**Lab 11 - Use DirectQuery in Power BI Desktop**

With Power BI Desktop, when you connect to your data source, it's always possible to import a copy of the data into the Power BI Desktop. For some data sources, an alternative approach is available: connect directly to the data source using DirectQuery.

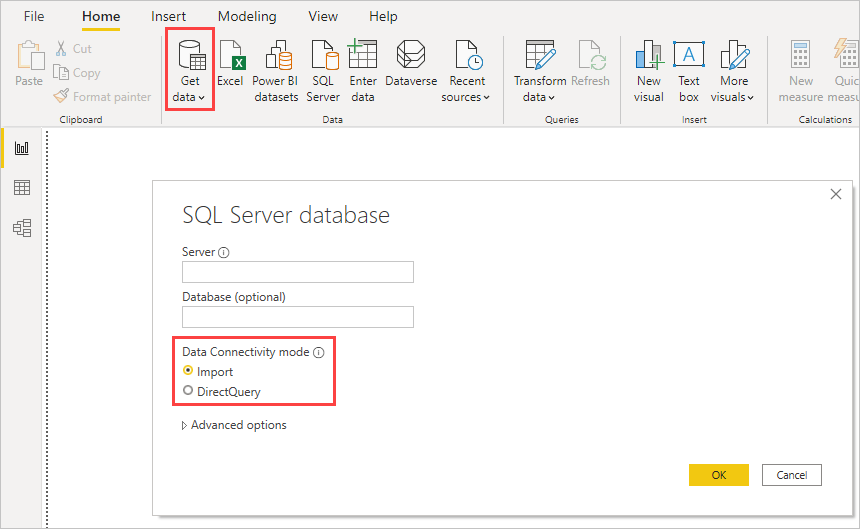
Supported data sources

For a full listing of data sources that support DirectQuery, see Data sources supported by DirectQuery.

<https://docs.microsoft.com/en-us/power-bi/connect-data/power-bi-data-sources>

## How to connect using DirectQuery

When you use **Get data** to connect to a data source supported by DirectQuery, the connection dialog box lets you select how you want to connect. For example, in Power BI Desktop, under the **Home** ribbon, select **Get data** > **SQL Server**. In the **SQL Server Database** dialog box, the **Data Connectivity mode** shows options of **Import** and **DirectQuery**:



Here are the differences between selecting **Import** and **DirectQuery**:

* **Import**: The selected tables and columns are imported into Power BI Desktop. As you create or interact with a visualization, Power BI Desktop uses the imported data. To see underlying data changes since the initial import or the most recent refresh, you must refresh the data, which imports the full dataset again.
* **DirectQuery**: No data is imported or copied into Power BI Desktop. For relational sources, the selected tables and columns appear in the **Fields** list. For multi-dimensional sources like SAP Business Warehouse, the dimensions and measures of the selected cube appear in the **Fields** list. As you create or interact with a visualization, Power BI Desktop queries the underlying data source, so you’re always viewing current data.

## Benefits of using DirectQuery

There are a few benefits to using DirectQuery:

* DirectQuery lets you build visualizations over very large datasets, where it would otherwise be unfeasible to first import all the data with pre-aggregation.
* Underlying data changes can require a refresh of data. For some reports, the need to display current data can require large data transfers, making reimporting data
* unfeasible. By contrast, DirectQuery reports always use current data.
* The 1-GB dataset limitation doesn't apply to DirectQuery.

## Limitations of DirectQuery

There are currently a few limitations to using DirectQuery:

* If the **Power Query Editor** query is overly complex, an error occurs. To remedy the error, either delete the problematic step in **Power Query Editor**, or import the data instead of using DirectQuery. For multi-dimensional sources like SAP Business Warehouse, there's no **Power Query Editor**.
* Calculated tables and calculated columns that reference a DirectQuery table from a data source with Single Sign-on (SSO) authentication are not supported in the Power BI Service.
* Auto date/time is unavailable in DirectQuery. For example, special treatment of date columns (drill down by using year, quarter, month, or day) isn't supported in DirectQuery mode.
* There's a one-million-row limit for cloud sources, with on-premises sources limited to a defined payload of about 4 MB per row (depending on proprietary compression algorithm) or 16MB data size for the entire visual. Certain limits may be raised when using Premium capacity. The limit doesn't affect aggregations or calculations used to create the dataset returned using DirectQuery. It only affects the rows returned. Premium capacities can set maximum row limits, as described in [this post](https://powerbi.microsoft.com/blog/five-new-power-bi-premium-capacity-settings-is-available-on-the-portal-preloaded-with-default-values-admin-can-review-and-override-the-defaults-with-their-preference-to-better-fence-their-capacity/).
  + - For example, you can aggregate 10 million rows with your query that runs on the data source. The query accurately returns the results of that aggregation to Power BI using DirectQuery if the returned Power BI data is less than 1 million rows. If over 1 million rows are returned from DirectQuery, Power BI returns an error (unless in Premium capacity, and the row count is under the admin-set limit).
* There's a 125 column limit in a table or matrix for results that have more than 500 rows for DirectQuery sources. When displaying a result that contains more than 500 rows in a table or matrix, you will see a scrollbar that enables you to fetch more data. In that situation, the maximum number of columns in the table or matrix is 125. If you must include more than 125 columns in a single table or matrix, consider creating measures using MIN, MAX, FIRST or LAST as they do not count against this maximum.
* **Performance and load**: All DirectQuery requests are sent to the source database, so the required visual refresh time depends on how long that back-end source takes to respond with the results from the query (or queries).