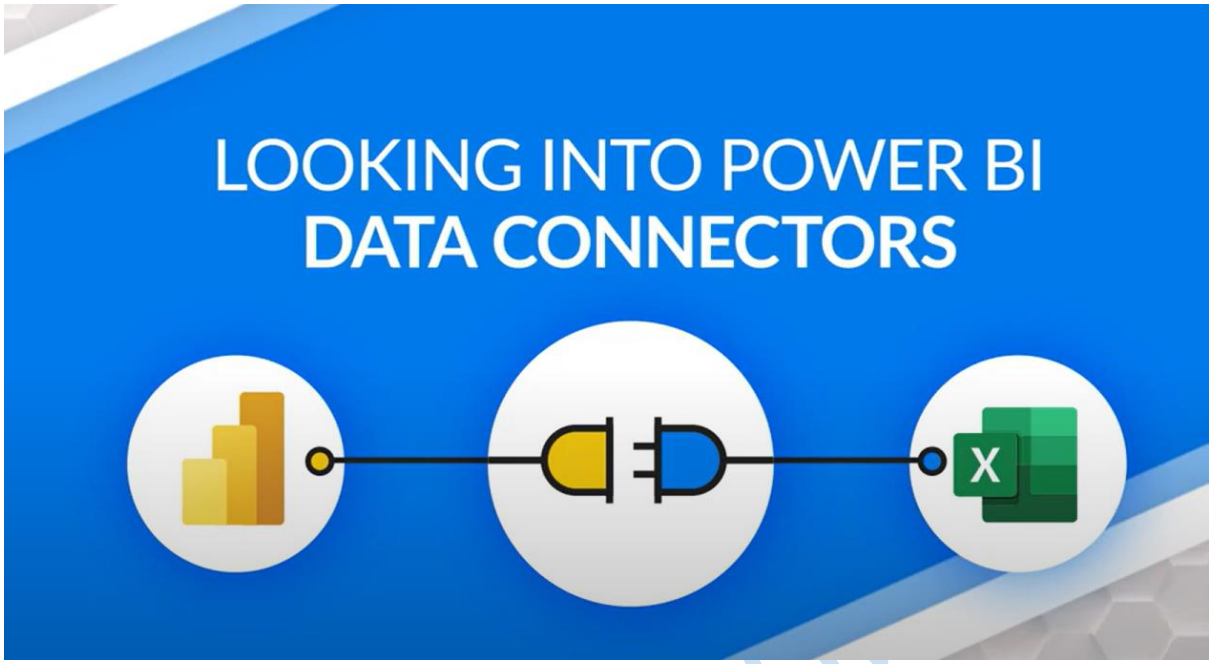
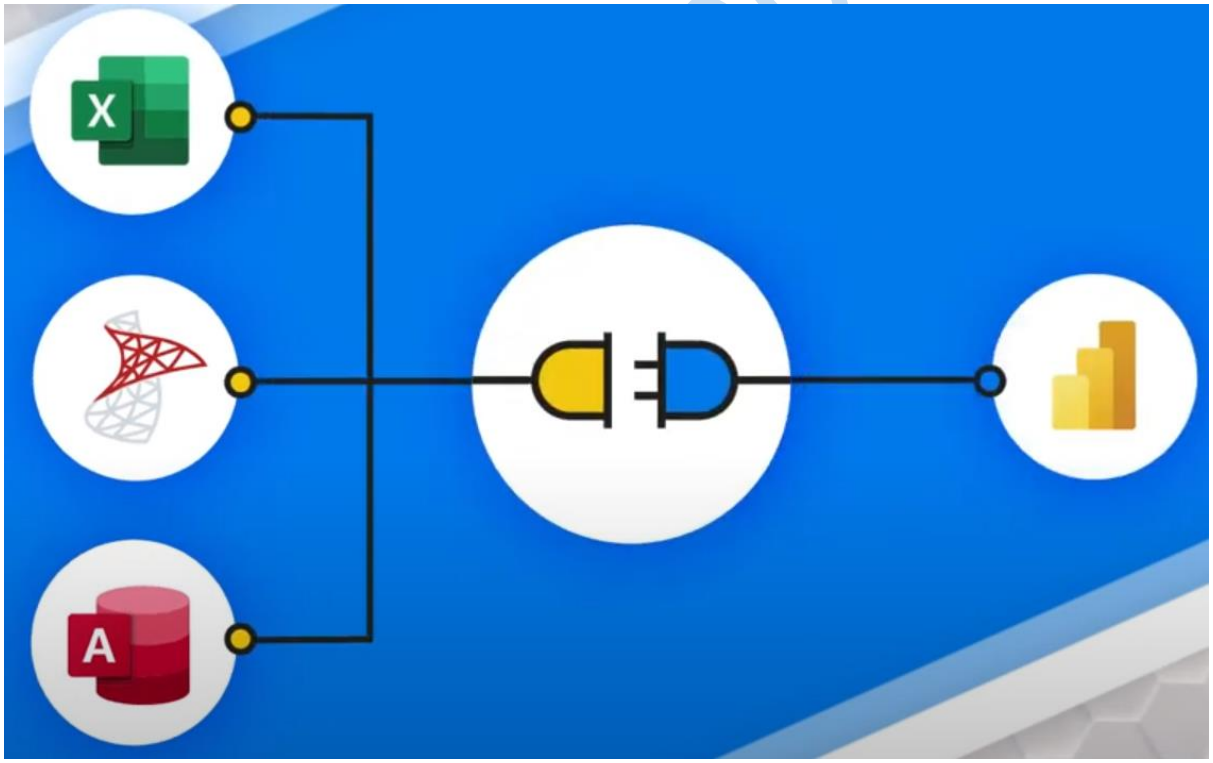
Power BI connectors:





What is a connector: Connecting data sources such as Excel, SQL server, Access database to Power BI. They allow communication between Power BI and Excel or SQL server or Access database or some other services. They allow to connect and use data from the data sources. This is all in the context of Power Query.

All connections made within Power BI Desktop are coming from your local machine.

Beware of network and authentication issues when connecting from your local computer.

Is the data source that we are using – is it available inside of Power BI or Excel?

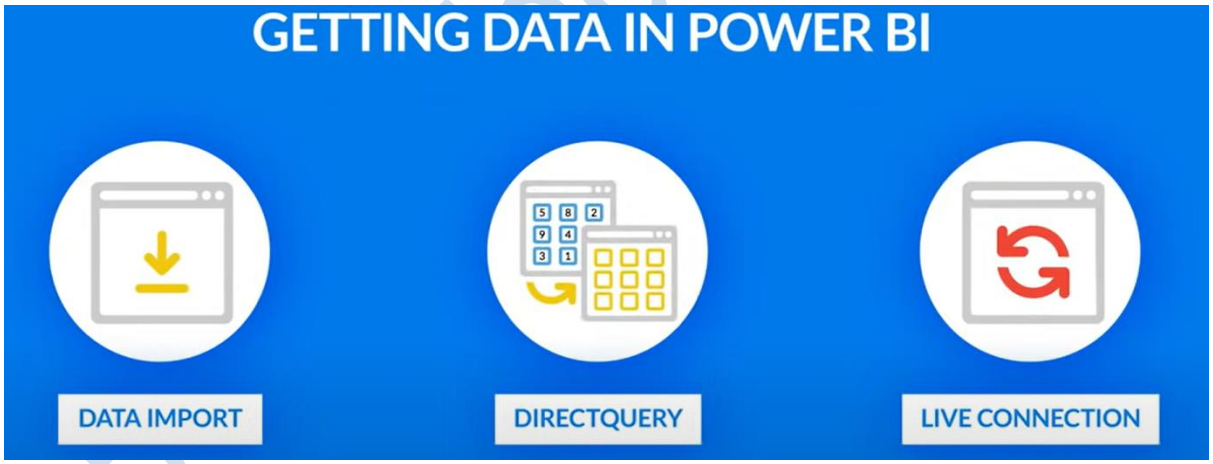
List of all Power Query connectors - Power Query | Microsoft Learn

Data sources in Power BI Desktop - Power BI | Microsoft Learn

To install Custom connectors, we would require to install the relevant drivers first. But custom connectors might have poor performance.

Type of connection – specific to Power BI – Import or Direct Query

Getting data in Power BI:



Data Import is referred to as cached data. It means that we are actually pulling data from your data source, bringing it to Power BI, as a copy of that data and then it sits inside of your Power BI desktop file. If you publish that to Power BI Service, that dataset resides inside of Power BI. And when we query items from a visual perspective, it's hitting that cached data.

Benefits:

Working with import and cached data is the fastest method. And there's data compression involved, so you can get a lot of data into a tiny space.

Also, whatever data sources we see in Power BI in Get Data menu, it is from Import data perspective. So, there's no restrictions on that. We'll have full capability of Power BI with any data type, including all of the DAX functions, all relationship types, all Power Query functions are available in Import mode. Default option.

Restriction: If we are not using premium size, our dataset size (on-disk size of Power BI desktop file) would be 1GB.

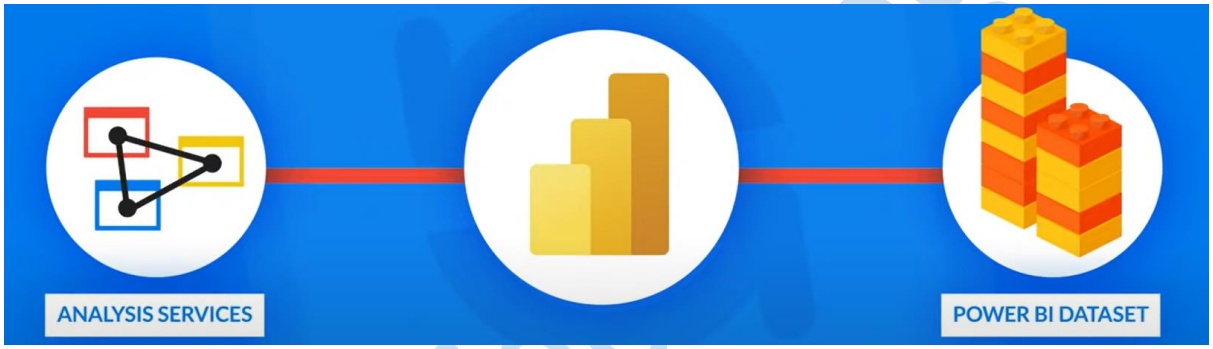
DirectQuery: We are going to connect to the data source and grab the schema of the data. So, information like table structure, the column names is stored inside of Power BI Desktop file in the data model, but the data itself stays at the data source.

Benefits: DirectQuery is the ideal method for working with massive amounts of data unlike Import type.

Restrictions:

1) Not all data sources support DirectQuery, typically relational type data sources.
2) With DirectQuery, there can be performance issues.
3) Also, we should not use Time Intelligence functions.

Live Connection: It is specific to when we're connecting to analysis services or a Power BI dataset. So, this could be Azure Analysis Services, Analysis services on prem, or the Power BI datasets.



From Live connection perspective, we just have the connection string. Everything is kept on the model itself. We don't bring anything over to Power BI Desktop file. If we connect to Power BI Semantic models, we will get that connection string to that semantic model. And no Data view, only Table and Model view available here. So, no changes to that model, because that model is an external source. And if we want to do, that's when composite models come into picture. Composite models are mix of Import, DirectQuery and Live connection types.

A semantic model in Power BI is a logical layer containing the transformations, calculations, and relationships between data sources needed to create reports and dashboards. It serves as the single source of truth for reports across an organization. The semantic model simplifies complex data, defining crucial relationships, calculations, and hierarchies needed for interpreting the data. The model consists of three main components: tables, columns, and measures. The semantic models can be hosted in the Power BI service or externally hosted by Analysis Services.

Dataverse in Power BI allows you to:

1) Connect directly to your data using Power BI Desktop to create reports and publish them to Power BI.
2) Integrate data from multiple sources into a single store, which can then be used in Power Apps, Power Automate, and Power BI.
3) Access the unified data schema for real-time analytics and data-driven decision-making.

In Power BI Desktop, a dataflow is a collection of tables. These tables are created and managed in workspaces in the Power BI service. A dataflow allows you to connect to, transform, combine, and distribute data for downstream analytics. You can add and edit tables in your dataflow, and manage data refresh schedules directly from the workspace in which your dataflow was created.

Row Level Security: It is a way to restrict data based on logged in user. So, when someone signs in, the data is filtered down based on who that person is.

Many to Many relationships:

Many-to-many relationships in Power BI Desktop - Power BI | Microsoft Learn

File Limit Restrictions: In Power BI Desktop, there is no data volume limitation for a load.

However, when you want to publish a .pbix file to Power BI Service, the dataset of a single .pbix file must be smaller than 1 GB. Power BI Premium supports uploads of Power BI Desktop (.pbix) files that are up to 10 GB in size. Once uploaded, a dataset can be refreshed to up to 12 GB in size. To use a large dataset, publish it to a workspace that is assigned to Premium capacity.

Other limitations:

- Distinct values in a column - When caching data in a Power BI dataset (sometimes called 'Import' mode), there is a 1,999,999,997 limit on the number of distinct values that can be stored in a column.
- Row limit - When using DirectQuery, Power BI imposes a limit on the query results that are sent to your underlying data source. If the query sent to the data source returns more than one million rows, you see an error and the query fails. Your underlying data can still contain more than one million rows. You're unlikely to run into this limit as most reports aggregate the data into smaller sets of results.
- Column limit - The maximum number of columns allowed in a dataset, across all tables in the dataset, is 16,000 columns. This limit applies to the Power BI service and to datasets used in Power BI Desktop. Power BI tracks the number of columns and tables in the dataset in this way, which means the maximum number of columns is 16,000 minus one for each table in the dataset.