



# Qlik Google BigQuery Connector Installation and User Guide

*Qlik Google BigQuery Connector Version 1.0*

*Newton, Massachusetts, July 2014*

*Authored by QlikTech International AB*

Copyright © QlikTech International AB 2014, All Rights Reserved Under international copyright laws, neither the documentation nor the software may be copied, photocopied, reproduced, translated or reduced to any electronic medium or machine-readable form, in whole or in part, without the prior written permission of QlikTech International AB, except in the manner described in the software agreement.

QlikTech®, Qlik® and QlikView® are registered trademarks of QlikTech International AB. All other company names, products and services used herein are trademarks or registered trademarks of their respective owners.

---

## Contents

1	Qlik Google BigQuery Connector .....	3
2	Install the Qlik Google BigQuery Connector.....	3
3	Create a Google BigQuery Connection.....	4
4	Create a Load Statement.....	7
5	Log Information.....	7

---

# 1 Qlik Google BigQuery Connector

The Qlik Google BigQuery Connector provides a means for the QlikView business discovery platform to efficiently load data into a QlikView application from a Google BigQuery dataset. The metadata is used by the **Create Select Statement** dialog to specify the table and columns to select for the SQL SELECT statement.

The Google BigQuery Connector includes the following significant features that enhance its performance and functionality:

- **Nested record support:** The Nested (Record) data type is now supported in queries.
- **Persistent destination tables:** The Connector now allows for specification of persistent destination tables instead of temporary tables for storing datasets.
- **Connection fault protection:** Provides settings to control handling of connection failures and interruptions.
- **Query cached results:** It is now possible to query an existing cache rather than repeating a query to repopulate a temporary table.
- **Google BigQuery integer type** and **64-bit integers** are now supported.

## 2 Install the Qlik Google BigQuery Connector

1. Download the `QvBigQueryConnector_setup.exe` file from the Qlik download site.
2. Run the `QvBigQueryConnector_setup.exe`.

The Qlik Google BigQuery Connector InstallShield Wizard sets up the installation environment and completes the installation .

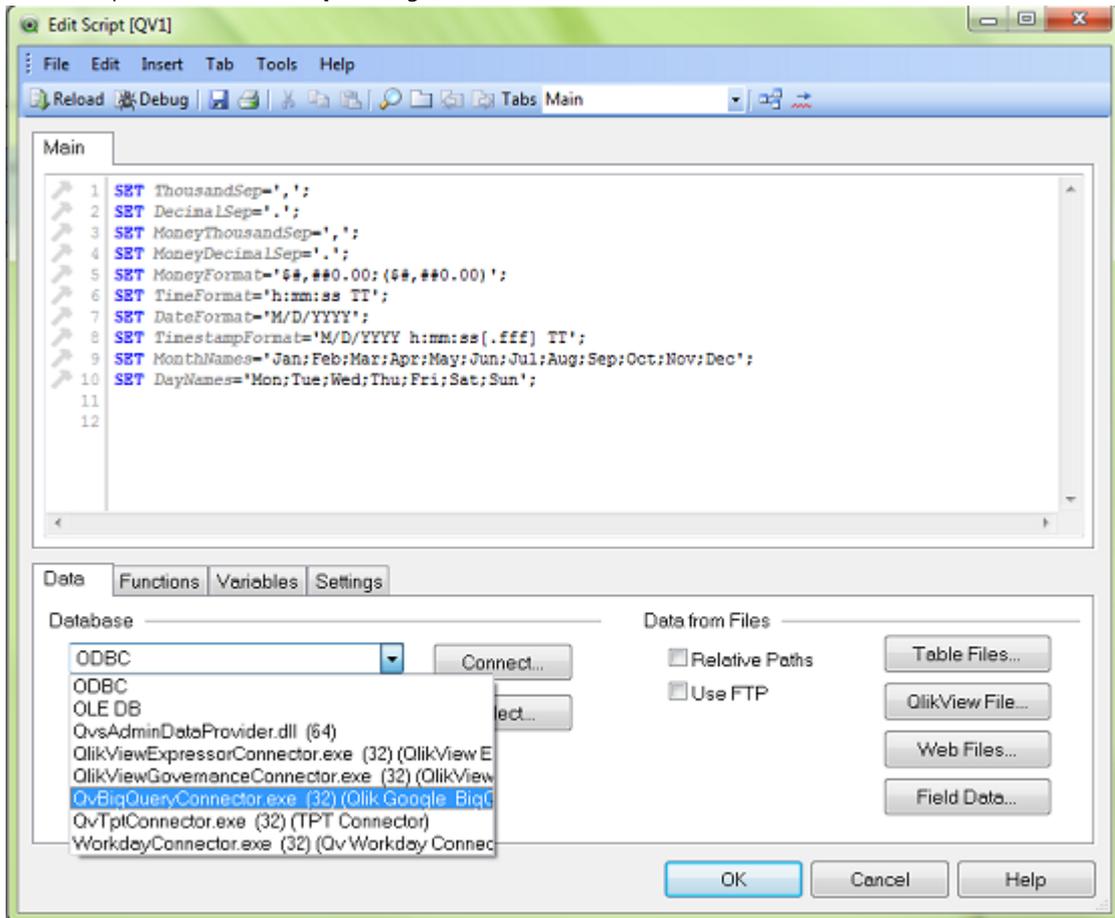
The Connector is then available from the QlikView Desktop with the name "QvBigQueryConnector.exe."

3. Open the QlikView11 Desktop application.  
If QlikView is open when you install the Google BigQuery Connector, you must close and reopen it in order for QlikView to recognize the Connector.

### 3 Create a Google BigQuery Connection

Access to the Qlik Google BigQuery Connector is provided by the QlikView business discovery platform. QlikView applications connect to data sources through the Script Editor.

1. Open a new or existing QlikView application.
2. Click **Edit Script...** on the **File** menu in the main QlikView window.
3. Select QvBigQueryConnector.exe from the drop-down list next to **Connect...** on the **Data** tab in the lower part of the **Edit Script** dialog.



4. Click **Connect...** to open the **Google BigQuery Connection** dialog.

The screenshot shows the 'Google BigQuery Connection' dialog box. It has a title bar with a close button. The dialog is organized into sections: 'Account properties' with input fields for 'Client id' and 'Client secret', a 'Timeout' field set to '60' with a 'Seconds' label, and an 'Authorize...' button; 'Destination table options' with a 'Destination table' field, a 'Select...' button, a 'Write preference' dropdown menu set to 'Write if empty', and an unchecked 'Allow large results' checkbox; 'Options' with a 'Query priority' dropdown menu set to 'Interactive', a checked 'Use cached results' checkbox, 'Connection restore attempts' set to '5', and 'Connection restore delay' set to '60' with a 'Seconds' label; and two empty text boxes for 'Project' and 'Dataset'. At the bottom are 'OK', 'Cancel', and 'Help' buttons.

5. Enter the Client id and Client secret credentials for the Google BigQuery account.  
The Client id and Client secret credentials are not the account name and password used to log into the Google account. The credentials are for the BigQuery projects available to the Google account.
6. Change the Timeout setting or use the default setting of 60 seconds.  
The timeout value specifies how much time (in seconds) the Connector waits for **Accept** to be clicked in the browser window opened after the user has clicked **Authorize** (see steps 7 and 8). The minimum Timeout setting is 10 seconds and the maximum is 300 seconds.
7. Click **Authorize** to verify that the access credentials are valid.
8. Click **Accept** on the Google account page, which displays in your web browser after **Authorize** has been clicked.  
If the credentials validate correctly, the BigQuery projects are listed in the **Google BigQuery Connection** dialog.  
If the Client id is invalid, an error message appears in the web browser instead of the Google account page.  
If the Client secret is invalid, the Connector displays an error message box after you click **Accept** on the Google account page.  
If neither the Google account page nor an error message appears in your web browser and you get a Timeout exceeded message, the Timeout setting is too short and must be increased.

- 
9. Click the **Select** button to specify the Destination Table if you want to create a persistent table to hold the data from the query.  
If a destination table is not specified, then the data is stored in a temporary table that exists for about 24 hours. A temporary table is accessible only for the current query, and it cannot be accessed for different query.  
A destination table, on the other hand, remains until the user deletes it. Because a destination table persists, it can be used for subsequent queries.  
You must specify a destination table if you intend select **Allow Large Results** in step 11.
  10. Select the Write Preference if you have specified a destination table.  
If you are writing to a destination table, it can be used for multiple queries.  
**Write if empty** means write the query results to the named destination table if the table is empty.  
**Append to table** allows the new query results to be added to the destination table if it already contains data. But the data from the new query must match the schema of the existing data in the table. If it does not, an error is produced.  
**Overwrite table** allows the new query results to replace existing data in the destination table. The new results do not have to match the schema of the existing data in the table. The table is recreated.
  11. Select **Allow Large Results** if you have specified a destination table and intend to exceed Google's limits on the size of temporary tables.
  12. Specify the Query Priority by clicking either **Interactive** or **Batch**.  
By default, Google BigQuery runs interactive queries. That is, the query is executed as soon as possible. Batch queries are placed in a queue and run as soon as idle resources are available. If batch query is selected and service is not available for three hours, the query priority automatically switches to Interactive.
  13. Select **Use Cached Results** if you want the query to pull data from a temporary table.  
If you are using a temporary table, that is, you have not specified a destination table, Google BigQuery can use the results of a previous query if you elect to use the cached results. If you clear **Use Cached Results**, the query is repeated and the temporary table repopulated. If you elect to use cached results, the original data source is checked to verify that the schema still matches the temporary table. If it does not, the temporary table is repopulated and the cached results are not used, even though **Use Cached Results** is selected.
  14. Specify the number of connection restore attempts.  
This is the number of times attempts are made to restore a connection if it is lost during a query. Once the connection is restored, the number of restore attempts is reset to the value specified.
  15. Specify the number of seconds to wait between attempts to restore a connection if it is lost during a query.
  16. Select a project.
  17. Select a dataset from the project.
  18. Click **OK** to create the connection string used when sending the load statement to Google BigQuery.

---

## 4 Create a Load Statement

1. Click **Select...** beneath **Connect...** on the **Data** tab in the lower part of the **Edit Script** dialog.
2. Select a database table from the list in the **Create Select Statement** dialog.
3. Select the fields to load from the **Fields** list.  
The SQL SELECT statement that will be used to load data from the Google BigQuery database displays under the Script tab in the bottom portion of the dialog.
4. Click **OK** when you are finished making your selections.  
The SQL SELECT statement does not include a WHERE clause, and one cannot be added in the Load Script editor. The SQL SELECT statement for the Google BigQuery Connector does not support the addition of WHERE clauses.

The Google BigQuery Connector can be used to load data with QlikView Direct Discovery. To use Direct Discovery, you must add the following set statements to the beginning of the script:

```
SET DirectDistinctSupport=false;  
SET DirectIdentifierQuoteChar='[]';  
SET DirectIdentifierQuoteStyle='big query';
```

SELECT DISTINCT is not supported by Google BigQuery. Also, Direct Discovery quotes field and table names with double quotation marks, but Google BigQuery SQL requires square brackets instead.

## 5 Log Information

The Google BigQuery Connector's log file is located at:

```
{ProgramData}\QlikTech\Custom Data\QvBigQueryConnector\Log
```

Where {ProgramData} is %ALLUSERSPROFILE%.