



BigQuery Connector for Qlik Sense User Guide



Content

- Configuring Google Access
- Creating a new connection
- Selecting and loading data
- Notes
- Tips and Tricks

This document describes how to use the BigQuery Connector for Qlik Sense. Please contact support at qlik.support@stretch.dk for help.

Creating Connections

Creating a new connection

The BigQuery connector for Qlik Sense has two modes, direct and fast.

- Direct mode is intended for the most use cases for most users, the script is simple and easy to read and modify as needed
- Fast mode is intended for large data transfers and production use for example in QVD-generation apps. The fast mode is as a rule of thumb twice as fast as the direct mode. But this mode requires additional initial configuration, and the load script is more verbose and require more expertise to modify if needed.

This document show the configuration of both modes, but it is not required to configure both for the connector to work. If the fast mode is not needed, then the step in this guide can be omitted.

This document assumes the users are familiar with basic Qlik tasks e.g. creating connections. If help is need with these tasks, please contact us for assistance.

Support: glik.support@stretch.dk

Configuring Google Access

1. Create a service account on the google cloud console
 - <https://cloud.google.com/iam/docs/creating-managing-service-accounts>
2. Create a JSON key for the service account
 - <https://cloud.google.com/iam/docs/creating-managing-service-account-keys>
3. Go the "iam-admin"
 - <https://console.cloud.google.com/projectselector/iam-admin>
4. Select the cache project and assign the following roles to the service account
 - BigQuery Data Editor
 - BigQuery Job User
 - Storage Object Admin
5. Select the data project(s) and assign the following roles to the service account
 - BigQuery Data Viewer

Edit permissions

Member	Project
qliksenseminimuser@stretch-retail-demo.iam.gserviceaccount.com	stretch-cache

- Role: BigQuery Data Editor
Access to edit all the contents of datasets
- Role: BigQuery Job User
Access to run jobs
- Role: Storage Object Admin
Full control of GCS objects.

[+ ADD ANOTHER ROLE](#)

[SAVE](#) [CANCEL](#)

Edit permissions

Member	Project
qliksenseminimuser@stretch-retail-demo.iam.gserviceaccount.com	Stretch Retail Demo

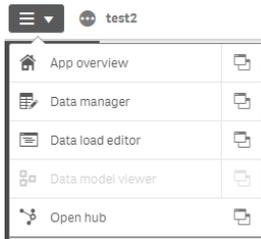
- Role: BigQuery Data Viewer
Access to view datasets and all of their contents

[+ ADD ANOTHER ROLE](#)

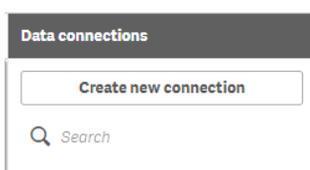
[SAVE](#) [CANCEL](#)

New connection – create

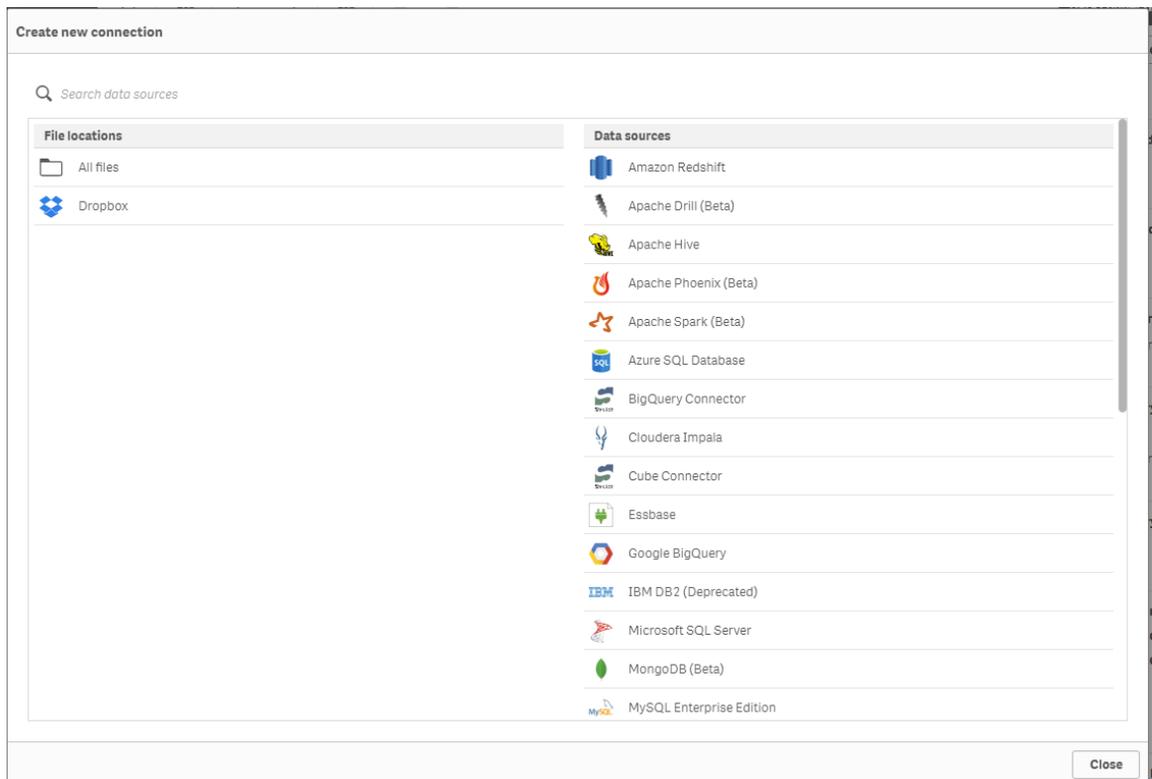
1. Go to the "data load editor" in the upper left corner:



2. Click the button "create new connection" in the upper left corner:



3. Select the BigQuery Connector from the list



Direct connection – creation

For a connection in direct mode the following information is required.

1. A connector name, which is used on the google BigQuery side for different operations.
 1. An example: "SalesQlikSenseConnection".
 2. This name should at all times be unique between all connector configurations
 3. No spaces is allowed in the name
2. A service account key in json format.
 1. This should be provided by the google cloud administrators or power user. For more details look in the google requirement in the installation guide.
3. The project id on the google cloud platform.
 1. An example: "stretch-cache"
 2. The project needs to contain a dataset and a bucket for the operations of this connector. Create one or reuse an existing one. For more details look in the google requirement in the installation guide.
4. The id of the dataset located in the project defined above.
 1. An example: "cache_dataset"
5. The id of the bucket on cloud storage located in the project defined above.
 1. An example: "cache_data_bucket"
6. The option for CSV load – Fast mode, must be unchecked
7. The connection name is the name visible in Qlik

Create new connection (BigQuery Connector)

BigQuery Connection

Connector identification at google, nospaces allowed

ConnectorName

Service account credentials

Service Account Key (JSON)

Project for where the Dataset and Bucket is located

Project

Dataset the connector can use

DatasetId

Bucket the connector can use

Bucket

Use CSV load (Faster, only for advanced users)

Qlik Properties

Name

Connection name

Test Connection Cancel Create

Fast Connection – Preparations

1. Create a folder on the Qlik server or a file share the qlik server can access
 - The folder or file share must support read/write speeds in excess of 50MB/s to maintain performance



2. Ensure that the Qlik Sense service account has read, write, create and delete right in the folder or file share
3. Create a folder connection in Qlik pointing to the exact folder or file share created above

A screenshot of a dialog box titled 'Edit connection (folder)'. It has a 'Path' field containing 'D > BigQuery Cache >' and a 'Name' field containing 'BigQuery CSV Cache'. Below the fields is a large empty text area with a back arrow icon in the top left corner. At the bottom right, there are 'Cancel' and 'Save' buttons.

Fast connection – creation

For a connection in direct mode the following information is required.

1. A connector name, which is used on the google BigQuery side for different operations.
 1. An example: "SalesQlikSenseConnection".
 2. This name should at all times be unique between all connector configurations
 3. No spaces is allowed in the name
2. A service account key in json format.
 1. This should be provided by the google cloud administrators or power user. For more details look in the google requirement in the installation guide.
3. The project id on the google cloud platform.
 1. An example: "stretch-cache"
 2. The project needs to contain a dataset and a bucket for the operations of this connector. Create one or reuse an existing one. For more details look in the google requirement in the installation guide.
4. The id of the dataset located in the project defined above.
 1. An example: "cache_dataset"
5. The id of the bucket on cloud storage located in the project defined above.
 1. An example: "cache_data_bucket"
6. The option for CSV load – Fast mode, must be checked
7. The location of the folder or file share created during the fast connection preparation step
8. The Qlik folder connection created during the fast connection preparation step
9. The connection name is the name visible in Qlik

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.
- 9.

Create new connection (BigQuery Connector)

BigQuery Connection

Connector identification at google, nospaces allowed

ConnectorName

Service account credentials

Service Account Key (JSON)

Project for where the Dataset and Bucket is located

Project

Dataset the connector can use

DatasetId

Bucket the connector can use

Bucket

Use CSV load (Faster, only for advanced users)

CSV loader (Advanced users)

Path for folder on the server.

D:\DATA\Cache

Name of folder connection pointing to the above folder

BigQuery Cache Folder Connection

Qlik Properties

Name

Connection name

Loading data

Loading data – User interface

The connector has a graphical user interface (GUI) to assist in selection and creating queries for data from BigQuery. For more advanced selections, the connector has a direct query mode where custom queries can be used. This covers the GUI.

1. Select the project from the list of available projects
 - Attention: If a project is changed the previous fields selections is forgotten
2. Select the data from the list of datasets in the project
 1. Attention: If a dataset is changed the previous fields selections is forgotten
3. Select the tables with the desired data
 - Using the checkmark to the left of the table name will select all fields in the table
4. Select the desired fields in the table
 - If all fields was selected, individual fields can be un-selected here
5. Repeat step 3 & 4 to select data from multiple tables
6. Click insert script when all desired fields has been selected

1. 2.

The screenshot shows the 'Select data to load' dialog box. At the top, there is a 'Project' dropdown menu and a 'DataSet' dropdown menu. To the right of these are buttons for 'Metadata' and 'Help', and a 'Switch to query' toggle switch. The main area of the dialog is empty, indicating that no data has been selected yet. At the bottom right, there are 'Cancel' and 'Insert script' buttons.

3. 4. 6.

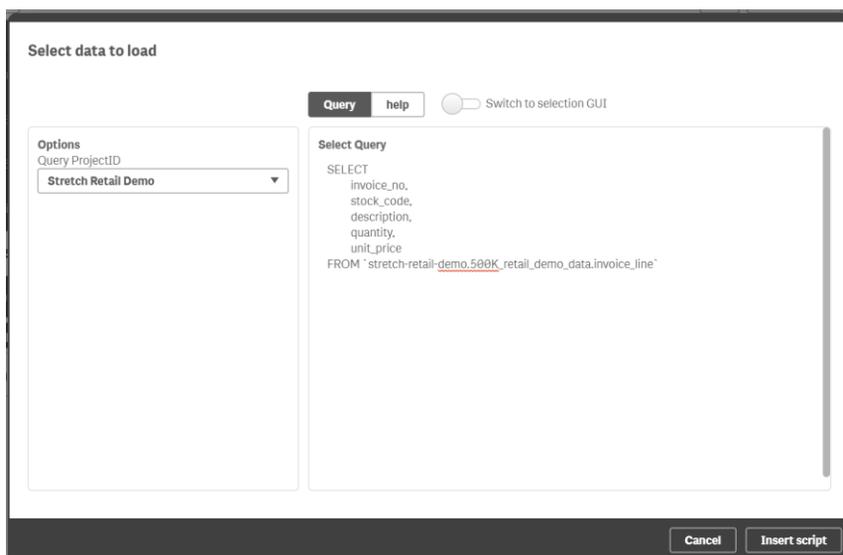
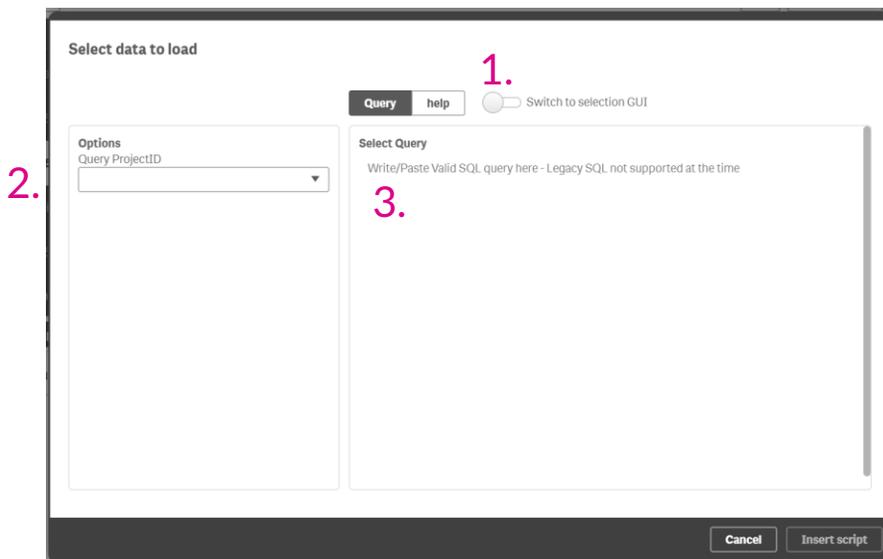
The screenshot shows the 'Select data to load' dialog box with the 'Project' dropdown set to 'Stretch Retail Demo' and the 'DataSet' dropdown set to '500K_retail_demo_data'. A table of fields is displayed, with checkboxes next to each field name. The 'invoice_view' field is selected. At the bottom right, there are 'Cancel' and 'Insert script' buttons.

Field name	Description	Type	Primary key
<input type="checkbox"/> invoice_no		STRING	No
<input type="checkbox"/> invoice_date		TIMESTAMP	No
<input type="checkbox"/> customer_id		INTEGER	No
<input type="checkbox"/> type		STRING	No

Loading data – Query

The connector has a graphical user interface (GUI) to assist in selection and creating queries for data from BigQuery. For more advanced selections, the connector has a direct query mode where custom queries can be used. This covers the direct query.

1. Switch to query mode
2. Select the project from the list of available projects
3. Write or paste in the valid BigQuery query
 - It might be helpful to use the query building in the BigQuery console
4. Click insert script when query is filled in



Notes Tips & Tricks

Notes

- Qlik Sense
 - The data manager is at present time not supported for this connector, if a user tries to use it, they will get a message pointing them to the data editor.
- BigQuery
 - Due to policies in BigQuery it is not possible to make selection across processing locations. The connector needs to be configured for each processing location

Tips & Tricks

- BigQuery
 - The Connector utilizes the build-in IAM functionality in BigQuery, hence all the rights are controlled from the Google Cloud console
- BigQuery logging
 - All the transactions performed by the connector are logged in the project defined during creation of a connection
 - A dedicated project for the connector can be helpful to contain the logging for all connectors in one place. Multiple connectors can use the same project
 - If an error occurs during the operation of the connector, it can be helpful to look at the logs
- BigQuery access & billing
 - Since the connector utilizes a service account, multiple connectors with a unique service account can be created with a specific sub-set of the data in BigQuery
 - The service accounts can be linked to a billing account owned by a department or organizational unit, ensuring the right unit is billed
- Qlik Sense – Mode uses
 - Direct mode is recommended for development and non-time-critical loads due to its ease of use
 - Fast mode is recommended for production for time-critical loads and very large loads, e.g. in QVD layer generation
- Cleaning of bucket and local cache folder
 - From May 2021, all new connections created will automatically have a cleaning feature enabled. This will periodically check and clean the cache bucket and the local cache folder. The default is that files older than 168 hours (7 days) are deleted and this is done on average for every 10th load.

Settings are:

- `bucketCleanup=xx` – How often to clean the bucket, 0 never 100 every time. Default is 10.
- `localCleanup=xx` – How often to clean the bucket, 0 never 100 every time. Default is 10.
- `deletePolicyInHours=xx` – How long to retain the files, 0 delete directly to a positive number. Default is 168.
- It is NOT recommended to do the following, it might cause unwanted interference if multiple loads are running in parallel.
 - To set the delete policy to 0
 - To set the bucket and local cleanup to 100



© Stretch Denmark ApS 2019