## Data Sources

* A Data Source in Power BI refers to any location, platform, or system from which Power BI can retrieve data for analysis, visualization & reporting.
* In simple terms, it is the origin of your data — whether it's a file like Excel or CSV, a database like SQL Server, or an online service like SharePoint or Salesforce.

**Examples of Data Sources:**

* **File-Based**: Excel, CSV, JSON, XML, PDF
* **Database**: SQL Server, MySQL, Oracle
* **Cloud Services**: Azure SQL Database, SharePoint Online, Salesforce
* **Web APIs**: REST APIs using JSON or XML
* **Online Services**: Google Analytics, Dynamics 365, GitHub

**Modes of Data Connection in Power BI**

Power BI has three main modes to connect to data:

1. **Import Mode**
2. **DirectQuery Mode**
3. **Live Connection**

**Import Mode:**

In this mode, Power BI imports data from the source and stores it inside the .PBIX file (Power BI internal storage using the VertiPaq engine). The data becomes static until it is refreshed.

**Features:**

1. Fast performance (Thanks to in-memory storage).
2. Full DAX support.
3. Most transformation options in Power Query are available.
4. Can work offline (no need for live connection after loading).
5. Supports scheduled refresh in Power BI service.

**Limitations:**

1. File size limit (1 GB per dataset in Power BI service for free; Pro increased).
2. Data is not real-time — it is only as fresh as the last refresh.
3. Large datasets may impact memory & performance.

**Ideal For:**

1. Small to medium size datasets.
2. Historical reporting where real-time is not required.

**DirectQuery Mode:**

In DirectQuery mode, Power BI does not bring the data into your report. Instead, every time you click a visual (like a chart or slicer), Power BI sends a live request to the original data source (like SQL database) to get the results.

So it is like Power BI is asking the database for answers in real-time instead of keeping a copy of the data.

**Features:**

1. Real-time or near real-time data access.
2. Saves memory because data is not stored in Power BI.
3. Good for large datasets that are too big to import.

**Limitations:**

1. Slower performance than Import mode (depends on source server).
2. Limited support for DAX and Power Query functions.
3. Cannot combine multiple sources easily.
4. Limited relationship — only one active relationship allowed between tables.
5. Data refresh is on-demand or periodic (not continuous real-time).

**Ideal For:**

1. Scenarios where data must always be up-to-date.
2. Large enterprise dataset (millions of rows).
3. Reports with strict data governance that prevents data movement.

**Live Connection:**

Live connection is used when your data is already prepared and stored in a powerful system — like SQL Server Analysis Services (SSAS) or a Power BI dataset that's already published.

In this mode, Power BI does not bring the data into your report. Instead, it just connects to that existing model and lets you build visuals on top of it.

**Features:**

1. No data stored in Power BI.
2. Real-time connection to analysis services.
3. Centralized model — single version of the truth.
4. Allows multiple reports to use the same-shared model.

**Limitations:**

1. No data modelling or transformation in Power BI.
2. You cannot create new tables, columns, or relationships in Power BI.
3. All calculations and relationships must exist in SSAS / Power BI dataset.

**Ideal For:**

1. Organizations using SSAS Tabular Cube models.
2. BI teams building centralized models for reuse across reports.
3. Enterprise scenarios with strict control over data models.

**Composite Model:**

A composite model allows you to combine different data connection modes (like Import & DirectQuery) in the same Power BI report.

**You Can:**

1. Use some tables in Import mode (fast, stored in memory).
2. Use other tables in DirectQuery mode (live connection to the database).
3. And create relationships between them.

**When to use a Composite Model:**

1. You have a large real-time table that must stay in DirectQuery.
2. You have reference or dimension tables that do not change often — better to use Import for performance.
3. You need to combine two or more data sources.
4. You want to reuse a published Power BI dataset and add more data to it.

**Key Features of Composite Models:**

**1. Mixed Storage Mode:**

* Each table can use Import or DirectQuery.
* You can even switch a table’s storage mode later.

**2. Model-Level Security:**

* You can still apply RLS, but with some limitations if mixing modes.

RLS Limitations and Best Practices in Power BI:-  
  
**a. RLS only works on Local Tables:**  
  
Example: If you are using a live connection to a published Power BI dataset, you can’t define new RLS rules on those tables — you must manage RLS at the dataset level where the model was created.  
  
  
  
**b. Security Filtering May Not Work Across Sources:**  
  
Example: Use Dual mode for dimension tables to ensure RLS filters apply correctly across both Import and DirectQuery tables.  
• If you have RLS on an Import table, but it has a relationship with a DirectQuery table, the RLS filter might not correctly propagate to the DirectQuery side.  
  
This is especially problematic when:  
• The relationship direction is single (RLS may not flow backward)  
• Storage mode is not set to Dual for dimension tables  
  
**c. Limited Debugging and Testing in Power BI Service:**  
• Testing RLS for composite models with mixed modes is more complex.  
  
You must publish the model to test RLS properly in the service, especially when combining multiple datasets or sources.  
  
**d. Incompatible Security Models in Live Connections:**  
  
If you are using a live connection to SSAS or a published Power BI dataset and try to add RLS in your local model, Power BI will block it.  
You will get an error like:  
  
“You cannot define RLS in this model because it is connected to a dataset with existing RLS or live connection.”  
  
You must manage RLS at the source dataset level, not in the local model.  
  
**Workarounds and Best Practices:**  
1. Push RLS to the source system (like SQL or the main dataset).  
2. Use Dual storage mode for shared dimension tables.  
3. If possible, apply RLS to import tables only and use those tables to filter the rest of the model.  
4. Avoid mixing too many datasets with conflicting security setups.  
  
  
  
**3. Relationships work across modes:**  
  
Relationships can be created between Import and DirectQuery tables.  
  
**4. Aggregation Tables:**  
  
You can use Import tables for aggregated data and DirectQuery for detailed drilldowns.  
  
**How to identify/change storage mode in Power BI:**  
1. In Model view, click on the table.  
2. Check storage mode: Import / DirectQuery / Dual  
3. You can change mode by right-clicking the table → Properties → Storage Mode  
  
**Benefits of Composite Models:**  
1. Combine multiple data sources in one model  
2. Use real-time data without giving up performance  
3. Build flexible, scalable dashboards  
4. Support for aggregated strategies (Import summary + DirectQuery details)  
5. Enables modular, reusable datasets  
  
**Data Source Settings in Power BI:**  
  
Power BI data source settings allow managing, viewing, and changing data source connections.  
  
**You can think of it as a control center for:**  
1. Managing credentials  
2. Changing file paths, servers, or databases  
3. Adjusting privacy settings  
4. Handling broken or outdated connections  
  
**Where to find it:**  
In Power BI Desktop:  
1. Go to Home tab  
2. Click on Transform Data  
3. Choose Data Source Settings.

**Gateway**:-  
**What are On-Premises Data Sources in Power BI?**  
On-premises data sources are data sources inside your organization’s network, not in the cloud.  
  
Examples:  
1. SQL Server databases  
2. Excel or CSV files on local drives or shared network folders  
Since Power BI Service is cloud-based, it cannot access these directly — that’s where the gateway comes in.  
  
**Power BI Gateway:**  
A Power BI gateway is a bridge that securely connects on-premises data to the Power BI Service (cloud).  
It allows your Power BI reports and dashboards to refresh data from local sources.  
  
**Types of Power BI Gateway:**  
**1. Personal Gateway:**  
• Runs as part of your Windows user account  
• Used for individual use  
• Only supports Import mode  
• Best for personal projects or small teams  
**2. On-Premises Data Gateway (Standard / Enterprise):**  
• Shared gateway  
• Supports Import, DirectQuery, and Live Connection  
• Best for organizations and enterprise projects.  
  
**How Power BI Gateway Works as a Bridge:**  
  
Imagine you have a Power BI report published on the Power BI service, but your actual data (e.g., Excel or SQL Server) is inside your office network (on-premises).  
Power BI in the cloud cannot directly access this internal data — it’s protected behind a firewall.  
  
Gateway acts as a secure tunnel or bridge to connect Power BI service to on-premises data — safely and securely.  
  
  
**Parameters in Power BI:**  
  
Parameters are dynamic values used to:  
1. Control what data is loaded  
2. Switch between environments (Dev, Test, Prod)  
3. Make queries flexible (e.g., filtering dates, regions)  
4. Change data source connections without manual edits  
  
Think of parameters as reusable variables that make reports smarter and more flexible.  
  
**Types of Parameters:**  
1. Text  
2. Decimal / Number  
3. Date / Time  
4. Boolean

**Best practices of working with Data Sources**  
1. Use parameters for dynamic connection  
2. Minimise data volume by filtering before loading  
3. Avoid unnecessary columns  
4. Use folders instead of single files for scalable imports  
5. Schedule refreshes in the Power BI service using a gateway  
6. Secure credentials and use organizational-level privacy settings  
7. Document source path and transformation