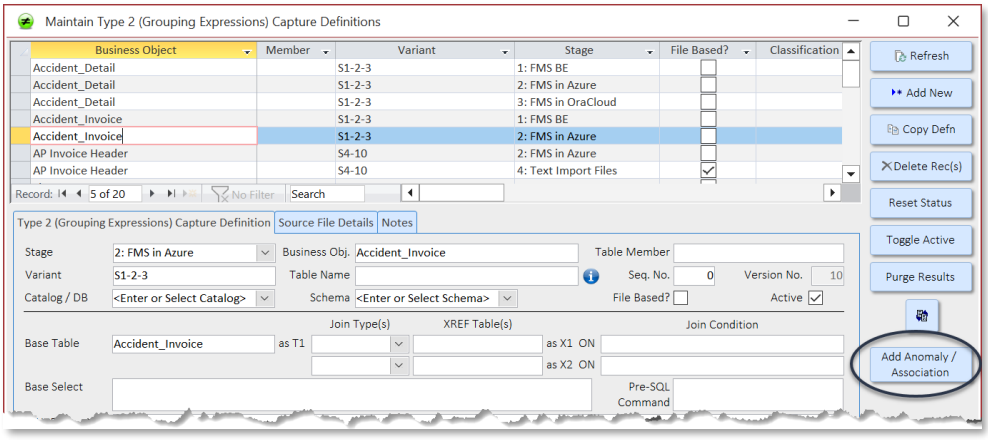
The four capture definition forms that support the creation of anomalies/associations are:

- Type 1 (Record Counts),
- Type 2 (Grouping Expressions),
- Type 3 (Entity Dimension Sampling), and
- Type PS (Processing Statistics)

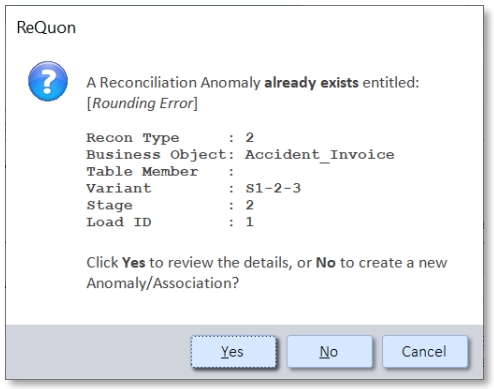
The capture definition forms would typically be used for creating anomalies and associations at Level 1 (Load ID & Capture Def’n) and Level 2 (Load ID, Capture Def’n & Stage). i.e. Level 1 and Level 2 are commonly caused by errors in the data migration or reconciliation logic that affects all the related captured metrics data.

An example of a **Level 1** anomaly/association could be a situation where rounding errors are occurring for **both stages** related to a capture definition.

An example of a **Level 2** anomaly / association might be an incorrect set of selection criteria on **one of the stages** associated with a capture definition. This might require changes to the data migration logic and/or the reconciliation logic.



Upon clicking the Add Anomaly / Association button, if an existing anomaly association already existed for this Capture Definition, then you would see a dialog box like the following:

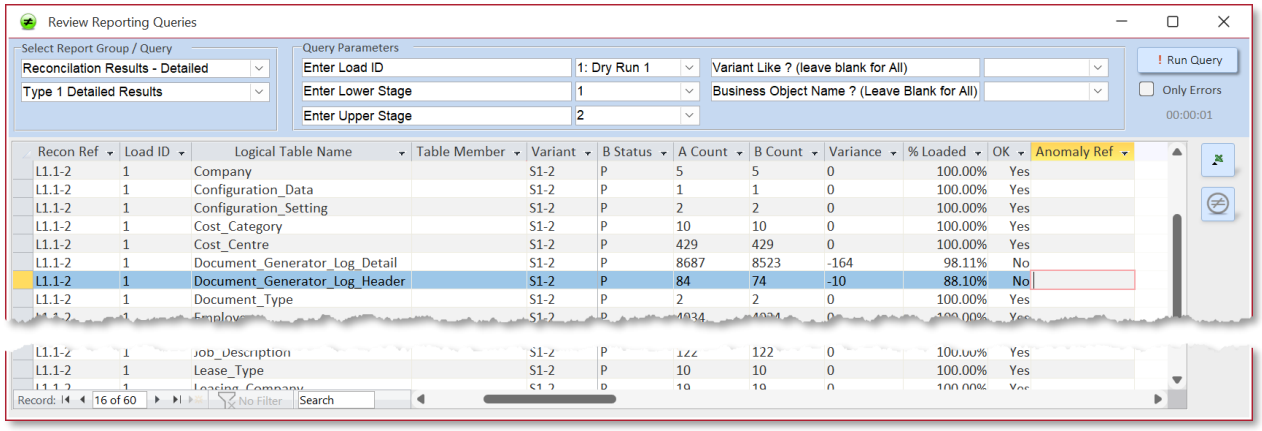


Clicking “Yes” will open the Maintain Anomalies form, while clicking “No” will launch the “New Anomaly and/or Association” Wizard. If no existing anomaly has been associated with this Capture Definition, then the “New Anomaly and/or Association” Wizard will be launched.

5.10.2.2 Anomalies via the Review Reporting Queries form

The Review Reporting Queries form is accessible via the Operational Reporting menu. If the reporting query being reviewed supports anomalies, then the Add Anomaly / Association button will be enabled. The Review Reporting Queries form is suitable as a launching pad for all types and levels of anomalies/associations, since the relationship between the results and the anomalies are clearly visible.

In the example below, we can see a variance in the record counts for Logical Tables "Document_Generator_Log_Detail" and "Document_Generator_Log_Header". The “Anomaly Ref” column is blank indicating that no anomaly record has been linked to these results.

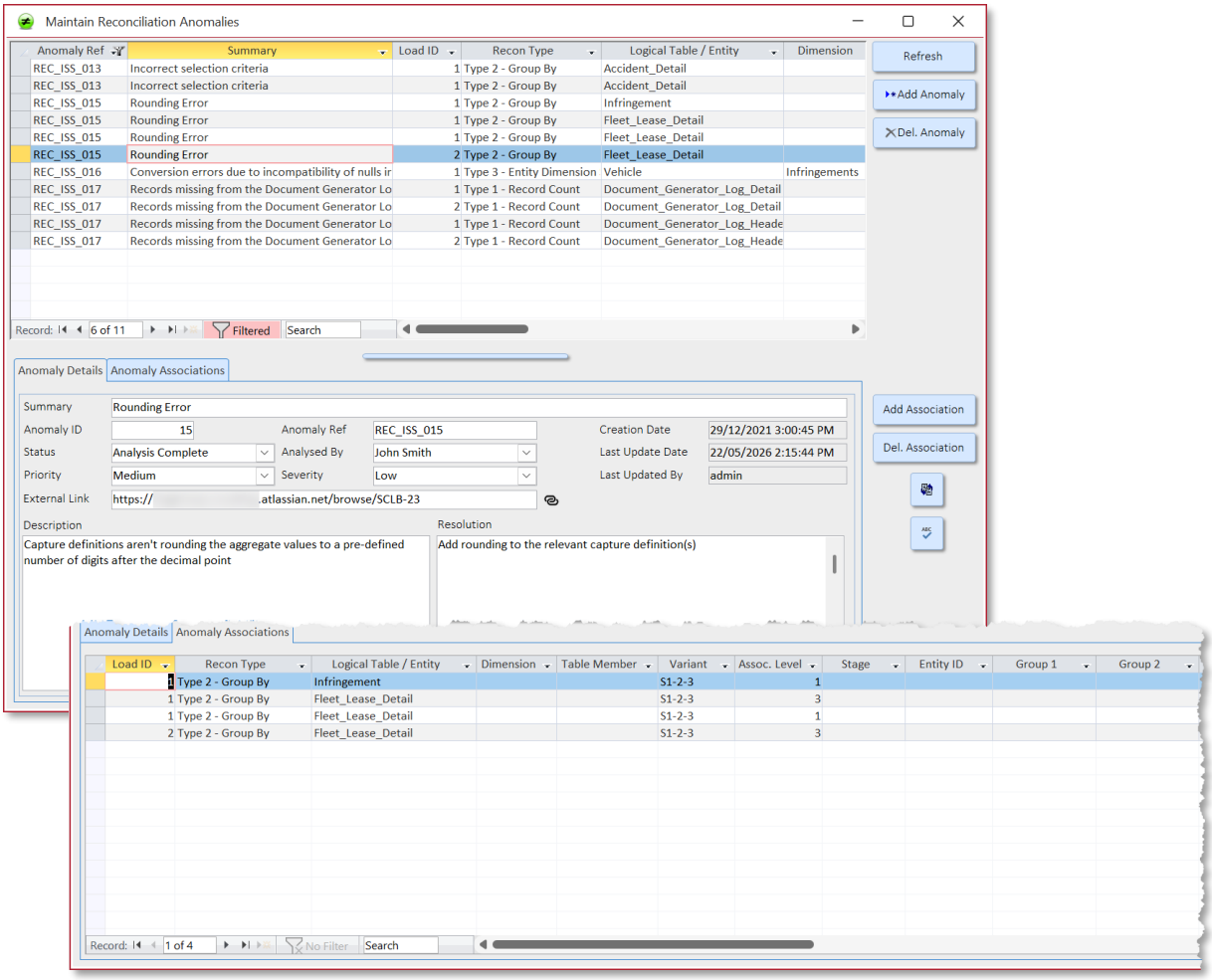


To add a new anomaly/association, select the relevant record, click on the icon, and follow the instructions.

5.10.2.3 Anomalies via the Maintain Reconciliation Anomalies form

The Maintain Reconciliation Anomalies form is accessible via the Publish Results menu. It not only supports the creation of any type of anomaly/association but is also the **only mechanism** that supports the deletion of either an anomaly association or an anomaly itself.

It is also the main channel for updating the anomaly details as the root cause analysis is conducted, a resolution identified and implemented, and subsequent retesting verifies that the issue has been resolved.



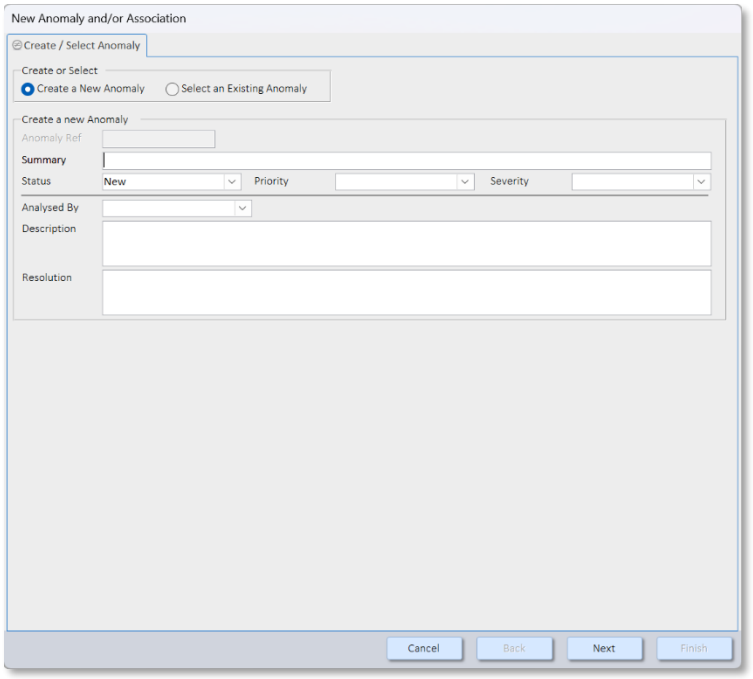
Clicking the Add Anomaly button inserts a new record into the table and positions the user into the Anomaly Details section of the form where the Description and other details can be entered. At this point no association has been created. Clicking the Add Association button will launch the "New Anomaly and/or Association" Wizard.

5.10.3 The "New Anomaly and/or Association" Wizard

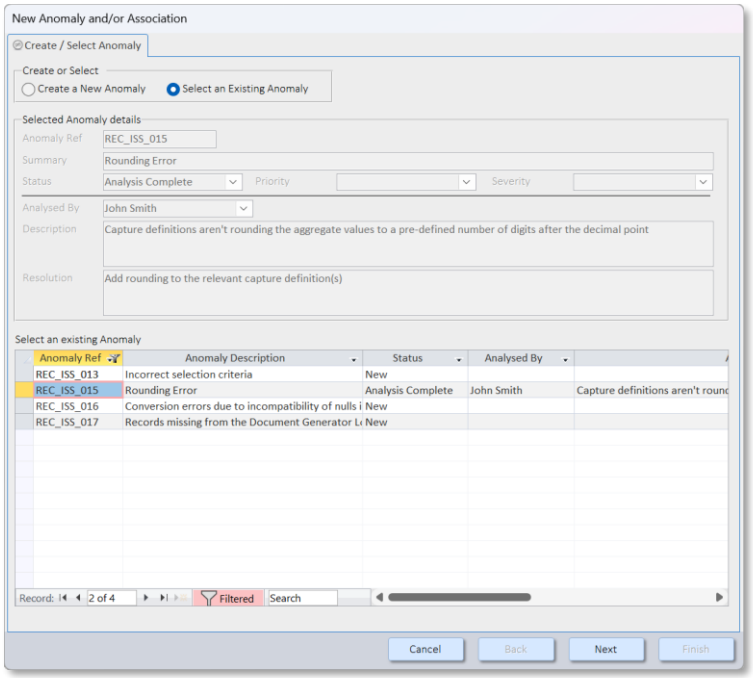
The operation of the "New Anomaly and/or Association" wizard is described below.

Step 1 – Choose "Create a New Anomaly" or "Select an Existing Anomaly"

If creating a new anomaly, enter the relevant details (Description as a minimum) and then click Next.

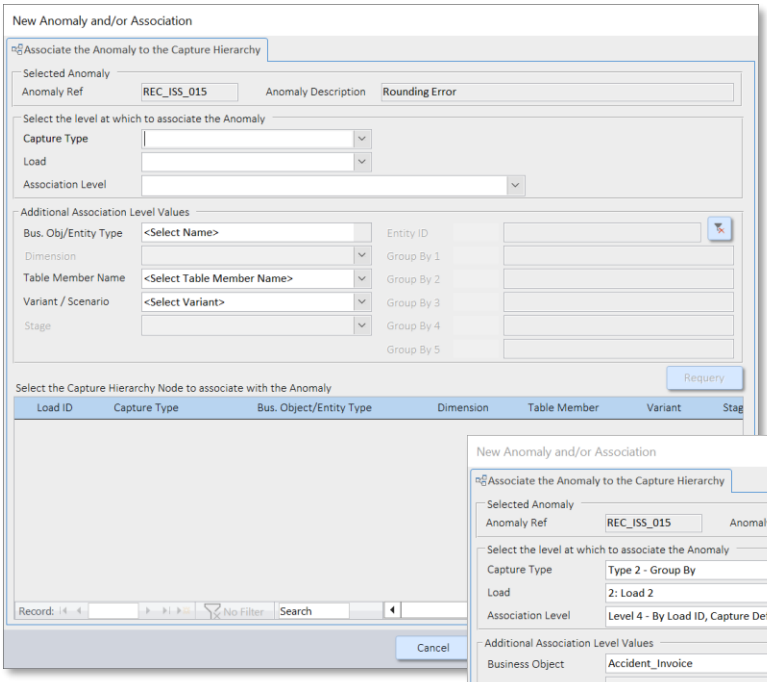


If looking to create a new association against an existing anomaly, select the appropriate anomaly from the list and then click Next.



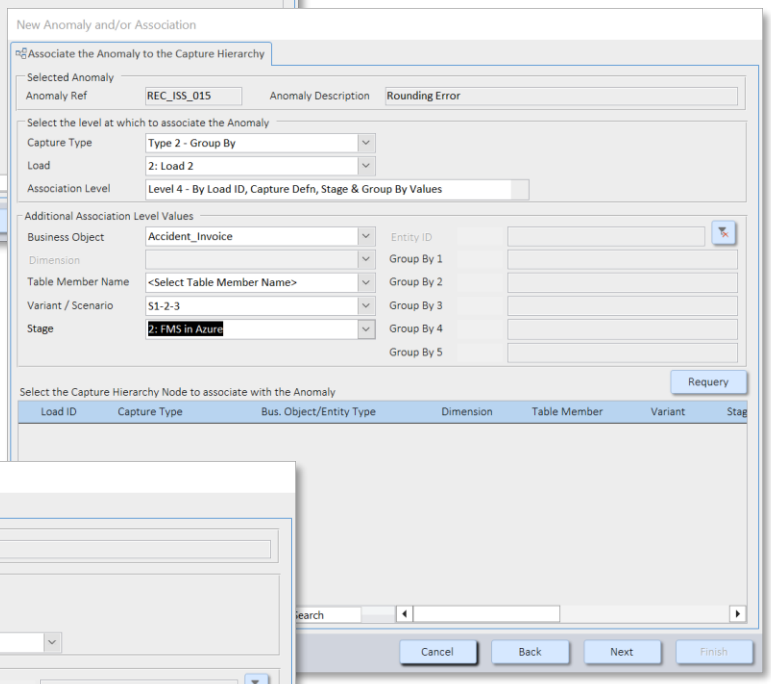
Step 2 – Select the level at which to associate the Anomaly

The next step is to decide the level, within the metrics capture hierarchy, to create an association to the anomaly. As a minimum, the Capture Type, Load Id and Association Level must be specified. For association levels > 1, additional parameter values need to be entered.

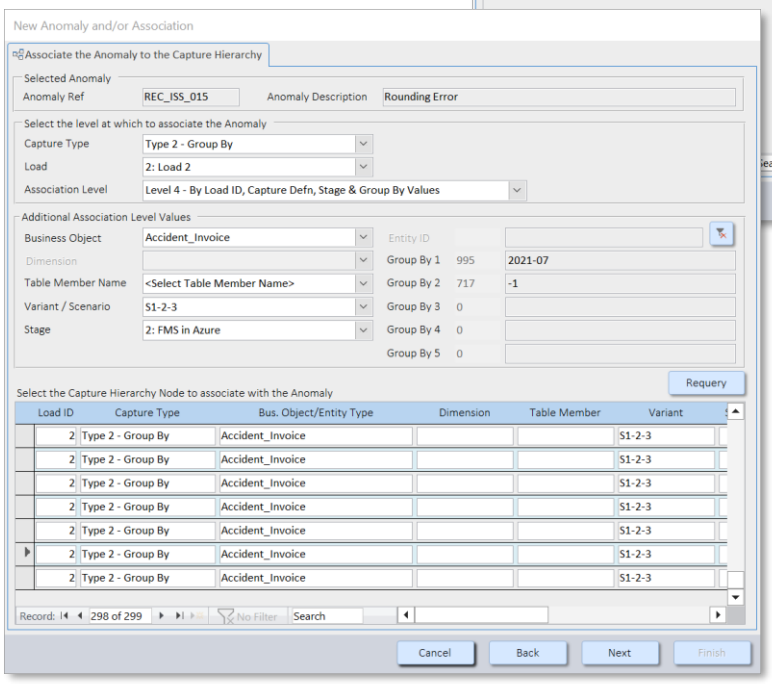


Fill out any additional parameters from the fields enabled.

Note: The Entity Id and Group By fields cannot be entered directly but are populated automatically when a record from the list of Capture Hierarchy Nodes has been selected.



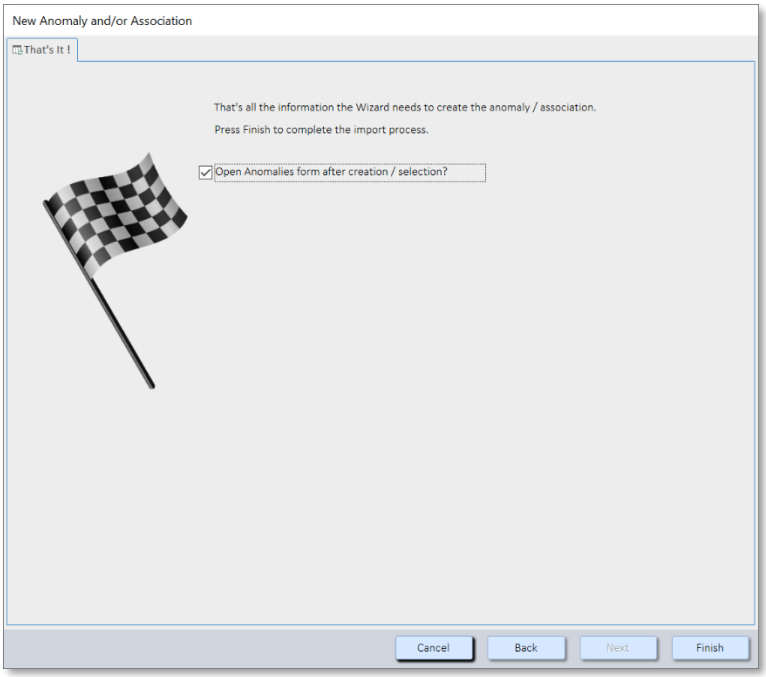
Click the Requery button to populate the Capture Hierarchy Node list based on the parameter values entered.



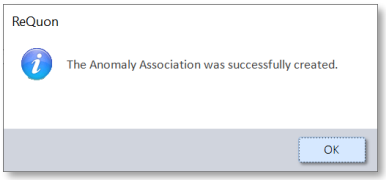
As you select different records from the list, you will see the relevant Entity ID and/or Group By values populated. Once you have selected an appropriate record, click on the Next button.

Step 3 – Confirm your selections.

If all required fields from the previous screen have been entered, the final screen will be displayed awaiting your confirmation to proceed. Click Finish to create the anomaly association. You will be presented with an appropriate message advising whether the anomaly association was successfully created or not.



If the creation was successful, and you ticked the checkbox “Open Anomalies form after creation / selection?”, the Anomalies form will be opened with the newly created Anomaly / Association pre-selected.



5.11 Reference Reports

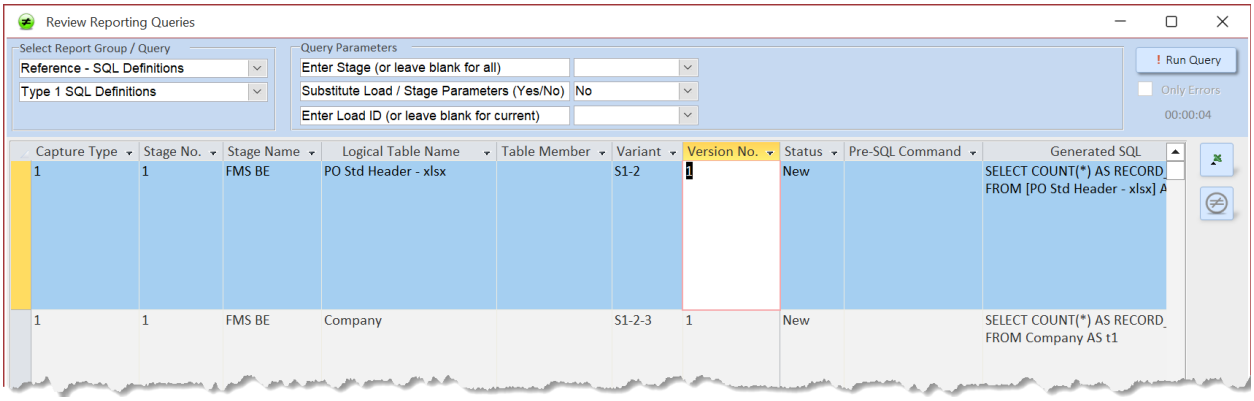
The reference reports use the same Review Reporting Queries form discussed under the Operational Reports, section 5.9. The types of reference reports are described in section 0,

Reference Reports Menu, but essentially cover:

- The SQL Definitions associated with the various types of Capture Definitions, and
- Data Quality Metrics related reference data (Metric definitions, DQ Owners and Targets)

SQL Definitions

The screenshot below shows an example of the output of a Type 1 SQL Definitions report executed for All Stages, no substitution of parameter values for the current Load ID.



The results can be exported to Excel by clicking on the export button and following the prompts.

5.12 Publishing Results

Generating reports and analysis during a data migration exercise, particular the final Go-Live conversion, can be a time-consuming and stressful process. ReQuon simplifies and fast-tracks this process by allowing the data reconciliation analyst to:

- Pre-configure one or more consolidated "Report Export Group" definitions, comprising multiple operational and reference data reports, for publishing to a single Excel Workbook.
- Utilize out-of-the-box Excel templates to accelerate setup of the reports. Can be customized as required.
- The consolidated reports can be run individually (on-demand), or bulk published as a set of reports via a single initiation.

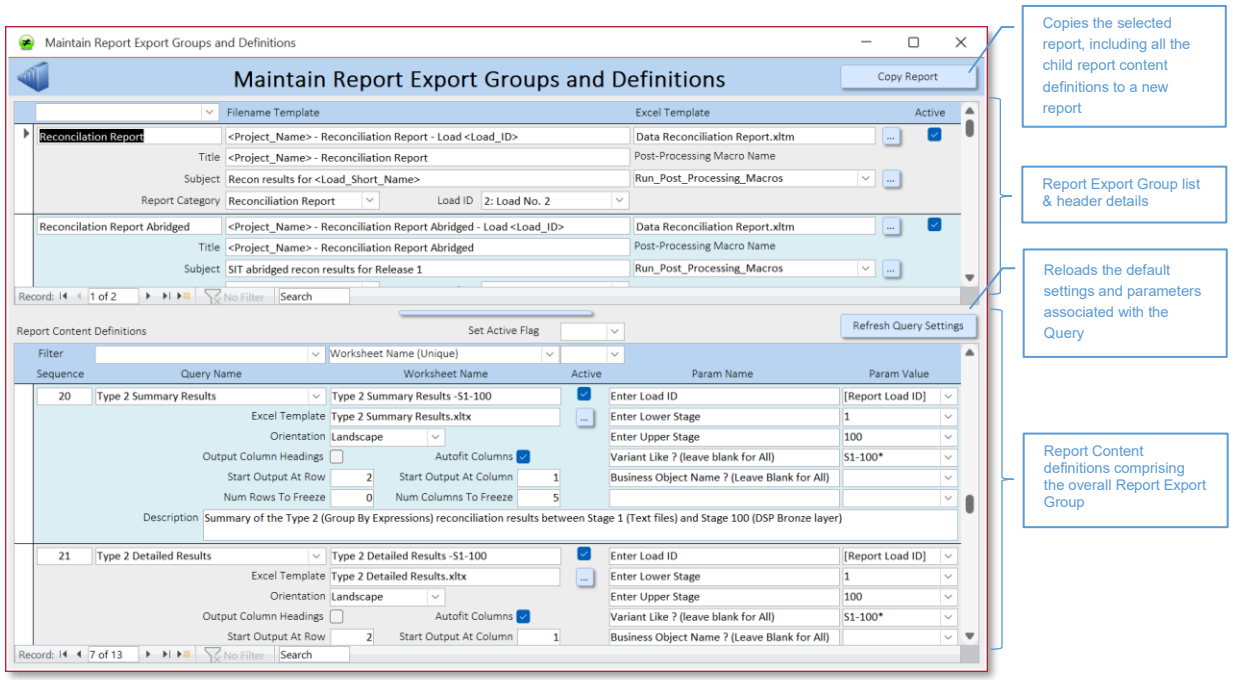
The various menu options under the "Publish Results" menu are described in section 4.1.9, Publish Results Menu. This section will focus primarily on the "Maintain Report Groups & Definitions" form.

5.12.1 Maintaining Report Export Groups and Definitions

The Maintain Report Export Groups and Definitions form is comprised of two (2) main sections:

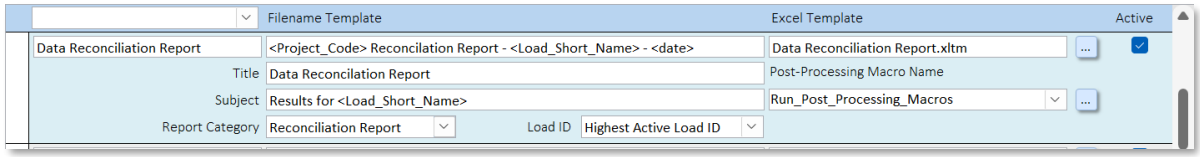
1. The Report Export Group list, where new report groups can be created, and header level parameters defined.

2. The Report Content Definitions list, where individual report queries are added to the report group and configured with the relevant input parameters required.



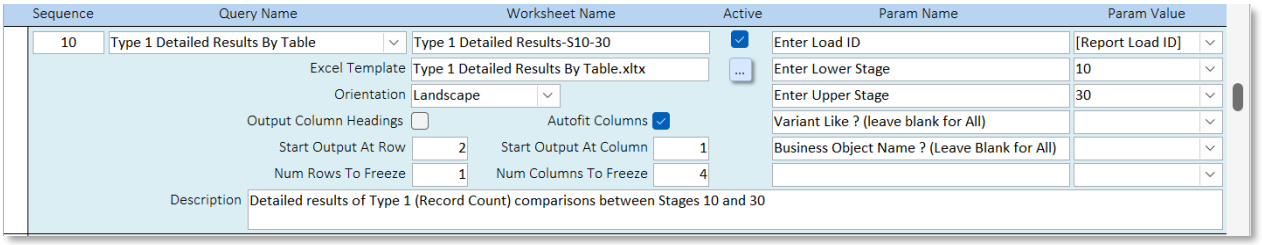
We will start by looking at the setup of a typical Reconciliation Report.

Report Export Group header



Field	Description
Short-Name	Used in the report menus to refer to the specific report
Filename Template	A template for the export filename (excluding the file extension). Can contain substitution parameters such as <Load_ID> and <date>. Right-click in the field to access the substitution parameter popup menu.
Title	Title for the report. Can be substituted into the report during generation.
Subject	Subject of the report. Can be substituted into the report during generation.
Report Category	Selects one of the pre-configure report categories, which can be used in the Output Directory configuration to determine where the generated report will be saved to.
Load ID	The Load ID specified here can be automatically passed to every query defined within the report group
Excel Template	The base Excel template to use when generating the report. Can contain static content not sourced from the report content queries
Post-Processing Macro Name	The name of a macro within the Excel Template that you want to run after the contents of the report have been generated. E.g to update a table of contents.
Active	Only active reports will be visible in the report menus for execution.

Report Content Definitions



Field	Description
Sequence	The sequence number field determine the order of the worksheets within the consolidated Excel workbook.
Query Name	Dropdown list containing the names of 30+ in-build operational and reference data reporting queries that can be incorporated into the report export group.
Worksheet Name	The name that will be given to the worksheet tab within the consolidated report
Active	Only active content definitions will be included in the generated report. Enabling or disabling individual definitions allows the user to customise the content for a particular report execution without having to permanently delete entries.
Excel Template	The name of an existing Excel Template suitable for the selected reporting query. Can be one of the out-of-the-box templates or a custom-made template
Orientation	Landscape or Portrait. Landscape is the most common orientation used.
Output Column Headings	If enabled, the first exported row will contain the column headings prior to the actual data. Usually disabled when using a template that has customized the headings.
Autofit Columns	If enabled, the column width with automatically resized fit the data. Useful when no template is used.
Start Output at Row	For extracts without an Excel Template, this will usually be row 1. Where a template it used it is usually row 2 and we disable "Output Column Headings"
Start Out at Column	Usually set to column 1
Num Rows to Freeze Num Columns to Freeze	Used to freeze the corresponding number of rows and columns in the Excel spreadsheet to assist when scrolling up and down through the sheet
Autofit Columns	Will automatically resize the columns to fit the data. Useful when no template is used.
Param Name 1 to 6 Param Value 1 to 6	Upon selecting the Query Name, the relevant parameter names will automatically be populated. The user specifies the appropriate parameter values suitable for the required output, utilizing the values presented in the Param Value dropdown lists, or free text (as applicable to the parameter). Note: For parameters involving Load ID, [Report Load ID] is a special value that means use the Load ID that was specified in the Report Export Group header record.
Description	A user-defined description to explain what the extract represents. The description can be substituted into the export as part of the Table or Contents. Refer to the out-of-the-box "Data Reconciliation Report.xltm" template.

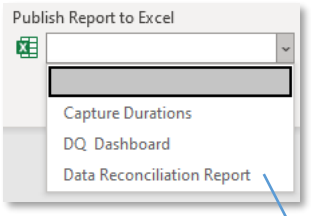
5.12.2 Running the Reports

After configuring the required Report Export Groups, they can be run via one of two ways:

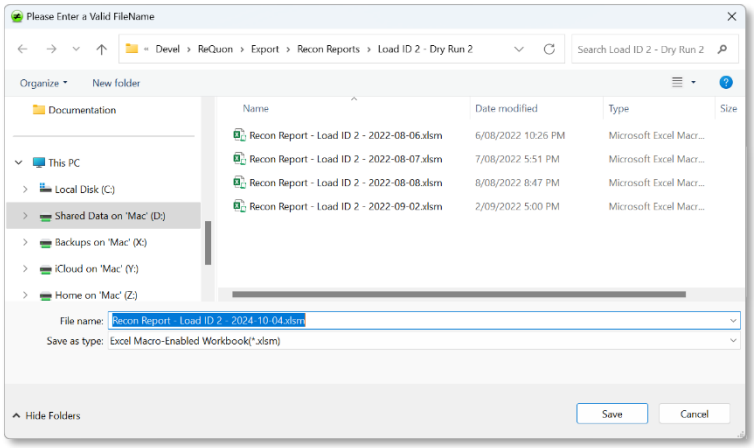
1. Selection from the "Publish Report to Excel" dropdown, or
2. Selection via the "Bulk Publish Report" form

Publish Report to Excel

The dropdown will list all active Report Export Group reports. To run the report, simply select it from the list.



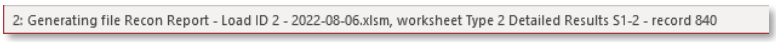
A dialog box will be displayed to specify the directory and filename to use when saving the report. Both the directory and the filename will be defaulted based on the Output Directory configuration and the filename template specified in the Report Export Group header (after any substitution parameters have been applied).

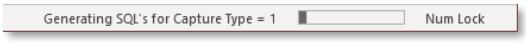


Click Save to start the generation process.

Warning: If a file already exists with the same name, it will be overwritten.

While the report is being generated the progress of the current query can be viewed in either the left or right-hand side of the Status / Message bar, at the bottom of the application window.



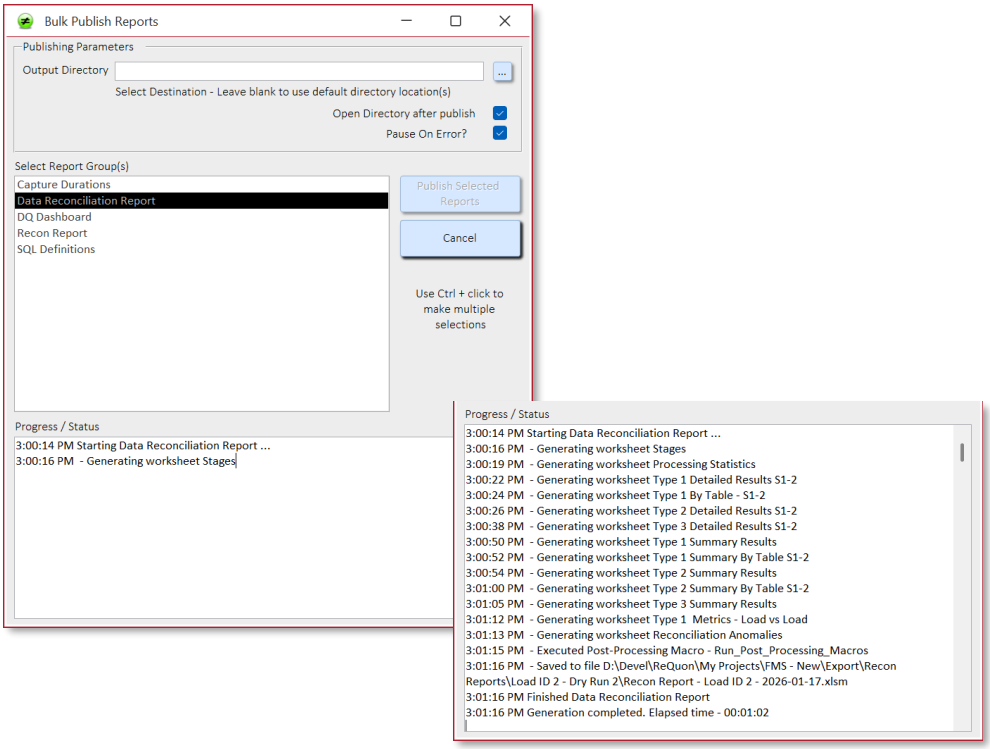


At the end of the generation process, the exported file will be automatically opened in Excel for viewing.

Bulk Publish Report

The Bulk Publish Reports option allows the user to select multiple reports to be generated and saved to either:

- a common folder (if the Output Directory is specified), or
- the default locations previously configured, based on the category of the report.



While the reports are being generated the progress of the current query can be viewed in either the left or right-hand side of the Status / Message bar, at the bottom of the application window. In addition, the overall progress (by Excel sheet within each report) can be viewed directly in the form.

5.12.3 Key Features of the Reconciliation Report Template

The reconciliation report template contains several specific features including:

1. Macro to generate a Table of Contents with hyperlinks to each worksheet.
2. Macro to setup hyperlinks between the anomaly references in the relevant reconciliation results and the Reconciliation Anomalies sheet, which contains the details of each anomaly.
3. Embedded contextual information (in the form of the expression definitions and column aliases from the Capture Definitions) for the **Group-By** and **Value** fields in Type 2, Type 3 and Processing Statistics query output.

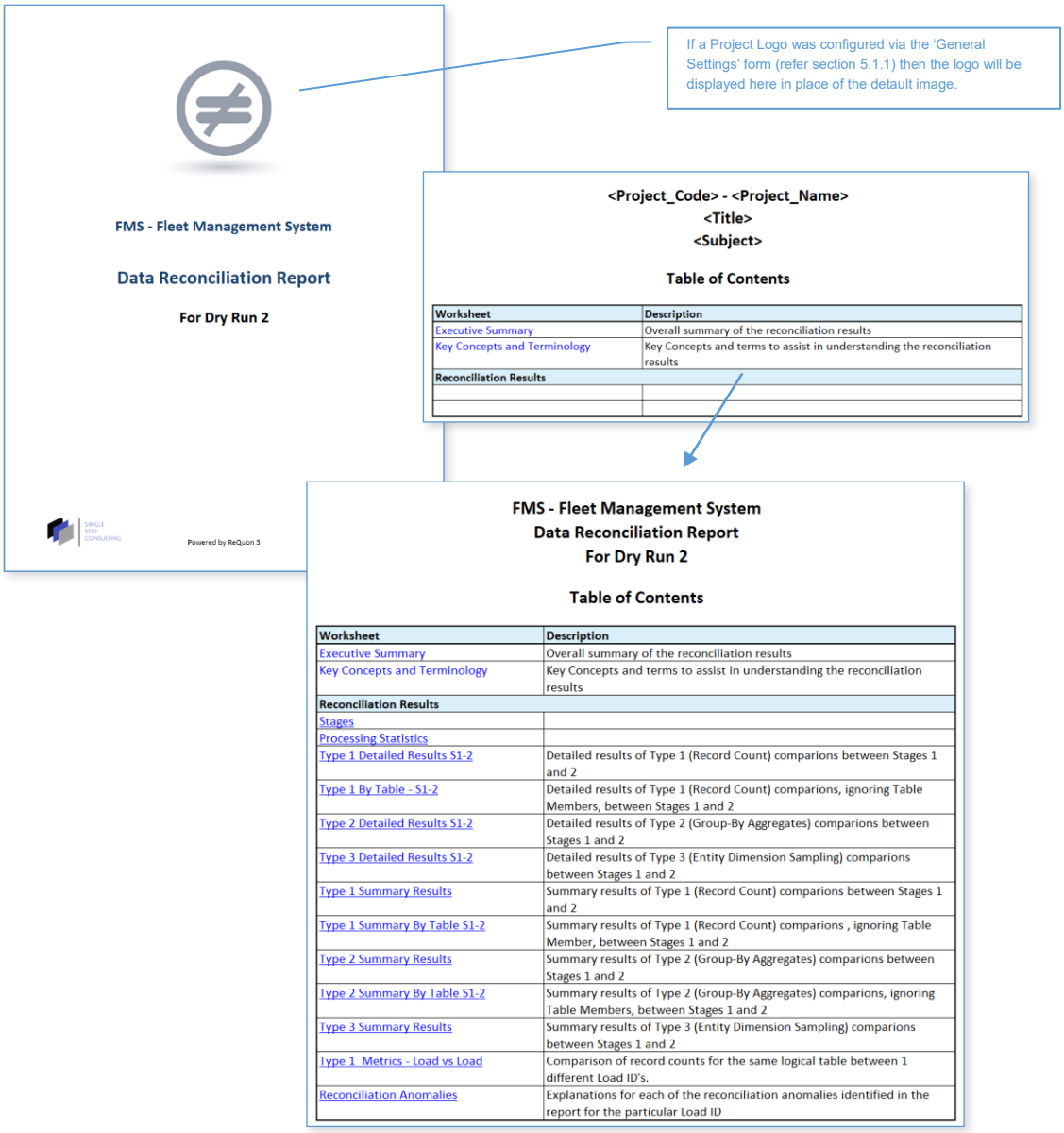
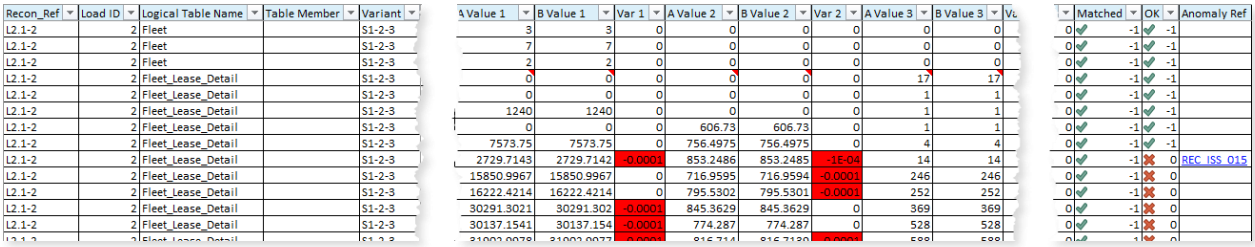
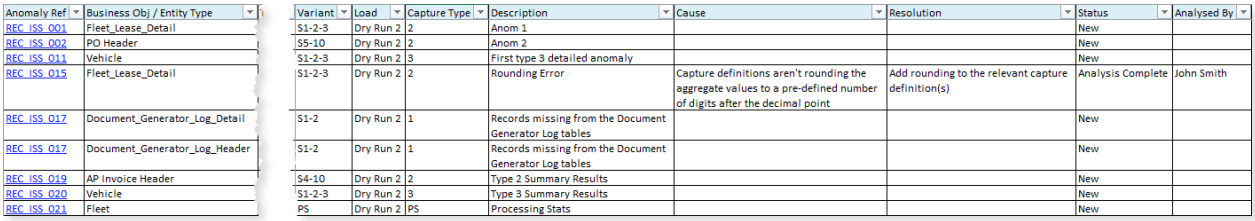


Figure 11 - Example Reconciliation Report Cover Page & Table of Contents



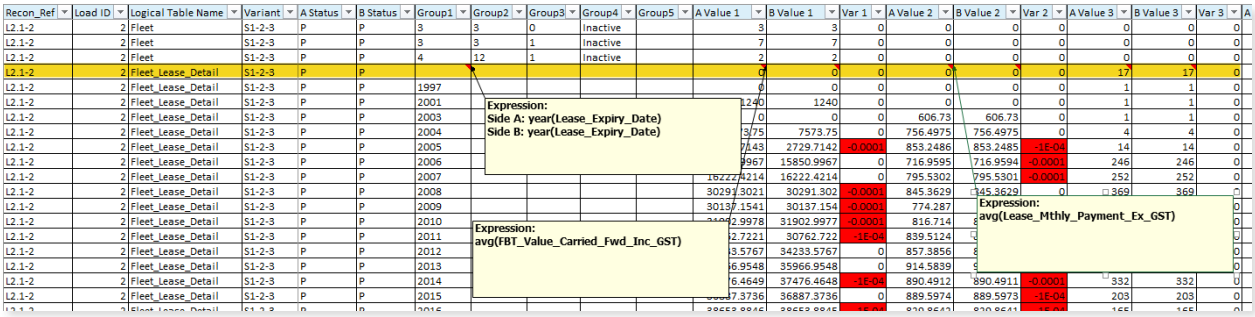
Recon_Ref	Load ID	Logical Table Name	Table Member	Variant	A Value 1	B Value 1	Var 1	A Value 2	B Value 2	Var 2	A Value 3	B Value 3	Matched	OK	Anomaly Ref
L2.1-2	2	Fleet		S1-2-3	3	3	0	0	0	0	0	0	-1	-1	
L2.1-2	2	Fleet		S1-2-3	7	7	0	0	0	0	0	0	-1	-1	
L2.1-2	2	Fleet		S1-2-3	2	2	0	0	0	0	0	0	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	0	0	0	0	0	0	17	17	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	0	0	0	0	0	0	1	1	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	1240	1240	0	0	0	0	1	1	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	0	0	0	606.73	606.73	0	1	1	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	7573.75	7573.75	0	756.4975	756.4975	0	4	4	-1	-1	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	2729.7143	2729.7142	-0.0001	853.2486	853.2485	-1E-04	14	14	-1	0	REC_ISS_015
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	15850.9967	15850.9967	0	716.9595	716.9594	-0.0001	246	246	-1	0	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	16222.4214	16222.4214	0	795.5302	795.5301	-0.0001	252	252	-1	0	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	30291.3021	30291.302	-0.0001	845.3629	845.3629	0	369	369	-1	0	
L2.1-2	2	Fleet_Lease_Detail		S1-2-3	30137.1541	30137.154	-0.0001	774.287	774.287	0	528	528	-1	0	

Figure 12 - Example of an Anomaly Hyperlink



Anomaly Ref	Business Obj / Entity Type	Variant	Load	Capture Type	Description	Cause	Resolution	Status	Analysed By
REC_ISS_001	Fleet_Lease_Detail	S1-2-3	Dry Run 2	2	Anom 1			New	
REC_ISS_002	PO Header	S5-10	Dry Run 2	2	Anom 2			New	
REC_ISS_011	Vehicle	S1-2-3	Dry Run 2	3	First type 3 detailed anomaly			New	
REC_ISS_015	Fleet_Lease_Detail	S1-2-3	Dry Run 2	2	Rounding Error	Capture definitions aren't rounding the aggregate values to a pre-defined number of digits after the decimal point	Add rounding to the relevant capture definition(s)	Analysis Complete	John Smith
REC_ISS_017	Document_Generator_Log_Detail	S1-2	Dry Run 2	1	Records missing from the Document Generator Log tables			New	
REC_ISS_017	Document_Generator_Log_Header	S1-2	Dry Run 2	1	Records missing from the Document Generator Log tables			New	
REC_ISS_019	AP Invoice Header	S4-10	Dry Run 2	2	Type 2 Summary Results			New	
REC_ISS_020	Vehicle	S1-2-3	Dry Run 2	3	Type 3 Summary Results			New	
REC_ISS_021	Fleet	PS	Dry Run 2	PS	Processing Stats			New	

Figure 13 - Example of an Anomaly Master List



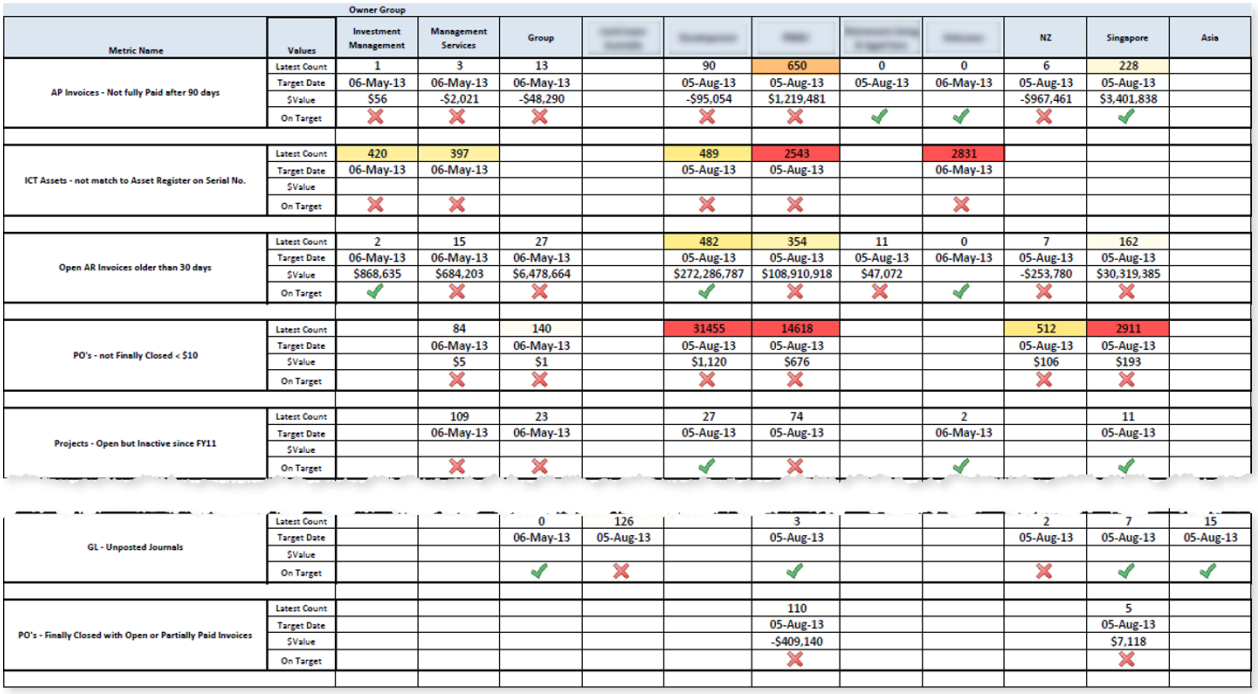
Recon_Ref	Load ID	Logical Table Name	Variant	A Status	B Status	Group1	Group2	Group3	Group4	Group5	A Value 1	B Value 1	Var 1	A Value 2	B Value 2	Var 2	A Value 3	B Value 3	Var 3
L2.1-2	2	Fleet	S1-2-3	P	P	3	3	0	inactive		3	3	0	0	0	0	0	0	0
L2.1-2	2	Fleet	S1-2-3	P	P	3	3	1	inactive		7	7	0	0	0	0	0	0	0
L2.1-2	2	Fleet	S1-2-3	P	P	4	12	1	inactive		2	2	0	0	0	0	0	0	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P						0	0	0	0	0	0	17	17	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	1997					0	0	0	0	0	0	1	1	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2001					1240	1240	0	0	0	0	1	1	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2003					0	0	0	606.73	606.73	0	1	1	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2004					7573.75	7573.75	0	756.4975	756.4975	0	4	4	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2005					2729.7143	2729.7142	-0.0001	853.2486	853.2485	-1E-04	14	14	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2006					15850.9967	15850.9967	0	716.9595	716.9594	-0.0001	246	246	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2007					16222.4214	16222.4214	0	795.5302	795.5301	-0.0001	252	252	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2008					30291.3021	30291.302	-0.0001	845.3629	845.3629	0	369	369	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2009					30137.1541	30137.154	-0.0001	774.287					0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2010						31902.9977	-0.0001	816.714					0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2011						30762.722	-1E-04	839.5124					0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2012						34233.5767	0	857.3856					0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2013						35966.9548	0	914.5839					0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2014						37476.4648	-1E-04	890.4912	890.4911	-0.0001	332	332	0
L2.1-2	2	Fleet_Lease_Detail	S1-2-3	P	P	2015						36887.3736	0	889.5974	889.5973	-1E-04	203	203	0

Figure 14 - Example of embedded contextual information for expressions

5.12.4 Key Features of the DQ Dashboard Template

The data quality dashboard template provides a high level, visual representation of the key DQ issues to complement the DQ Detailed Results Report.

1. Matrix Based Summary Dashboard (DQ Metrics vs Owner Groups).
2. DQ Detailed Results Report
3. DQ Metrics Trending Charts



Metric Name	Values	Owner Group: Investment Management	Management Services	Group	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NZ	Singapore	Asia
AP Invoices - Not fully Paid after 90 days	Latest Count	1	3	13		90	650	0	0	6	228	
	Target Date	06-May-13	06-May-13	06-May-13		05-Aug-13	05-Aug-13	05-Aug-13	06-May-13	05-Aug-13	05-Aug-13	
	$Value	$56	-$2,021	-$48,290		-$95,054	$1,219,481			-$967,461	$3,401,838	
	On Target	✗	✗	✗		✗	✗	✓	✓	✗	✓	
ICT Assets - not match to Asset Register on Serial No.	Latest Count	420	397			489	2543		2831			
	Target Date	06-May-13	06-May-13			05-Aug-13	05-Aug-13		06-May-13			
	$Value											
	On Target	✗	✗			✗	✗		✗			
Open AR Invoices older than 30 days	Latest Count	2	15	27		482	354	11	0	7	162	
	Target Date	06-May-13	06-May-13	06-May-13		05-Aug-13	05-Aug-13	05-Aug-13	06-May-13	05-Aug-13	05-Aug-13	
	$Value	$868,635	$684,203	$6,478,664		$272,286,787	$108,910,918	$47,072		-$253,780	$30,319,385	
	On Target	✓	✗	✗		✓	✗	✗	✓	✗	✗	
PO's - not Finally Closed < $10	Latest Count		84	140		31455	14618			512	2911	
	Target Date		06-May-13	06-May-13		05-Aug-13	05-Aug-13			05-Aug-13	05-Aug-13	
	$Value		$5	$1		$1,120	$676			$106	$193	
	On Target		✗	✗		✗	✗			✗	✗	
Projects - Open but Inactive since FY11	Latest Count		109	23		27	74		2		11	
	Target Date		06-May-13	06-May-13		05-Aug-13	05-Aug-13		06-May-13		05-Aug-13	
	$Value											
	On Target		✗	✗		✓	✗		✓		✓	
GL - Unposted Journals	Latest Count			0	126		3			2	7	15
	Target Date			06-May-13	05-Aug-13		05-Aug-13			05-Aug-13	05-Aug-13	05-Aug-13
	$Value											
	On Target			✓	✗		✓			✗	✓	✓
PO's - Finally Closed with Open or Partially Paid Invoices	Latest Count						110				5	
	Target Date						05-Aug-13				05-Aug-13	
	$Value						-$409,140				$7,118	
	On Target						✗				✗	

Figure 15 - DQ Summary Dashboard

Stage Name	Metric Name	Owner Grou	Priority	Baseline Date	Baseline Value	Latest Date	Latest Valu	Target Date	Target Valu	Avg Daily Rate	Req'd Run Rate	% Comple	Est. Completion Date	On Target
Oracle R11	AP Vendor - Mandatory Address Line 1 Missing	Singapore	High	1/28/2013	1	25-Feb-13	1	12-Apr-13	0	0	0.03	0.00%		No
Oracle R11	Mismatch between AP Retention and Custom Retention	Singapore	High	1/28/2013	922	25-Feb-13	922	12-Apr-13	0	0	27.46	0.00%		No
Oracle R11	PO's - not Finally Closed < $10	Singapore	High	1/28/2013	4580	25-Feb-13	3629	12-Apr-13	0	45.9103	108.1	20.76%	15-Jun-2013	No
Oracle R11	Projects - Open but Inactive since FY11	Singapore	High	1/28/2013	20	25-Feb-13	10	12-Apr-13	0	0.48276	0.3	50.00%	26-Mar-2013	Yes
Oracle R11	GL - Unposted Journals	Singapore	Low	1/28/2013	13	25-Feb-13	8	12-Apr-13	0	0.24138	0.24	38.46%	12-Apr-2013	No
Oracle R11	AP Invoices - Not fully Paid after 90 days	Singapore	Med	1/28/2013	387	25-Feb-13	340	12-Apr-13	0	2.26897	10.13	12.14%	22-Sep-2013	No
Oracle R11	Open AR Invoices older than 30 days	Singapore	Med	1/28/2013	237	25-Feb-13	254	12-Apr-13	0	-0.8207	7.57	-7.17%		No
Oracle R11	PO's - Finally Closed with Open or Partially Paid Invoices	Singapore	Med	1/28/2013	3	25-Feb-13	4	12-Apr-13	0	-0.0483	0.12	-33.33%		No

Figure 16 - DQ Detailed Results Report

Figure 17 - DQ Metrics Trending Charts

Appendix A Acronyms, Abbreviations & Definitions

Term	Meaning
Load ID	Reconciliation metrics need to be captured at an appropriate point in time to be effective. The concept of a **Load ID** is used to define a temporal reference point / baseline against which a set of metrics can be captured and subsequently reported. Refer to section 2.2.2 for more details.
ODBC	Open Database Connectivity (ODBC) is a standard application programming interface (API) for accessing database management systems (DBMS). Refer https://en.wikipedia.org/wiki/Open_Database_Connectivity
OLEDB	OLE DB (Object Linking and Embedding, Database, sometimes written as OLEDB or OLE-DB), an API designed by Microsoft, allows accessing data from a variety of sources in a uniform manner. The API provides a set of interfaces implemented using the Component Object Model (COM); it is otherwise unrelated to OLE. Microsoft originally intended OLE DB as a higher-level replacement for, and successor to, ODBC, extending its feature set to support a wider variety of non-relational databases, such as object databases and spreadsheets that do not necessarily implement. Refer https://en.wikipedia.org/wiki/OLE_DB
SQL Variant	A **SQL Variant** recognizes that different databases (e.g., Ms Access vs SQL Server vs Oracle, etc.) contain variations in supported SQL syntax. ReQuon comes with several pre-configured SQL Variants, however the user can configure several key syntax settings to support other variations.
Stage	Each **Stage** typically represents a physical layer from which reconciliation metrics can be captured and subsequently compared (reconciled) against metrics captured from a different **Stage**. Refer to section 2.2.1 for more details.
Variant	To differentiate between multiple set of metrics that need to be defined and captured from the **same Stage and Logical table**, the identifying keys used to identify a specific metric definition includes an additional label referred to as the "**Variant**". Refer to section 2.2.5 for more details.

Appendix B Microsoft ODBC Desktop Database Drivers – Scalar Functions

The Microsoft ODBC Desktop Database Drivers support the following **minimum** set of scalar functions:

CONCAT	LCASE	RIGHT
CONVERT	LEFT	RTRIM
CURDATE	LENGTH	SUBSTRING
CURTIME	LOCATE	UCASE
DATABASE	LTRIM	USER
DAYOFMONTH	MOD	YEAR
DAYOFWEEK	MONTH	

Further details can be found at:

https://learn.microsoft.com/en-us/sql/odbc/reference/appendixes/scalar-function-escape-sequence

https://docs.microsoft.com/en-us/sql/odbc/microsoft/scalar-functions

https://docs.microsoft.com/en-us/sql/odbc/reference/appendixes/appendix-e-scalar-functions

https://support.microsoft.com/en-us/office/odbc-scalar-functions-4d311cc4-aadb-486c-a11b-3bf8b77543f1

https://learn.microsoft.com/en-us/sql/odbc/reference/develop-app/interoperability-of-sql-statements

https://learn.microsoft.com/en-us/sql/odbc/reference/appendixes/appendix-c-sql-grammar

https://learn.microsoft.com/en-us/sql/odbc/reference/appendixes/odbc-escape-sequences

https://learn.microsoft.com/en-us/sql/odbc/reference/appendixes/date-time-and-timestamp-escape-sequences

Appendix C General Housekeeping

C.1 Ms Access Backend

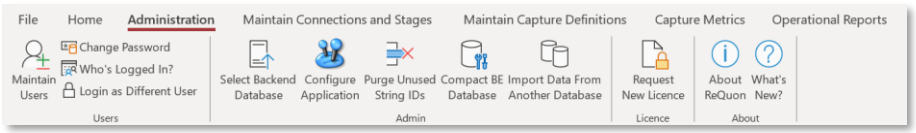
The primary housekeeping activities for a Microsoft Access database are:

- Ensure you take regular backups of your project backend database(s).
- Manage the size of the databases (which is limited to 2Gb).

> *"Database files can grow quickly as you use them, sometimes impeding performance. They can also occasionally become corrupt or damaged. You can use the Compact and Repair Database command to prevent or fix these problems. The compact process does not compress your data — it makes your database file smaller by eliminating unused space. The Compact and Repair Database command can also help improve performance of your database."* [2]

There are three ways to manage the size of your project backend databases:

1. Purge old metric capture results once they are no longer required. E.g. once the results have been published. There are 3 ways to purge results:
   - Purge results for an individual Capture Definition via the Capture Definition forms. Refer section 5.6.5.
   - Purge all results for a specific Load ID. Refer section 5.1.10 – "Maintain Load ID's".
   - Purge results for a specific Capture Definition and Load ID. Refer section 5.8.3 – "Review Metric Counts".
2. Purge Unused String IDs. ReQuon stores "Group By" strings once and then references them in the capture results, via a String Id. If the results that referenced a particular string have all been deleted, then the unused String Id's can also be deleted to free up space. To purge any unused strings from the currently selected backend database, navigate to the Administration menu and click on the "Purge Unused String IDs" option.



3. Run Compact and Repair on the backend (BE) database. To compact and repair the currently selected backend database, navigate to the Administration menu and click on the "Compact BE Database"

Note: This saves a copy of the original (uncompacted) database and then compacts the backend. The last 3 versions of the backend are retained in case you need to recover to an earlier version.

[2] https://support.microsoft.com/en-au/office/compact-and-repair-a-database-6ee60f16-aed0-40ac-bf22-85fa9f4005b2

Appendix D Common Pull-Thru and File-Splitting Scenarios

The following information lists some common use-cases / scenarios where Pull-Through Queries, with or without file-splitting functionality can be used

D.1 Scenario 1 - Prepare a set of data quality extracts to be sent to various Data Owners

Description of the Scenario

While the Data Quality Metrics capture definitions can provide the statistics, Pull-Through queries can be used to provide the record level details of the erroneous data. There are instances however, where different subsets of the data need to be actioned by different people. E.g. Data ownership may be split along organisational lines (e.g. Department or Business Unit) or geographical areas (such as States or Regions)

In this scenario, the file-splitting capabilities of Pull-Through queries can be used to prepare individual files for each of the relevant data owners.

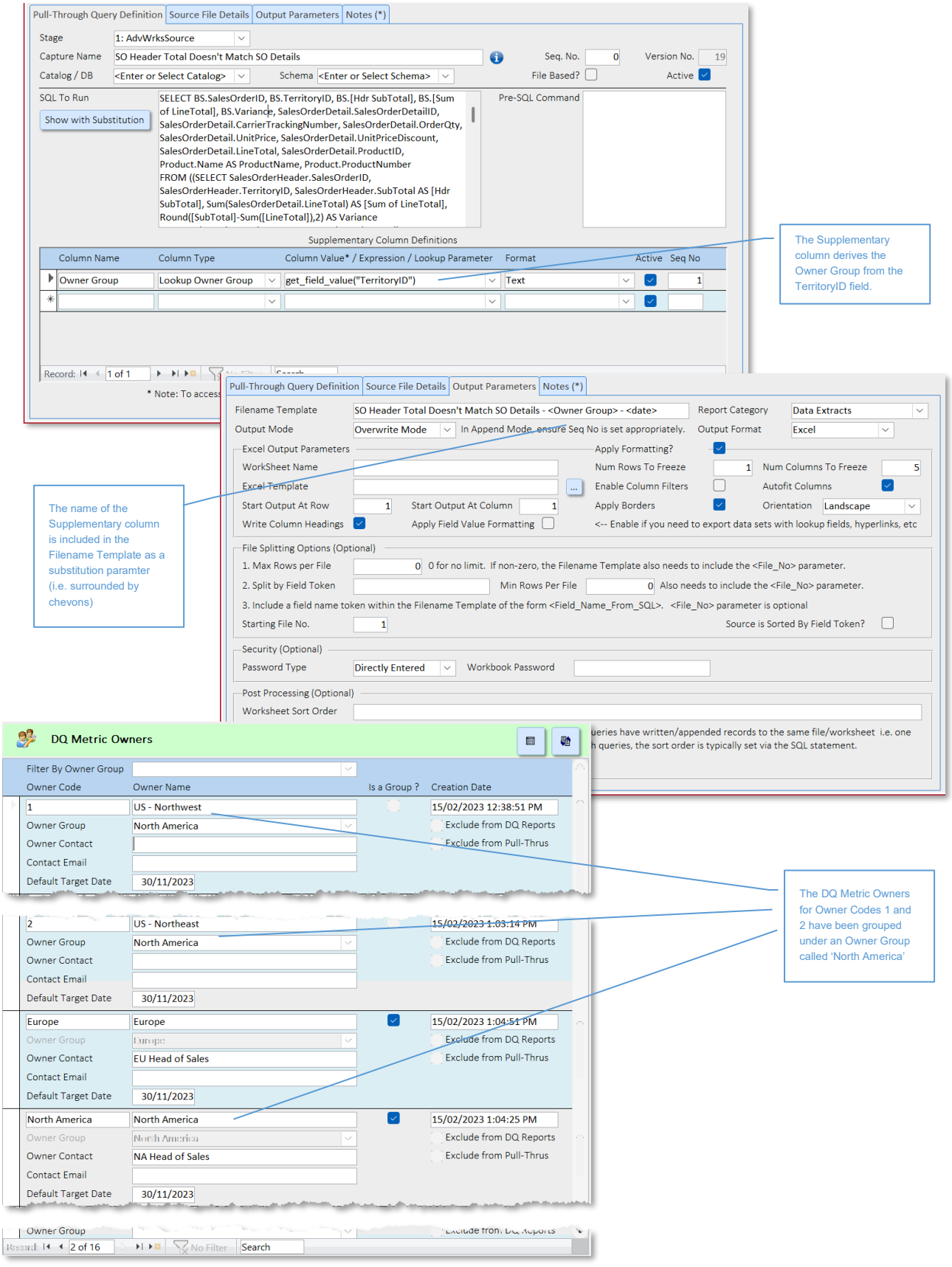
Implementation Instructions

The Pull-Through Query SQL will require inclusion of a field that represents each of the data owners or can be used to derive the owners. This field will represent the 'Owner Code' (refer section 5.1.15), and then multiple Owner Codes can be grouped under a common 'Owner Group'.

For example, if there is a Data Owner responsible for data related to East Coast of Australia, then a State_Code field would serve the role of the 'Owner Code' and the values of QLD, NSW, VIC & TAS could be grouped under an Owner Group of 'East_Coast' via the DQ Metric Owners form.

The nominated Owner Code can be used directly as a Field Token (refer File-Splitting options 2 and 3 in section), or indirectly via a Supplementary Column definition that reads the Owner Code and looks up the corresponding Owner Group. In the indirect scenario, the name of the Supplementary Column is included in the Filename Template (File-Splitting Option 3).

In the following example (from the ReQuon AdventureWorks2019 Demo project file), the 'Owner Code' as been identified as a field called 'TerritoryID', and a number of territories have been grouped under a number of 'Owner Group's.



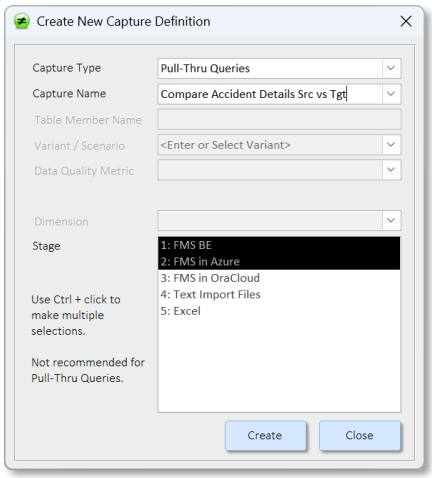
D.2 Scenario 2 - Compare two data sets at the field level by using Row Hashes.

Description of the Scenario

In this scenario, we want to compare every field in every row of 2 data sets to verify that they are identical.

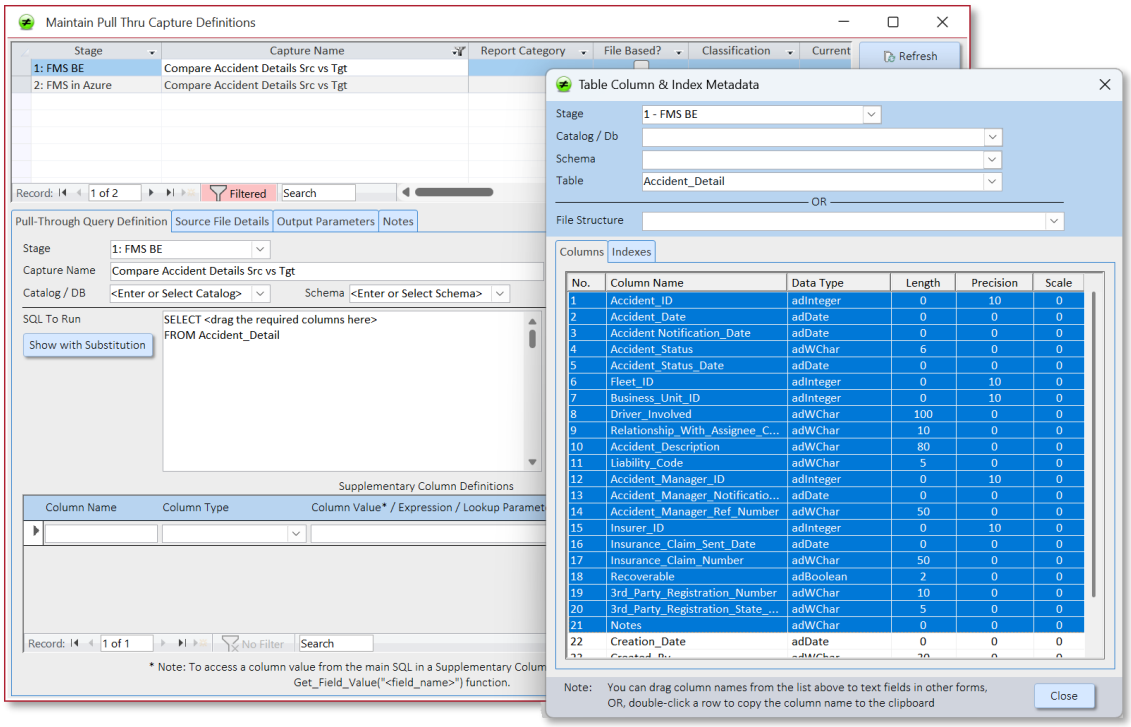
Implementation Instructions

This can be achieved with **a set of paired** Pull-Through queries that utilise the **Row Hash** functionality within the **Supplementary Columns**.



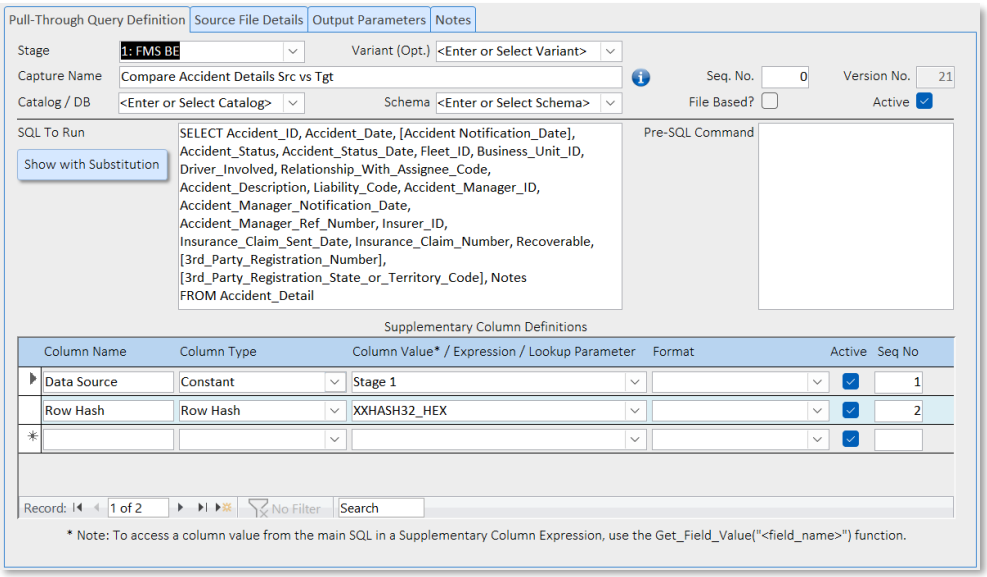
1. Create a paired set of Pull-Thru Queries which will have the same ‘Capture Name’ but for 2 different Stages

2. Select the first stage entry from the list and populate the ‘SQL to Run’ column with the set of columns that we want to compare. You can use the Table Column & Index Metadata popup form to select the select and drag fields from the relevant tables into the ‘SQL to Run’ field. Manually complete any final changes to the SQL as required.



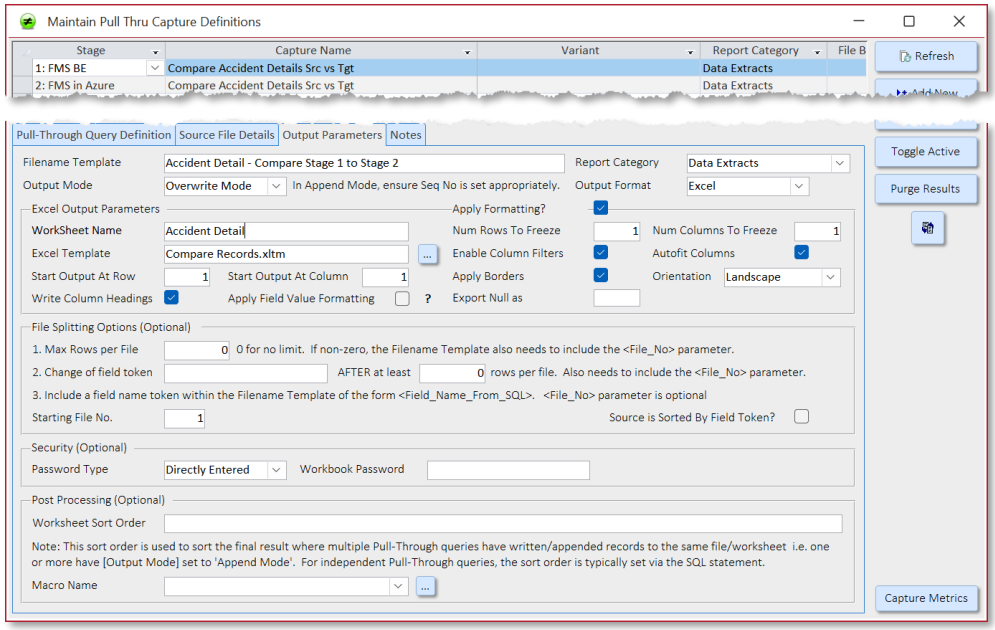
3. Add two (2) Supplementary Column definitions, one to identify the data source and one to calculate the row hash value.

> **Tip:** Either of the XXHash version (32 or 64 bit) represented in Hex format are an excellent choice for this scenario.
>
> **Tip:** If you chose to use the 'Compare Records.xltm' Excel template (refer to Step 4 below), and apply the "ApplyConditionalFormatting" macro (in Step 6) it expects the relevant Supplementary Column to be named "Row Hash".



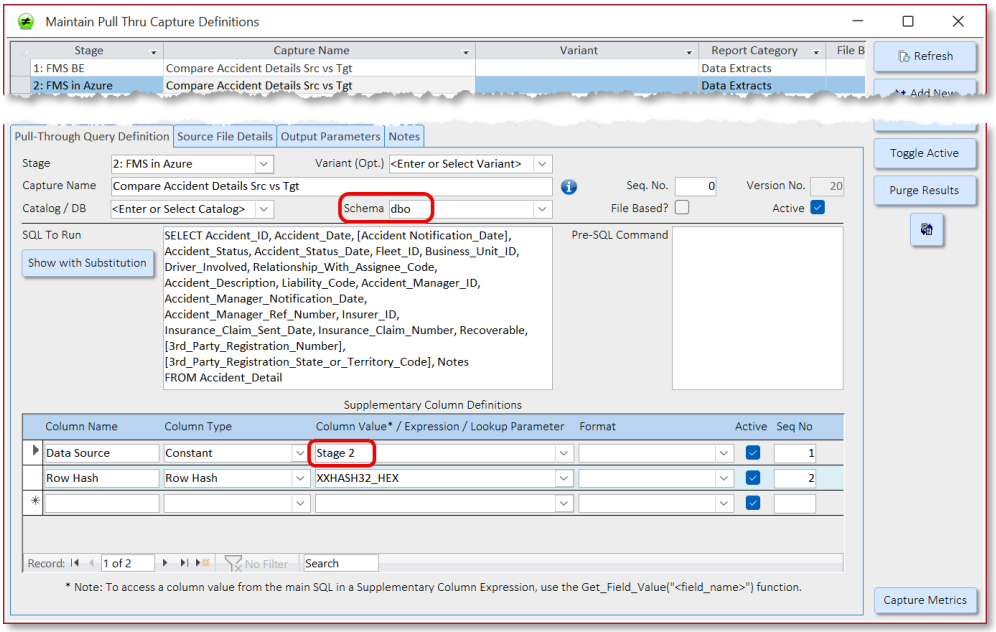
4. Go to the 'Output Parameters' tab and configure the following parameters

- Set the 'Filename Template' – the same name will apply to both Stages. If no name is provided it will default to the value of the 'Capture Name'.
- Set the 'Report Category'. If you want to save the output of these queries in their own folder, then you could add a new Report Category such as "Hash Comparisons" and select that.
- Ensure the 'Output Mode' for at least one of the stages (usually the first), is set to "Overwrite Mode"
- ReQuon comes with a pre-defined 'Excel Template' called "Compare Records.xltm". While optional, this template contains some conditional formatting and macros that help visualize the results and is highly recommended.
- Ensure the 'Write Column Headings' checkbox is ticked.
- For the first (overwrite) stage **do NOT populate** the 'Worksheet Sort Order' or 'Macro Name' parameters as they only need to be applied to the second (final) stage.
- You can also set any other optional parameters such as 'Worksheet Name', formatting related parameters, etc., as desired.



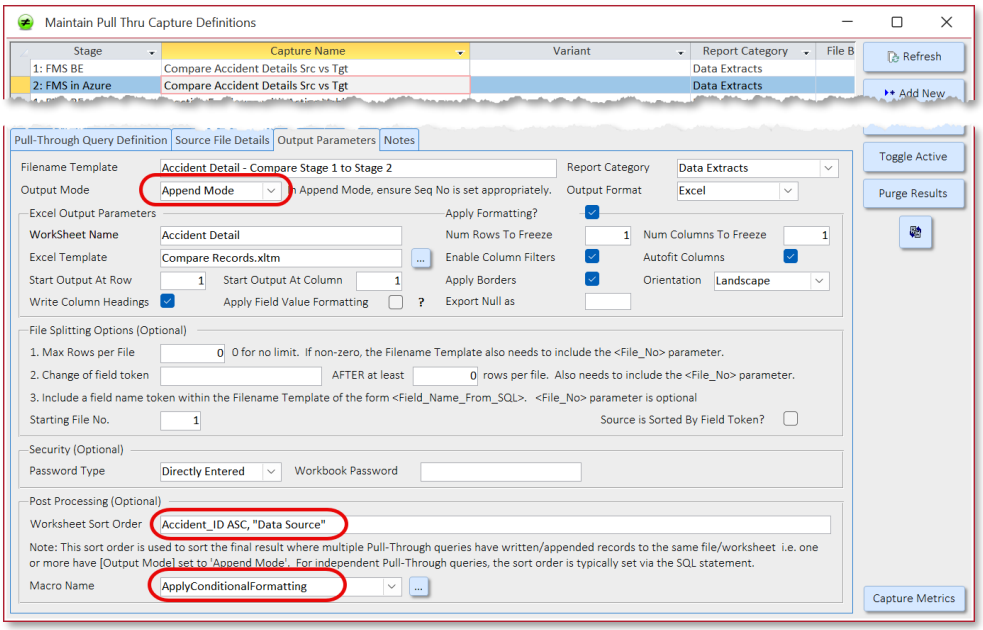
5. Now select the second Stage capture definition and repeat steps 2 and 3 for this stage.
   - Ensure the same set of fields are included in the same order as the first stage.
   - If there have been any transformations between the 2 stages, you will have to include logic in one, or both, of the 'SQL to Run' definitions to normalise the values. For example, a boolean value of "Y" or "N" in one stage may have been transformed into True or False in the other. Other examples may include date formats, different status code values, etc.

Tip: You can copy and paste the Supplementary Column definitions from one stage to the other and then edit them as required



6. Now go to the 'Output Parameters' tab for the **second stage**. Most of the values would have already been populated via Step 4, however, make the following changes: -

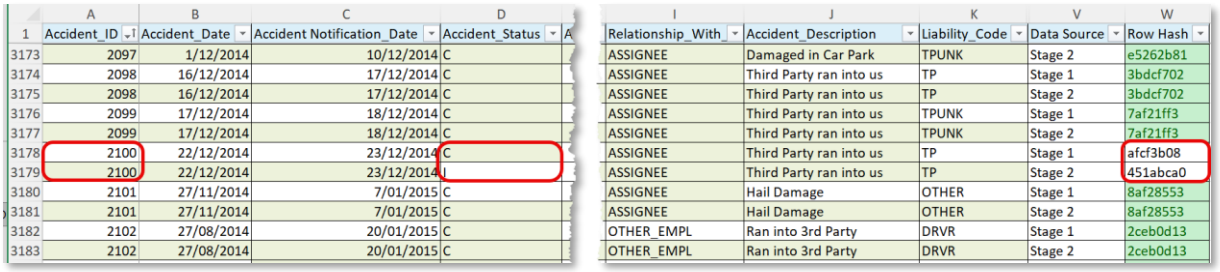
   - Change the 'Output Mode' to "Append Mode"
   - Populate the 'Worksheet Sort Order' parameter, with a combination of the fields that represent the primary key of the base data sets, followed by the name of the Supplementary Column representing the Data Source / Stage.
   - If you chose to use the 'Compare Records.xltm' template, click on the ... button to the right of the 'Macro Name' parameter (to populate the dropdown with a list of the macros in the template) and then select the "ApplyConditionalFormatting" macro.



How Does It Work?

- The first Pull-Through query creates / overwrites the output file, writing out the column headings from the base SQL, as well as the Supplementary Columns.
- The 'Row Hash' column is calculated based on the full set of field values from the base SQL i.e. it excludes the Supplementary Columns.
- Once all the rows have been written, the output file is closed
- When the second Pull-Through query runs, it will open the existing output file, locate the last row, and start outputting rows. The 'Write Column Headings' flag is ignored when outputting to an existing file.
- Once all the rows from the second data set have been written, it will now sort the full set of rows so that the corresponding rows (based on the primary key) are positioned next to each other.
- The final step runs the 'ApplyConditionalFormatting' macro. The macro applies two types of formatting:
    - Firstly, it will look for a column named 'Row Hash'. If it finds the column, it will apply formatting to highlight duplicate hash values i.e. if the pair of rows are identical then the hash values will be duplicated. Any hash value NOT highlighted indicates an issue to be investigated.
    - For the remaining columns, it applies formatting that compares each cell to the cell above. If they are the same, the lower cell will be highlighted. Non-highlighted cells do NOT necessarily indicate an issue, given we expect records with different primary keys to contain some different values. But where the hash values are different, the secondary formatting can assist in identifying which of the relevant column values may be the cause of the difference.

In the following example, the mismatching 'Accident_Status' values for records with Accident_ID = 2100, are the cause of the mismatching 'Row Hash' values.



	A	B	C	D	I	J	K	V	W
1	Accident_ID	Accident_Date	Accident Notification_Date	Accident_Status	Relationship_With	Accident_Description	Liability_Code	Data Source	Row Hash
3173	2097	1/12/2014	10/12/2014	C	ASSIGNEE	Damaged in Car Park	TPUNK	Stage 2	e5262b81
3174	2098	16/12/2014	17/12/2014	C	ASSIGNEE	Third Party ran into us	TP	Stage 1	3bdcf702
3175	2098	16/12/2014	17/12/2014	C	ASSIGNEE	Third Party ran into us	TP	Stage 2	3bdcf702
3176	2099	17/12/2014	18/12/2014	C	ASSIGNEE	Third Party ran into us	TPUNK	Stage 1	7af21ff3
3177	2099	17/12/2014	18/12/2014	C	ASSIGNEE	Third Party ran into us	TPUNK	Stage 2	7af21ff3
3178	2100	22/12/2014	23/12/2014	C	ASSIGNEE	Third Party ran into us	TP	Stage 1	afcf3b08
3179	2100	22/12/2014	23/12/2014	I	ASSIGNEE	Third Party ran into us	TP	Stage 2	451abca0
3180	2101	27/11/2014	7/01/2015	C	ASSIGNEE	Hail Damage	OTHER	Stage 1	8af28553
3181	2101	27/11/2014	7/01/2015	C	ASSIGNEE	Hail Damage	OTHER	Stage 2	8af28553
3182	2102	27/08/2014	20/01/2015	C	OTHER_EMPL	Ran into 3rd Party	DRVR	Stage 1	2ceb0d13
3183	2102	27/08/2014	20/01/2015	C	OTHER_EMPL	Ran into 3rd Party	DRVR	Stage 2	2ceb0d13

D.3 Scenario 3 - Prepare a set of data load files that contain size & content constraints.

Description of the Scenario

Pull-Through queries can also be used to prepare data load files that have specific contraints. One real-life example required employee detail load files prepared with the following characteristics:

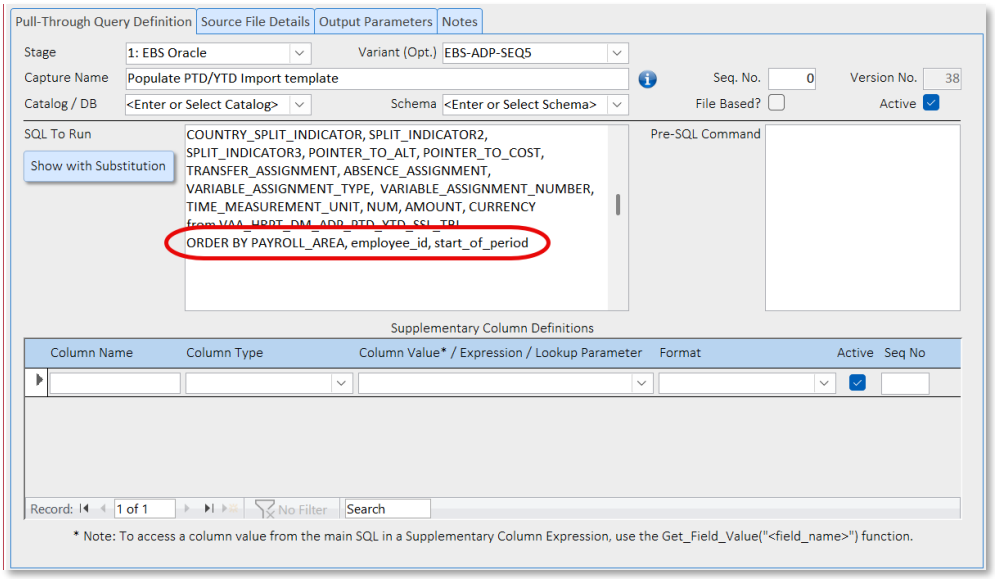
1. Each file can only contain data related to a single Payroll Area (fortnightly vs monthly)
2. A single file can contain records for multiple employees, however, records for a single employee must NOT be split across multiple files.
3. The maximum number of records in each file should be approximately 100,000.
4. The output is required to be populated in a pre-defined template with a fixed number of fields

Initial analysis indicated that the maximum expected number of records for any single employee was around 1000.

Implemention Instructions

This scenario will combine 2 types of file splitting: Option #2 and Option #3

1. Setup the Pull-Through query ensuring that the 'SQL to Run' sorts the source data by Payroll_Area, Employee_Id (and optionally start_of_period).

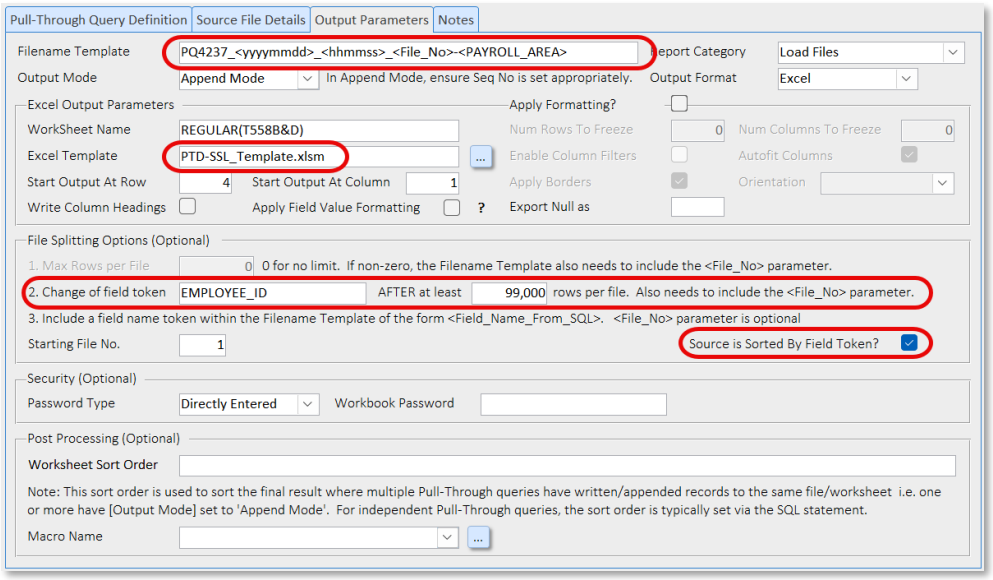


2. Configure the Output Parameter settings

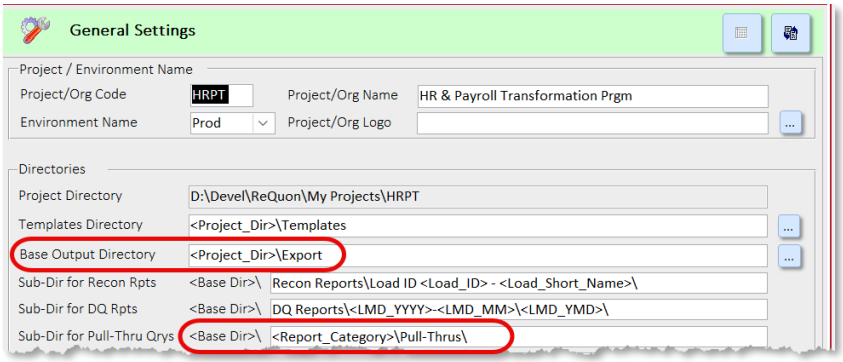
- Ensure the output 'Filename Template' includes the <File_No> and <PAYROLL_AREA> substitution parameters.
- Set the required 'Excel Template' that defines the structure of the output files. Note: The template needs to be copied into either the shared Templates directory (under the

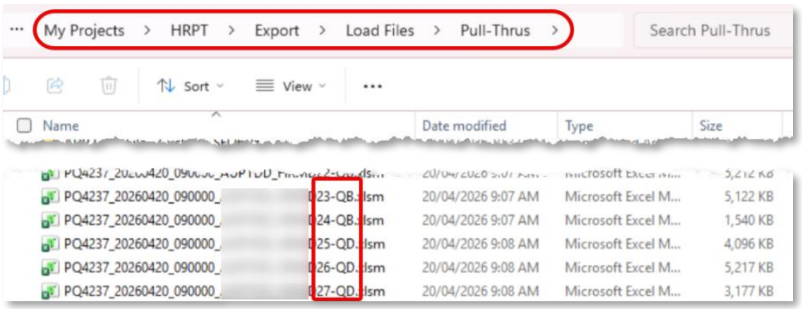
ReQuon application directory), OR, a project-specific templates directory configured under the Administration / Configure Application / General Settings form.

- Set the 'Split by Field Token' parameter to the name of the employee number column i.e. 'EMPLOYEE_ID'
- Set the 'Min Rows Per File' parameter to 99,000.
- Ensure the 'Source is Sorted by Field Token?' checkbox is ticked



The output files will be written out to the sub-directory configured under the General Settings for Pull-Thru queries. The sub-directory definition can utilise a number of substitution parameters including the <Report_Category> defined in the Pull-Thru query output parameters.





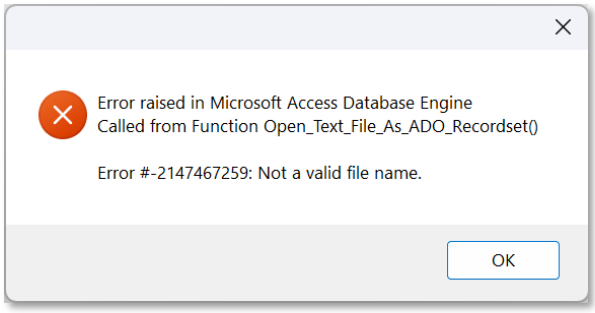
Appendix E Troubleshooting

E.1 Connectivity

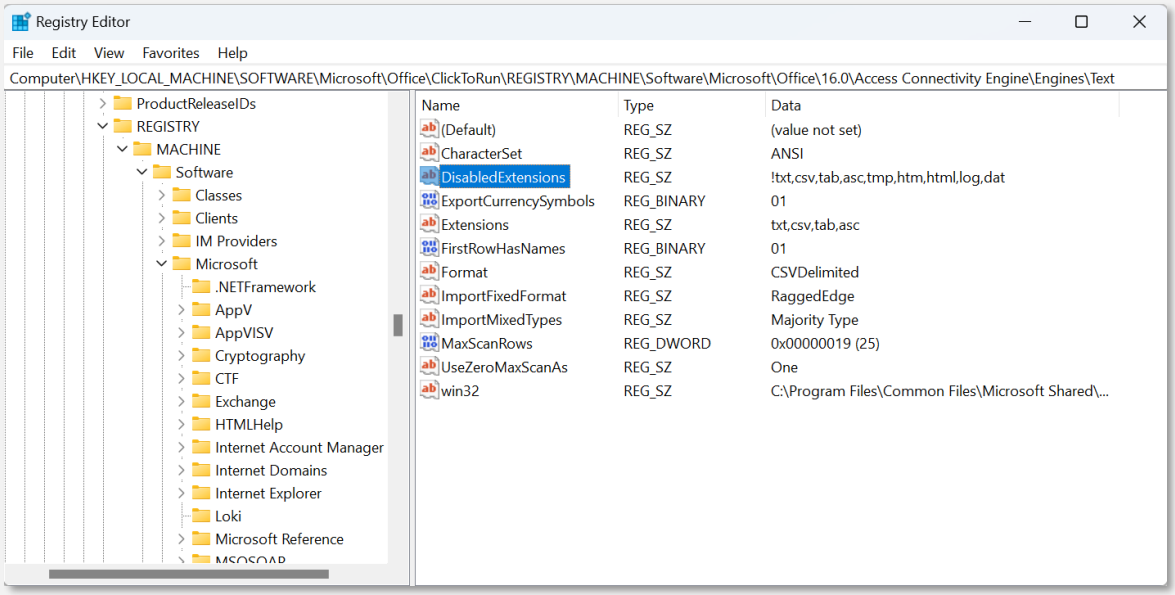
There are a wide variety of reasons / causes that result in the inability to connect to a data source. The following table provides some common causes. Wherever possible, refer to the vendor documentation for the relevant ODBC or OLEDB drivers.

Symptom	Possible Causes / Resolution
Unable to establish a connection to a data source	• Wrong driver installed or wrong version. Ensure the bitness (32-bit vs 64-bit) of the driver you have installed matches the bitness of the version of Ms Office / Access that you installed. Unless a specific reason exists, generally aim to install the latest version of the driver.
	• Invalid / mis-formed connection string. Check the relevant documentation to ensure the correct parameter names and values have been specified. Some drivers can be quite finicky with the structure of the connection string. E.g. an unexpected space char can cause some drivers to error.
	• Invalid login / authentication credentials. Depending upon the security regime setup for the specific data source, verify that you have the correct username/password details or authentication keys / certificates, etc.
	• Firewall settings are blocking the relevant IP-address(es) and/or ports. Consult with the data source and/or networking support teams to determine if this may be the cause of the error.
Connection works sometimes and fails other times	• Instability of the source system. Check if there are any known issues with the stability of the target data source
	• The Connection is via a load balancer, and not all the required ip-address / ports have been opened. Consult with the data source and/or networking support teams to determine if this may be the cause of the error.

E.2 Importing Metadata

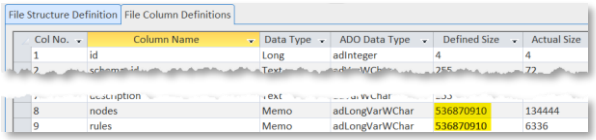
Symptom	Possible Causes / Resolution
Import fails with “Not a valid file name.” but the filename is correct. 	• The file extension is not in the list of accepted extensions. Resolution: Refer to Much ADO About Text Files | Working with Custom File Extensions - https://learn.microsoft.com/en-us/previous-versions/windows/internet-explorer/ie-developer/scripting-articles/ms974559(v=msdn.10)#working-with-custom-file-extensions While the above link refers to the older Jet database engine, the same applies to the newer ACE (Access Connectivity Engine). Depending upon your version of Ms Access installed, you may need to search for the applicable registry path.

Example configuration within the register to allow the processing of *.dat files via the Text reader.



E.3 Capture Metric Failures

Symptom	Possible Causes / Resolution
Metric capture fails with an error such as: Syntax error (missing operator) in query expression 'SUM(75thpercentile)'	• Field name commences with a numeric character. E.g. 75thPercentile. Some drivers (such as the csv text driver) don't expect field names to commence with a number, so an expression such as `SUM(75thPercentile)` will cause the capture to fail. To resolve this, ensure the field name is enclosed in the appropriate field name delimiters. E.g. `SUM([75thPercentile])`
Use of the standard format function on a date field within a text-based source file returns null	• The format function available in Ms Access does not error when processing Text files but returns null when trying to format a date field to 'yyyy-mm'. In this scenario, the same result can be achieved with the follow expression: YEAR(<datefield>) & '-' & RIGHT('0' & MONTH(<datefield>),2)
Formatting of field data types for text file imports not as expected based on the schema.ini definition	• The schema.ini has not be saved after making changes to the field type / formatting definitions in the database. Re-save the schema.ini file, making sure that you select ALL the files that are referenced in the capture definition. Repeat, every time new files are added to the capture definition.
	• One or more of the file names contain special characters that cause the txt driver to NOT find the relevant entry in the schema.ini file. e.g. a filename containing % (such as %3A being a html encoded replacement for the colon : character) will cause the lookup within the schema.ini file to not find a match. The workaround is to rename the file to remove any special characters. Note: Remember to re-generate / save the schema.ini file using the new file names.

Symptom	Possible Causes / Resolution
Error “In the text file specification ‘<filename>’ the Col<*n*> option is invalid”	• If the imported metadata included a memo / adLongVarWChar field with a Defined Size > 99999 then upon execution of the Capture Definition you will get this error.  There are two options to address this issue: 1. Manually modify the relevant column definitions and change the Defined Size to 99999 or less. 2. If there were no overridden column definitions, then you can delete all of the File *Column* Definitions for the specific file and simply rely on the File *Structure* Definition to correctly import the data.