

Introduction to Power BI & It's Important for Data Analytics

What is Power BI?

Power BI is a business intelligence (BI) and data visualization tool developed by Microsoft. It enables users to connect to various data sources, transform raw data into meaningful insights, and create interactive dashboards and reports. Power BI is widely used by businesses and analysts to monitor key performance indicators (KPIs), detect trends, and make data-driven decisions.

Key Features of Power BI:

- **Data Connectivity:** Connects to multiple data sources like Excel, SQL Server, cloud services, and APIs.
- Data Transformation: Uses Power Query to clean and shape data before analysis.
- Interactive Dashboards: Provides visually appealing charts, graphs, and reports.
- DAX (Data Analysis Expressions): Enables advanced calculations and custom measures.
- AI-Powered Insights: Uses machine learning to detect trends and anomalies.
- Real-Time Data Updates: Supports live data streaming for up-to-date insights.

Why is Power BI Important for Data Analytics?

1. User-Friendly Interface

Power BI offers a drag-and-drop interface, making it accessible even for users without advanced technical knowledge.

2. Enhanced Data Visualization

It allows users to create visually compelling reports and dashboards that simplify complex data.

3. Seamless Integration

Power BI integrates with Microsoft tools (Excel, Azure, Teams) and third-party applications, making data analysis more efficient.

4. Improved Decision-Making

By providing real-time insights, businesses can make data-driven decisions quickly and effectively.

5. Scalability and Security

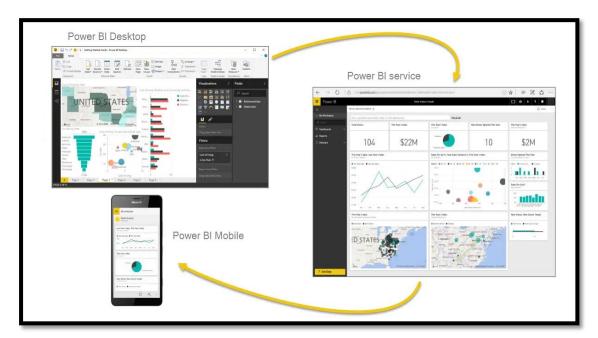
Power BI scales with business needs and offers robust security features, including role-based access control and data encryption.

6. Cost-Effective Solution

Power BI offers both free and premium versions, making it an affordable option for organizations of all sizes.

There are three major components of Power BI ecosystem -

- 1. **Power BI Desktop**: Microsoft Windows desktop application
- 2. Power BI service: online SaaS (Software as a Service) service
- 3. **Power BI Mobile apps:** that are available on any device, with native mobile BI apps for Windows, iOS, and Android.



Getting Started with Power BI

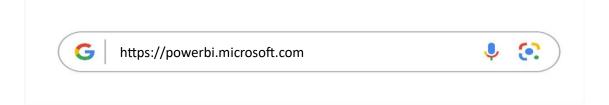
To begin using Power BI, you need to install it, set it up, and familiarize yourself with its interface.

1. Installing and Setting Up Power BI

Step 1: Download Power BI Desktop

Power BI Desktop is a free application that allows users to create reports and dashboards.

- Visit the Microsoft Power BI website
- Click on Download Free or go to the Microsoft Store and search for "Power BI Desktop."
- Install the application by following the on-screen instructions.



Step 2: Sign in or Create Microsoft Account

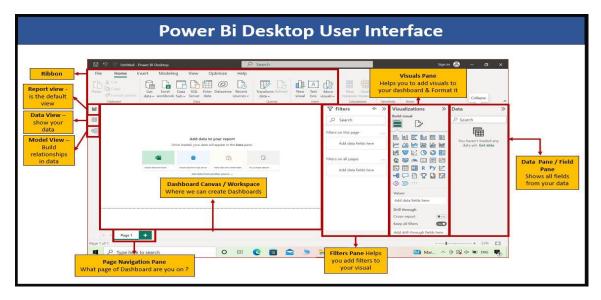
- Open Power BI Desktop after installation.
- If you have a work or school account, sign in; otherwise, create a Microsoft account to access cloud services.

Step 3: Connecting to Data Sources

Power BI supports various data sources, including:

• Excel spreadsheets, SQL databases, Cloud-based services (Google Analytics, Azure, SharePoint, etc.), Online APIs

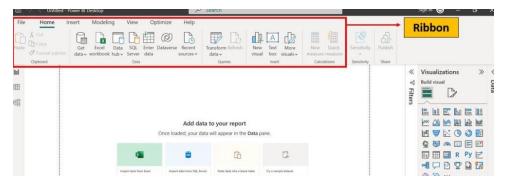
2. Understanding the Power BI Interface



1. Home Ribbon

Contains commonly used options such as:

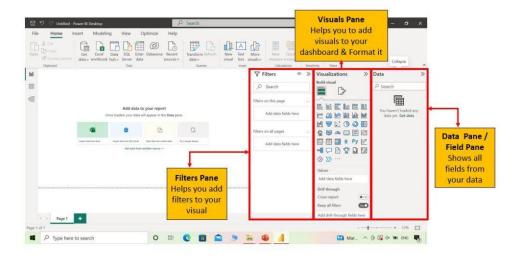
- Get Data: Connect to various data sources.
- Transform Data: Open Power Query Editor for data cleaning.
- Manage Relationships: Define relationships between datasets.
- Publish: Share reports to Power BI Service.



2. Report View

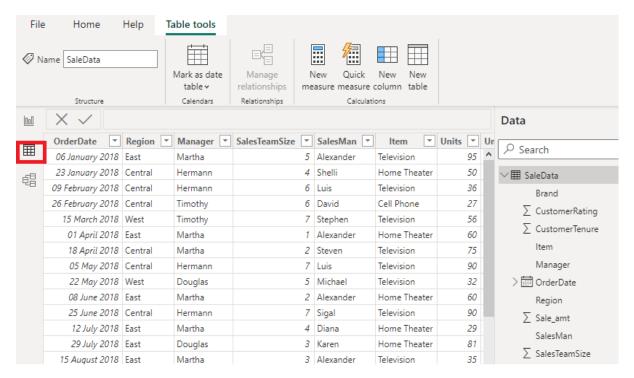
The main workspace where users design reports and visualizations. It consists of:

- Fields Pane: Lists available datasets and fields.
- Visualizations Pane: Allows you to add and customize charts.
- Filters Pane: Apply filters to refine data views.



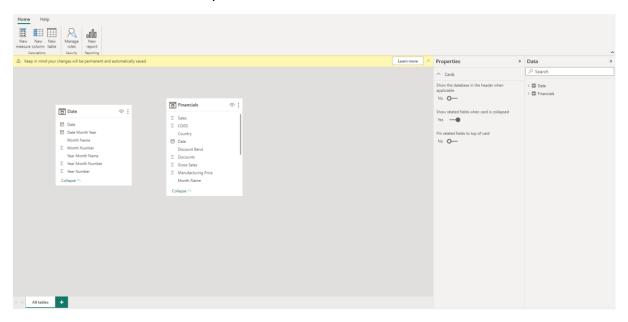
3. Data View

This section shows the underlying data in table format, allowing users to inspect imported datasets.



4. Model View

Used to define and manage relationships between different tables. Helps in creating a structured data model for analysis.



5. Power Query Editor

A separate window where users can clean, transform, and shape data before loading it into Power BI.

Connecting to Data Sources in Power BI

Power BI allows users to connect to various data sources, import data, and transform it before creating reports and visualizations. This step is crucial for preparing clean and structured data for analysis.

A. Importing Data from Excel

Excel is one of the most common data sources for Power BI.

Steps to Import Excel Data:

- 1. Open Power BI Desktop
- 2. Click Home > Get Data > Excel
- 3. Browse and select the Excel file (.xlsx or .xls)
- 4. Choose the sheet or table you want to load
- 5. Click Load to import data directly or Transform Data to clean it in Power Query

B. Importing Data from SQL Databases

SQL databases store large datasets and are widely used in organizations.

Steps to Import SQL Data:

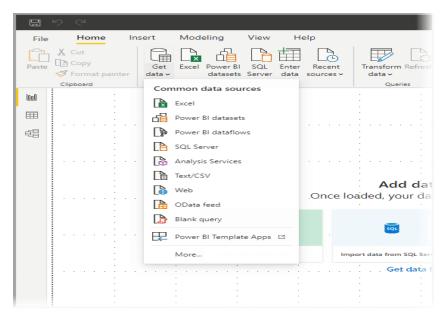
- 1. Click Home > Get Data > SQL Server
- 2. Enter the Server Name and Database Name
- 3. Choose either **Direct Query**(live connection)or **Import Mode**(loads data into Power BI)
- 4. Select the tables or views you need
- 5. Click Load to import or Transform Data to clean it

C. Connecting to Other Data Sources

Power BI supports multiple data sources, including:

- Cloud Services: Google Analytics, Azure, SharePoint, Salesforce
- Online Services: APIs, Power BI datasets, Web Scraping
- Big Data: Hadoop, Spark, Snowflake

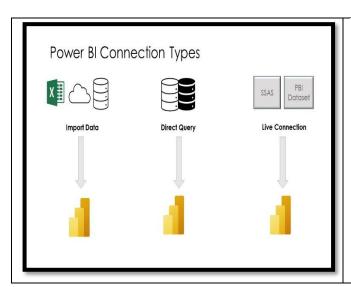
To connect, go to **Get Data** and choose the relevant source, then follow the authentication and selection process.

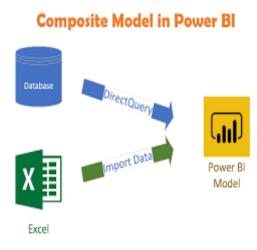


Direct Query & Import Mode

Feature	Import Mode	Direct Query
Data Storage	Stored in Power BI memory	Data remains in the source
Performance	Faster (in-memory)	Slower (live queries)
Data Updates	Requires refresh	Always up-to-date

Dataset Size	Limited to 1 GB (Free & Pro)	No limit (depends on source)
Transformations	Extensive Power Query support	Limited transformations
Best For	Small to medium datasets, performance	Large datasets, real-time data





Data Cleaning & Preparation:

Data Transformation Using Power Query

After importing data, you may need to clean and transform it using **Power Query Editor**.

Power Query is a data transformation and data preparation engine. Power Query comes with a graphical interface for getting data from sources and a Power Query editor for applying transformations. Using Power Query, you can perform the extract, transform, and load (ETL) processing of data.

Opening Power Query Editor:

- Click **Transform Data** after loading a dataset
- Power Query Editor will open, showing data in a tabular format

Common Data Transformations:

- ✓ Removing Unnecessary Columns: Right-click the column > **Remove**
- ✓ Filtering Data: Click on a column's dropdown > Select required values
- ✓ Replacing Null or Missing Values: Select column > Transform > Replace Values
- Splitting Columns: Use **Split Column** to separate text values (e.g., first name & last name)

- Merging Queries: Use **Merge Queries** to combine datasets based on common fields (like a SQL Join)
- ✓ Changing Data Types: Ensure numbers, dates, and text are in the correct format

Applying and Loading Data:

• Click Close & Apply to load the cleaned data into Power BI

Handling Missing Values

Missing values can affect the accuracy of reports and insights. Power Query in Power BI provides multiple ways to handle them.

A. Identifying Missing Values

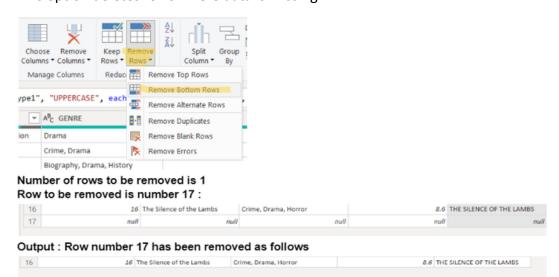
- 1. Open Power Query Editor (Click Transform Data in Power BI).
- 2. Check for **null** or blank values in columns.
- 3. Use filters (dropdown in column headers) to find missing data.

B. Replacing Missing Values

- Select the column with missing values.
- Click Transform > Replace Values
- Enter the value to replace (e.g., 0 for numerical data, "Unknown" for text).
- Click OK to apply changes.

C. Removing Rows with Missing Data

- Select the column with null values.
- Click Home > Remove Rows > Remove Blank Rows
- This option deletes rows where data is missing.



D. Filling Missing Values (Forward/Backward Fill)

- Select the column with missing values.
- Click **Transform > Fill Down** (fills missing values with the previous row's data).
- Click **Fill Up** to fill missing values using the next row's data.

Handling Duplicates

Duplicate data can lead to inaccurate calculations and insights.

A. Identifying Duplicates

- Open Power Query Editor.
- Select the column(s) where duplicates might exist.
- Click **Transform > Group By** to check for repeating values.

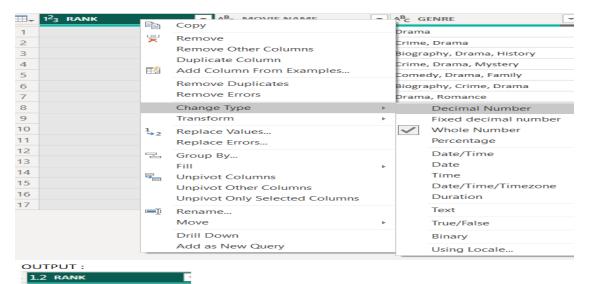
B. Removing Duplicates

- Select the column(s) to check for duplicates.
- Click Home > Remove Duplicates.
- Power BI keeps only the first occurrence and deletes the rest.

Basic Transformations in Power BI

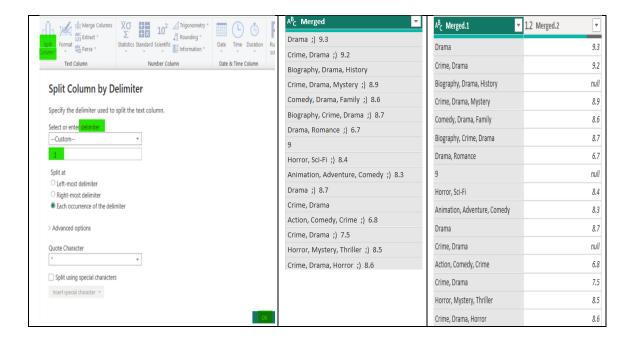
A. Changing Data Types

- Select a column.
- Click Transform > Data Type and choose the correct type (Text, Number, Date, etc.).



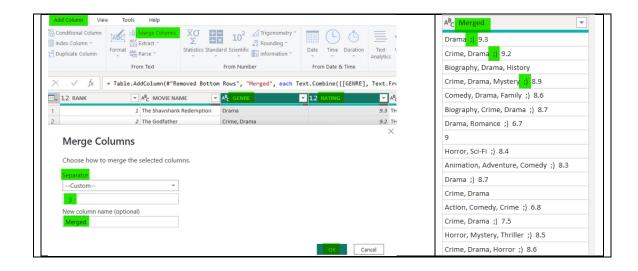
B. Splitting Columns

- Select a column with combined data (e.g., "Full Name").
- Click **Transform > Split Column** and choose **By Delimiter** (e.g., space or comma).



C. Merging Columns

- Select two or more columns.
- Click Transform > Merge Columns, choose a separator (e.g., space or comma), and click OK.

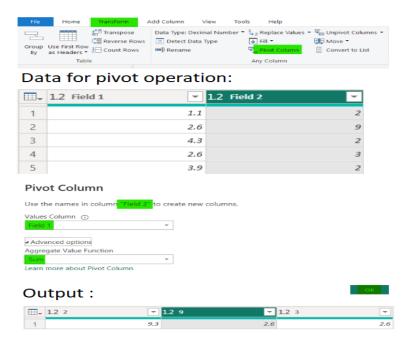


Adding Conditional Columns

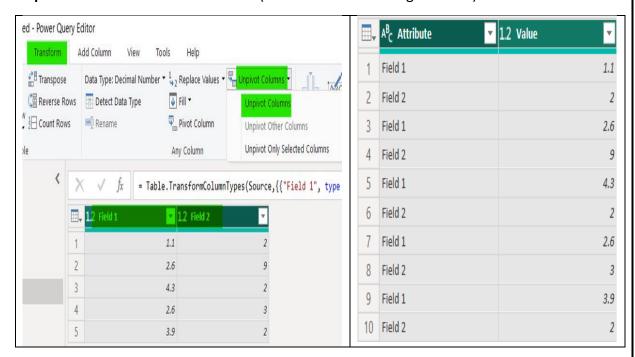
• Click **Add Column > Conditional Column** to create new columns based on conditions (e.g., categorize age groups).

Pivoting & Unpivoting Data

• **Pivot:** Converts unique values in a column into new columns (useful for summarizing data).



• **Unpivot:** Converts columns into rows (useful for restructuring datasets).



Data Modeling in Power BI

Data modeling is a crucial step in Power BI that ensures structured, efficient, and meaningful data relationships. It involves defining **relationships between tables** and using **DAX (Data Analysis Expressions)** to create custom calculations for insightful reporting.

A. Understanding Relationships in Power BI

- A relationship connects tables based on a common column (like a foreign key in SQL).
- Power BI uses primary keys and foreign keys to define relationships.
- Relationships help in data integration and enable efficient aggregations across tables.

B. Types of Relationships

1 One-to-Many (1:M) - Most Common

- One record in Table A relates to multiple records in Table B.
- Example: A Customers table (one row per customer) linked to an Orders table (multiple orders per customer).

2 Many-to-Many (M:M)

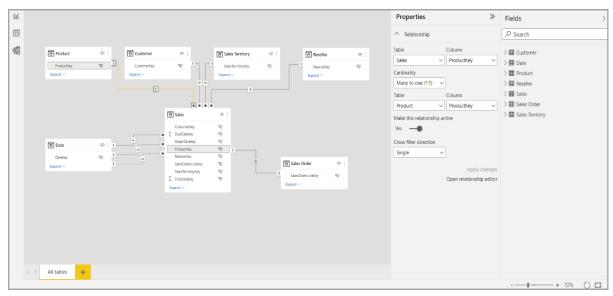
- When multiple records in both tables relate to each other.
- Example: A Products table linked to Sales by multiple categories.
- Requires a bridge table to resolve ambiguity.

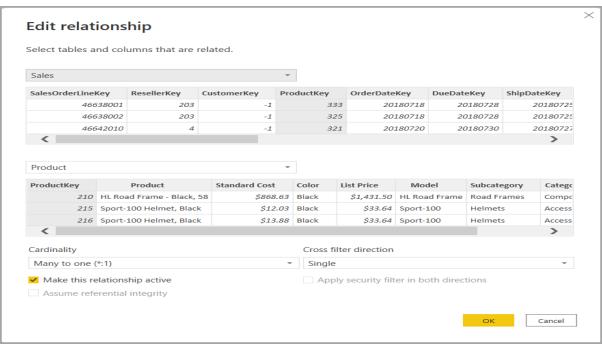
3 One-to-One (1:1)

- One record in Table A relates to exactly one record in Table B.
- Example: A User Details table linked to User Credentials.

C. Creating Relationships in Power BI

- 1. Go to Model View in Power Bl.
- 2. Drag and drop the common field between tables to establish a relationship.
- 3. Power BI automatically sets a One-to-Many (1:M) relationship if a primary key exists.
- 4. Adjust the cardinality (1:M, M:M, or 1:1) as needed.
- 5. Set Cross-filter direction:
 - Single: Filters data in one direction (best for performance).
 - Both: Filters work both ways (useful for many-to-many relationships).
- 6. Click Manage Relationships (under the "Modeling" tab) to review or modify relationships.





Introduction to DAX (Data Analysis Expressions)

DAX is a formula language used in Power BI to create custom measures, calculated columns, and tables for advanced analytics.

A. Types of DAX Calculations

Calculated Columns

- Added within tables to derive new values based on existing data.
- Stored physically in the dataset, increasing file size.

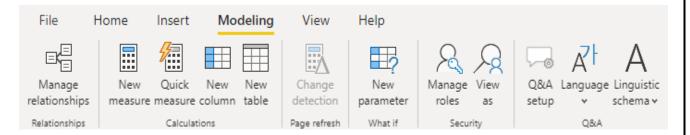
Measures

• Used for aggregations and calculations in reports.

- Calculated dynamically based on filters.
- Efficient and does not increase file size.

Calculated Tables

- Creates a new table based on calculations.
- Used for advanced modeling and filtering.



There are two main types of context in DAX: row context and filter context.

Row context

This refers to just "the current row" across all columns of a table and extends to all columns in related tables. This type of context lets the DAX formula know which rows to use for a specific formula.

Here is an example of a formula for a calculated column that has a row context:

Cost Price Per Unit = financials[COGS] / financials[Units Sold]

Use iterator functions like SUMX or AVERAGEX to explicitly define row contexts in measures.

Filter context

Filter context is applied on top of a row context and refers to a subset of rows or columns that are specified as filters in the report. Filters can be applied in a few ways:

- Directly in a DAX formula
- Using the filters pane
- Using a slicer visual

An example of adding a filter context to a DAX formula is using the CALCULATE function

USA Profit Margin = CALCULATE (SUM (financials[Profit]) / SUM (financials[Sales]), financials[Country] = "United States of America")

DAX (Data Analysis Expressions) is a formula language used in Power BI, Excel Power Pivot, and SQL Server Analysis Services (SSAS). It enables users to create custom calculations, aggregations, and advanced data models.

Types of DAX Functions

1. Aggregation Functions

• SUM: Adds up all values in a column.

```
Total Revenue = SUM(Sales[Revenue])
```

• AVERAGE: Calculates the mean value.

```
Avg Price = AVERAGE(Products[Price])
```

• COUNT: Counts the number of rows.

```
Order Count = COUNT(Sales[OrderID])
```

2. Logical Functions

• IF: Returns different results based on a condition.

```
Category Type = IF(Products[Price] > 1000, "Premium", "Standard")
```

• SWITCH: Works like a case statement.

```
Product_Type = SWITCH( Products[Category], "A", "Electronics", "B", "Furniture",
"C", "Clothing", "Others")
```

3. Filter Functions

• FILTER: Returns a subset of a table.

```
HighSales = FILTER(Sales, Sales[Amount] > 5000)
```

• ALL: Ignores all filters on a table or column.

```
AllSales = CALCULATE(SUM(Sales[SalesAmount]), ALL(Sales))
```

• RELATED: Fetches values from a related table.

```
Customer Region = RELATED(Regions[RegionName])
```

4. Time Intelligence Functions

• DATESYTD: Year-to-date total.

```
YTD Sales = TOTALYTD(SUM(Sales[SalesAmount]), Sales[Date])
```

• SAMEPERIODLASTYEAR: Compares values with the same period last year.

```
Last_Year_Sales=CALCULATE(SUM(Sales[SalesAmount]),SAMEPERIODLASTY EAR(Sales[Date]))
```

• TOTALYTD: Computes cumulative year-to-date values.

```
YTD Profit = TOTALYTD(SUM(Sales[Profit]), Sales[Date])
```

5. Text Functions

• CONCATENATE: Joins two text values.

FullName = CONCATENATE(Employees[First Name], Employees[Last Name])

• SEARCH: Finds a substring within text.

Position = SEARCH("Data", Jobs[Job Title], 1, 0)

6. Mathematical Functions

• ABS: Returns the absolute value of a number.

AbsValue = ABS(-100)

• ROUND: Rounds a number to the specified number of decimal places.

RoundedSales = ROUND(Sales[Amount], 2)

7. Statistical Functions

• MEDIAN: Returns the median of a column.

MedianSales = MEDIAN(Sales[Sales Amount])

• STDEV.P: Returns the standard deviation for an entire population.

StdDevSales = STDEV.P(Sales[Sales Amount])

8. Date & Time Functions

• TODAY: Returns the current date.

Current Date = TODAY()

• NOW: Returns the current date and time.

CurrentDateTime = NOW()

• YEAR: Extracts the year from a date.

Order Year = YEAR(Sales[OrderDate])

9. Rank Functions

• RANKX: Ranks values in a column based on an expression.

RankSales = RANKX(ALL(Sales), Sales[Sales_Amount])

Best Practices for DAX in Power BI

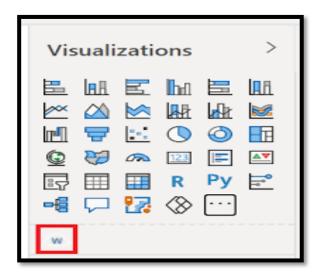
- Use Measures Instead of Calculated Columns
- Optimize Relationships: Ensure proper table relationships to avoid performance issues.

- Avoid Using Too Many Filters: Excessive filtering can slow down reports.
- Use Variables (VAR) for Complex Calculations:
- VAR AvgSales = AVERAGE(Sales[SalesAmount])
- RETURN SUMX(Sales, Sales[SalesAmount] AvgSales)
- Minimize Row Context in Measures: Aggregations should be handled efficiently.

Creating Visualizations in Power BI

Different Types of Charts in Power BI & When to Use Them

Power BI provides a variety of chart types to visualize data effectively. Choosing the right chart depends on the type of data and the insights you want to convey.

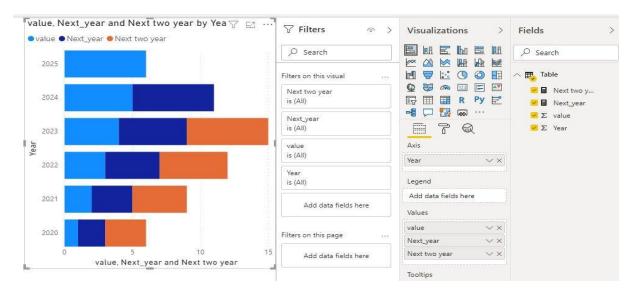


1. Column & Bar Charts

- Comparing different categories or groups.
- Showing rankings and distributions.
- Analyzing discrete data points.

Types:

- Clustered Bar/Column Chart: Compares multiple categories side by side.
- Stacked Bar/Column Chart: Shows the composition of categories.
- 100% Stacked Bar/Column Chart: Displays the percentage contribution of each category.
- **Example:** Sales by product category, number of employees by department.



2. Line Chart

- Showing trends over time.
- Visualizing continuous data.
- Identifying patterns and seasonality.
- **Example:** Monthly sales trends, temperature changes over time.



3. Area Chart

- Highlighting cumulative trends.
- Showing the magnitude of change over time.

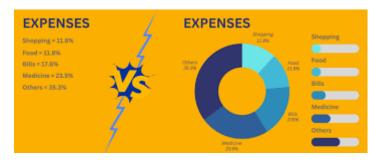
Types:

- Basic Area Chart: Fills the area under the line for emphasis.
- Stacked Area Chart: Shows contributions of multiple categories over time.
- **Example:** Revenue growth comparison across different years.



4. Pie & Donut Charts

- Displaying proportions of a whole.
- · Showing percentage distributions.
- Comparing categorical data with few categories (ideally <5).
- **Example:** Market share of different products, customer demographics.



5. Treemap

- · Visualizing hierarchical data.
- Representing part-to-whole relationships.
- **Example:** Sales contribution by region and sub-region.

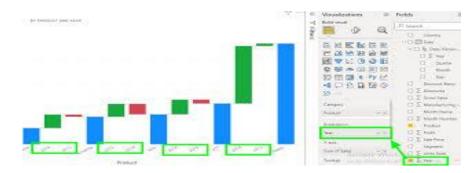


6. Scatter Plot

- Identifying relationships and correlations between two numerical variables.
- Finding outliers in data.
- **Example:** Sales vs. advertising spend, customer age vs. purchase frequency.

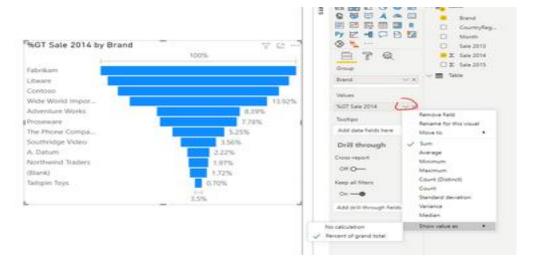
7. Waterfall Chart

- Showing the cumulative effect of sequentially added or subtracted values.
- Understanding financial statements like profit & loss.
- **Example:** Revenue breakdown, budget variance analysis.



8. Funnel Chart

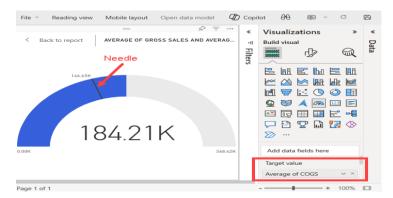
- Visualizing sequential stages in a process.
- Tracking conversion rates (e.g., leads to customers).
- **Example:** Sales pipeline stages, website user journey.



9. Gauge Chart

- Measuring progress toward a goal.
- Showing key performance indicators (KPIs).

Example: Sales target vs. actual performance, customer satisfaction score.



10. Map Visuals (Filled & Bubble Maps)

- Representing geographical data.
- Analyzing location-based insights.
- **Example:** Sales by country, customer distribution across regions.

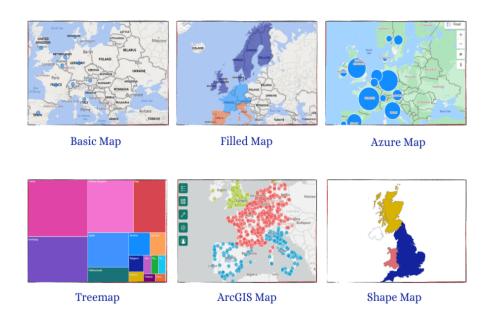


Chart Type	Best For
Column/Bar Chart	Category comparisons
Line Chart	Trends over time
Pie/Donut Chart	Percentage distribution
Treemap	Hierarchical data
Scatter Plot	Correlations between two variables
Waterfall Chart	Financial analysis
Funnel Chart	Conversion tracking
Gauge Chart	KPI measurement
Map Chart	Geographical insights

Using Slicers and Filters for Interactive Reports in Power BI

Slicers and filters help users interact with reports dynamically, enabling them to focus on specific data points without modifying the underlying dataset.

Slicers in Power BI

A slicer is a visual tool that allows users to filter data by selecting values from a dropdown, list, or range. It helps in segmenting data for better analysis.

When to Use Slicers?

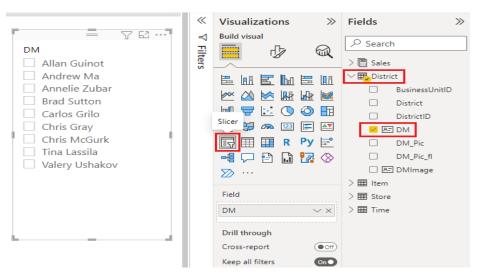
- To provide a user-friendly way to filter reports.
- When a report viewer needs quick, on-screen filtering options.
- When analyzing categorical data like regions, products, or years.

Types of Slicers:

- 1. **List Slicer:** Displays options as a vertical or horizontal list.
- 2. **Dropdown Slicer:** Saves space and works like a dropdown menu.
- 3. **Date Range Slicer:** Allows users to select data within a date range.
- 4. Numeric Range Slicer: Filters numerical data using a slider.
- 5. **Hierarchical Slicer:** Filters based on nested categories (e.g., Country \rightarrow State \rightarrow City).

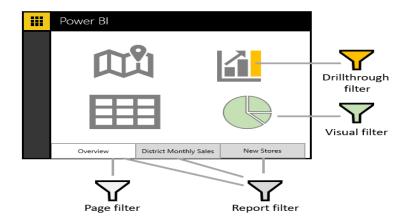
How to Add a Slicer in Power BI?

- 1. Click on Slicer in the Visualizations Pane.
- 2. Drag the field (e.g., "Region") into the slicer visual.
- 3. Resize and position it as needed.
- 4. Choose between **List**, **Dropdown**, or **Range Slider** from the **Format Pane**.



2. Filters in Power BI

Filters allow users to refine data across entire reports, specific pages, or individual visuals.



Types of Filters in Power BI

Filter Type	Scope	Use Case
Visual-Level Filter	Affects a single visual only	Showing only top 5 products in a sales chart
Page-Level Filter	Affects all visuals on a specific page	Filtering a dashboard by a particular year
Report-Level Filter	Applies to all report pages	Filtering the entire report for a selected region
Drill-Through Filter	Filters data based on user interactions	Clicking on a product name to see its detailed performance
Cross-Filtering	Filters data across different visuals	Selecting a category in a bar chart filters a table below

Now to Apply Filters?

- 1. Open the **Filters Pane**.
- 2. Drag the field you want to filter (e.g., "Year") into the appropriate filter level.
- 3. Set filter conditions (e.g., "Show only 2024 data").
- 4. Adjust interaction settings to control how filters affect other visuals.

Example:

- A **Page-Level Filter** on "Product Category" ensures that only selected product data appears across all charts on that page.
- A **Report-Level Filter** on "Region" ensures that only data for India is displayed in all report pages.

Slicers vs. Filters - When to Use Which?

Feature	Slicer	Filter
User Interaction	Directly selectable on the report	Set up in the Filters Pane
Visibility	Always visible to the user	Can be hidden from viewers
Type of Filtering	Best for interactive, on-screen filtering	Works at different levels (visual, page, report)
Use Case	Ideal for categorical selections like Region, Product, Year	Suitable for advanced filtering needs (Top N, Range Filters)

What is a Report?

A report is a multi-page document that contains various visualizations based on a single dataset. It allows deep data exploration, filtering, and detailed analysis.

What is a Dashboard?

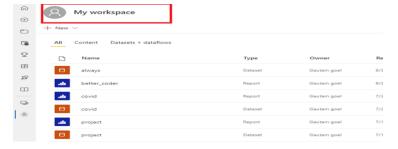
A dashboard is a single-page, high-level summary of key metrics. It combines visuals from multiple reports and datasets, providing an interactive overview.

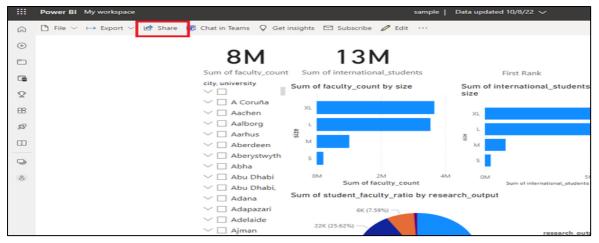
Publishing and Sharing Reports in Power BI

Once you have created your Power BI report, the next step is to **publish** and **share** it with others. Power BI Service allows you to share reports securely with colleagues, stakeholders, or external users.

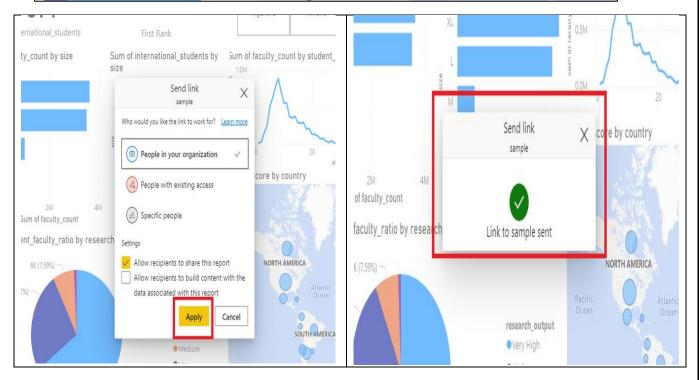
1. Publishing Reports to Power BI Service

- 1. Sign in to Power BI Desktop with your Microsoft account.
- 2. Click on File \rightarrow Publish \rightarrow Publish to Power BI.
- 3. Select a Workspace (e.g., "My Workspace" or a shared workspace).
- 4. Wait for the report to be published.
- 5. Click Open in Power BI to view the report in Power BI Service.
- ♦ You need a Power BI **Pro** or **Premium** license to share reports with others.









2. Sharing Reports in Power BI Service

Once your report is published, you can **share** it with specific users or entire teams.

Sharing Reports via Direct Link:

- 1. Open the report in Power BI Service.
- 2. Click Share at the top right.
- 3. Enter the email addresses of the users.
- 4. Set permissions: Can view (read-only access).

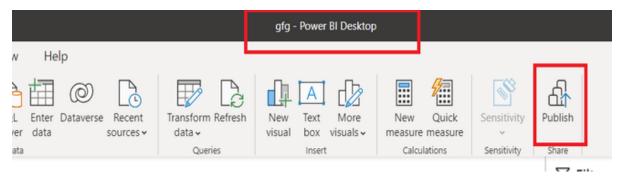
Can reshare (allows users to share further).

5. Click Send.

Sharing Reports via a Power BI App:

If you want to share multiple reports in an organized way, use Power BI Apps:

- 1. Go to Workspaces in Power BI Service.
- 2. Click Create App → Select Reports & Dashboards.
- 3. Customize branding & navigation.
- 4. Click Publish App and share the link.
- **Best For:** Sharing with large teams in an organized way.



3. Setting Up Automatic Data Refresh

Ensures that Power BI reports always show the latest data. Works with cloud-based and on-premise data sources.

Steps to Schedule Automatic Refresh:

- 1. Open Power BI Service and go to the Workspace.
- 2. Click on Datasets → Select the dataset.
- 3. Click **Scheduled Refresh** on the right panel.
- 4. Enable Keep Data Updated and set refresh frequency (Daily, Hourly, etc.).

- 5. Enter database credentials (if required).
- 6. Click Apply.

For On-Premise Data Sources:

Install Power BI Gateway to refresh data from local databases (SQL Server, Excel, etc.).

Refresh only when necessary (avoid overloading). Use **Incremental Refresh** for large datasets. Check **refresh history** for errors.

Managing Access & Permissions

Role-Based Permissions:

Role	Permissions
Viewer	Can view reports but not edit
Contributor	Can edit & publish reports
Admin	Full control over workspace & permissions

Restricting Access:

- Use Row-Level Security (RLS) to control data visibility.
- Set up Power BI Permissions to limit report access.

Row-Level Security (RLS) in Power BI restricts user access to specific data based on their role or attributes. This ensures that users see only the data relevant to them, improving data security and governance.