

Power BI

# POWER BI MASTER CLASS

---

# Overview of the course



## Introduction to Business Intelligence

Understand the basics of Business Intelligence and its importance in today's data-driven world.



## Getting Started with Power BI

Overview of Power BI, including Installation and navigation of Power BI Desktop.



## Data Transformation and Modeling

Learn to clean and transform raw data using Power Query Editor.  
Introduction to data modeling and relationships in Power BI.  
Basics of DAX for data analysis.



## Creating Reports and Visualizations

Explore creating interactive reports and dashboards.  
Practice with different visualizations to tell compelling data stories.



## Publishing and Sharing Reports

How to publish reports from Power BI Desktop to Power BI Service.  
Share insights with dashboards and collaborate with your team.



## Advanced Data Analysis and DAX

Dive deeper into DAX for complex data analysis.  
Learn advanced techniques for more dynamic reports and analysis.



## Power BI in Practice

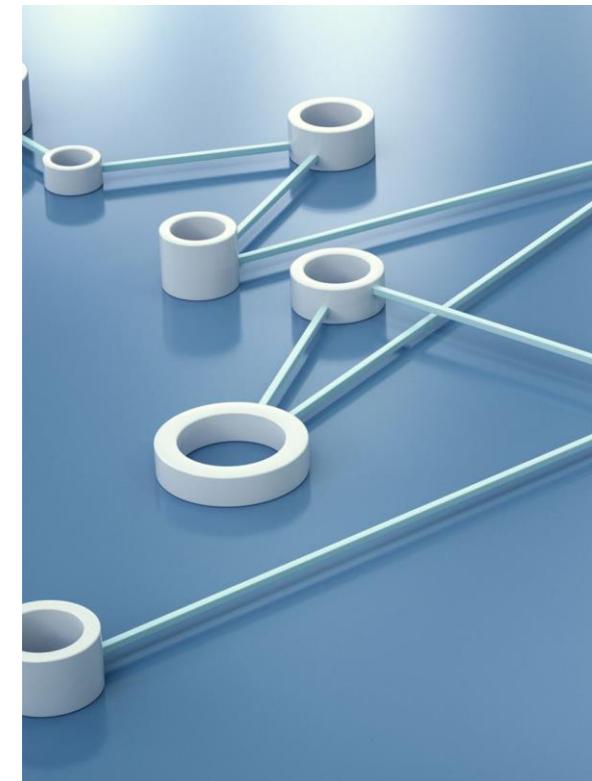
Integration of Power BI with other tools like Excel.  
Real-world case study to apply what you've learned.

# Introduction to Business Intelligence

---

# What is Power BI?

- Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights.



# Why Power BI ?



SEARCH VOLUME IN  
GOOGLE



MAXIMUM NUMBERS  
OF FEATURES

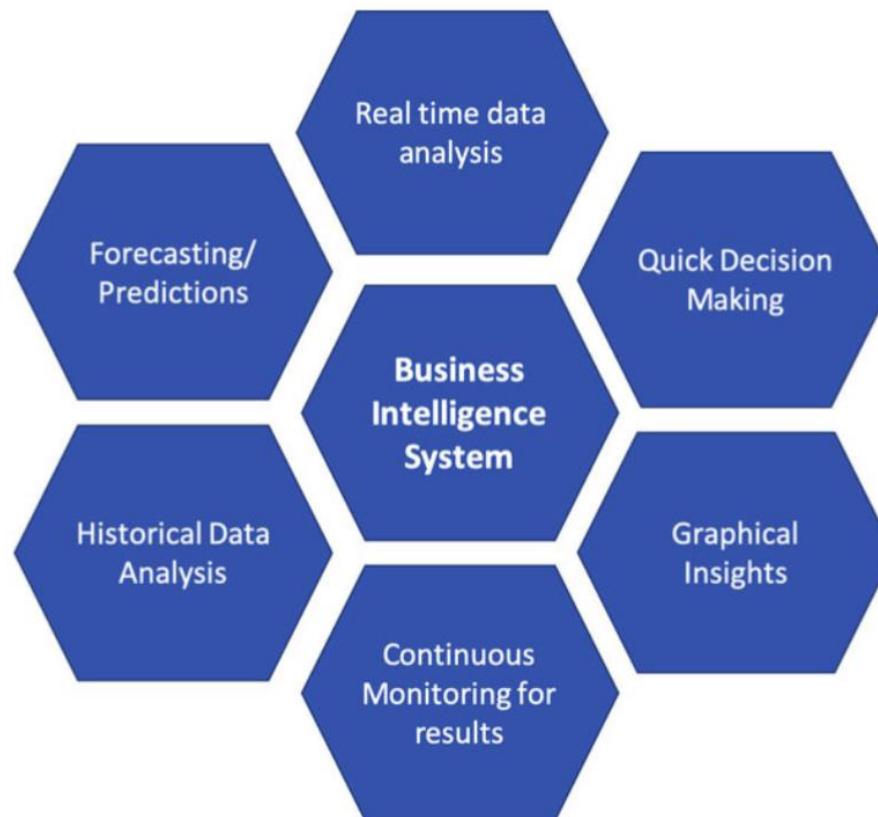


COST OF POWER BI



DATA CONNECTIVITY

# Components Business Intelligence



- **The Essence of Business Intelligence (BI):**

- Simplified: BI is the transformation of data into actionable insights for better decision-making in the present and future.

- **Role of BI Solutions:**

- Aids in understanding business activities.
- Helps answer key questions: Who? What? When? How much?
- Predicts future trends and outcomes for strategic planning.

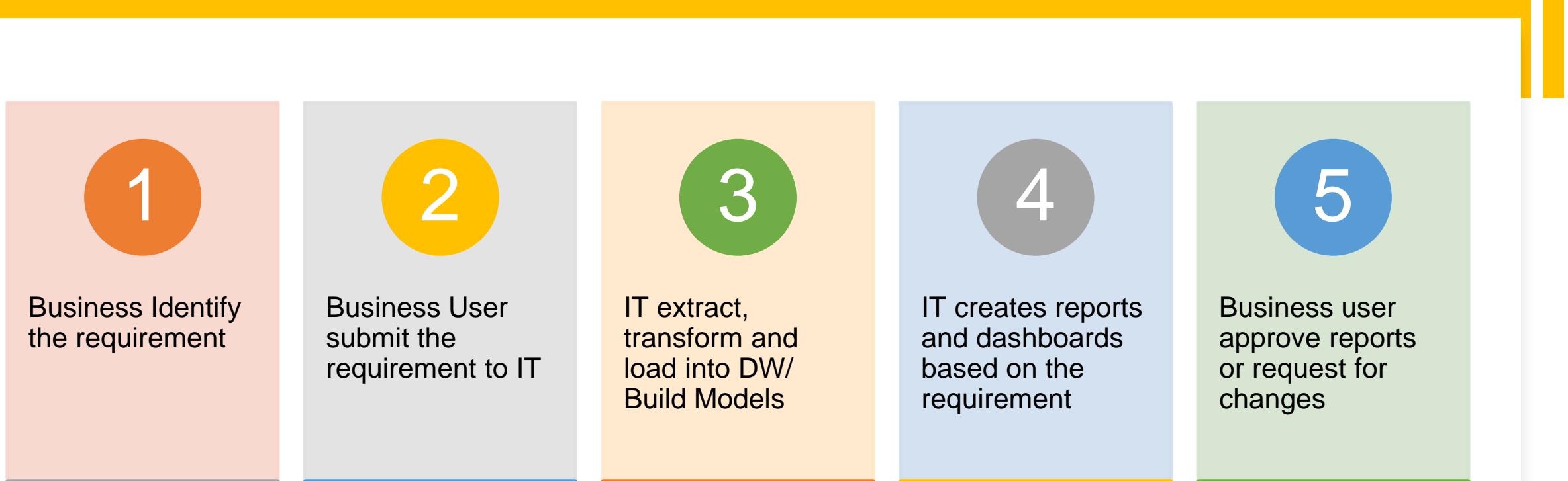
- **BI's Evolution with Big Data:**

- Modern BI leverages cloud computing for cost-efficiency.
- Real-time data processing is integral for up-to-date insights.
- BI tools are expanding to include components like real-time and predictive analytics.

- **The Power of BI Analysis:**

- Beyond traditional data - Capable of analyzing text, sentiment, and more.
- Helps understand public opinion of products and services.

# Traditional BI Approach



# Getting Started with PowerBI

---

# Power BI Ecosystem



**POWER BI  
DESKTOP**



**POWER BI  
SERVICE**



**POWER BI MOBILE**

# Installing Power BI Desktop



## System Requirements:

Check if your system meets the minimum requirements:  
Windows 7 / Windows Server 2008 R2, or later.  
Internet Explorer 9 or above.



## Download:

Visit the official Power BI Desktop download page.  
Choose the correct version (32-bit or 64-bit) based on your system.



## Installation:

Run the downloaded installer file.  
Follow the on-screen instructions to complete the installation process.



## Sign In:

Once installed, open Power BI Desktop.  
Sign in with your Microsoft account or work/school account.



## First Launch:

Familiarize yourself with the welcome screen.



The screenshot shows the Microsoft Power BI desktop application. The ribbon at the top includes tabs for File, Home, Insert, Modeling, View, Optimize, and Help. The Home tab is selected, displaying various data import options like 'Import data from Excel' and 'Import data from SQL Server'. The Data pane on the left shows icons for tables, queries, and data flows. The main workspace is currently empty, with a placeholder message 'Add data to your report'. The Visualizations pane on the right lists available visual types such as bar charts, line graphs, and maps. A 'Filters' pane is also visible on the right side of the workspace.

# Power BI Interface Tour

# PowerBI Desktop Home Tab

**A. Clipboard :** In this option, you can find similar other Microsoft office products.

**B. Get Data :** Where you can ingest data into a Power BI file by connecting different sources.

**C. Excel workbook :** You can upload excel workbook in Power BI.

**D. SQL Server :** You can connect Power BI to SQL server through a data gateway.

**E. Enter Data :** This will open a new window to enter data manually and create a table.

**F. Recents Sources :** Link to open the most recent data sources you have been working on.

**G. Transform data :** To access Power Query Editor in Power BI Desktop.

**H. Refresh :** You can refresh all the tables in the model with the latest data from the source.

**I. New Visual :** By clicking, you can add a new visual to the Report page.

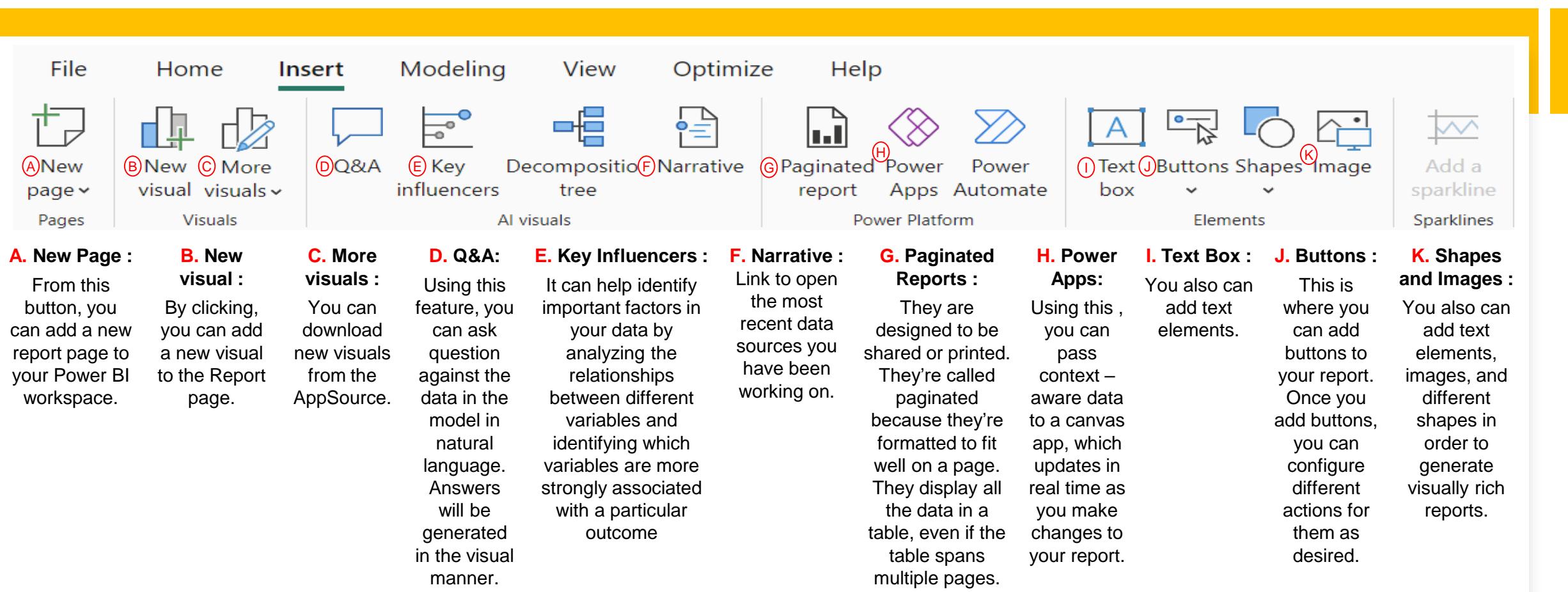
**J. Text Box :** You also can add text elements.

**K. New Measure :** Using DAX queries, you can create new (pre – calculate) measures in your model. This type of calculation is useful when you need to create aggregations that do not require to evaluate row-by-row. Calculate measures are evaluated during the query execution.

**L. Quick Measure :** Allows quickly creating a measure based on the measures and numerical columns in the table

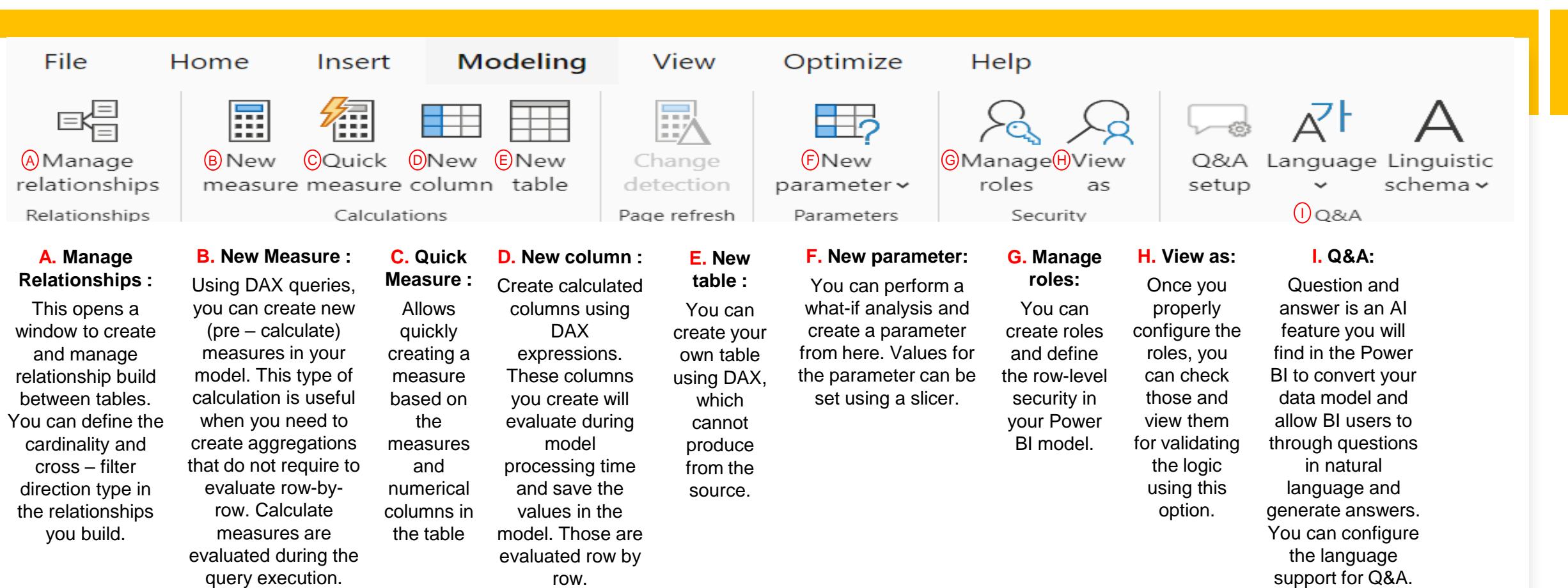
**M. Publish :** You can publish your report on Power BI service using this icon

# PowerBI Desktop Insert Tab



<b>A. New Page :</b> From this button, you can add a new report page to your Power BI workspace.	<b>B. New visual :</b> By clicking, you can add a new visual to the Report page.	<b>C. More visuals :</b> You can download new visuals from the AppSource.	<b>D. Q&amp;A:</b> Using this feature, you can ask questions against the data in the model in natural language. Answers will be generated in the visual manner.	<b>E. Key Influencers :</b> It can help identify important factors in your data by analyzing the relationships between different variables and identifying which variables are more strongly associated with a particular outcome	<b>F. Narrative :</b> Link to open the most recent data sources you have been working on.	<b>G. Paginated Reports :</b> They are designed to be shared or printed. They're called paginated because they're formatted to fit well on a page. They display all the data in a table, even if the table spans multiple pages.	<b>H. Power Apps:</b> Using this, you can pass context-aware data to a canvas app, which updates in real time as you make changes to your report.	<b>I. Text Box :</b> You also can add text elements.	<b>J. Buttons :</b> This is where you can add buttons to your report. Once you add buttons, you can configure different actions for them as desired.	<b>K. Shapes and Images :</b> You also can add text elements, images, and different shapes in order to generate visually rich reports.
---	---	--	--	--	--	---	--	---	---	---

# PowerBI Desktop Modelling Tab



A. Manage Relationships :	B. New Measure :	C. Quick Measure :	D. New column :	E. New table :	F. New parameter:	G. Manage roles:	H. View as:	I. Q&A:
This opens a window to create and manage relationship build between tables. You can define the cardinality and cross – filter direction type in the relationships you build.	Using DAX queries, you can create new (pre – calculate) measures in your model. This type of calculation is useful when you need to create aggregations that do not require to evaluate row-by-row. Calculate measures are evaluated during the query execution.	Allows quickly creating a measure based on the measures and numerical columns in the table	Create calculated columns using DAX expressions. These columns you create will evaluate during model processing time and save the values in the model. Those are evaluated row by row.	You can create your own table using DAX, which cannot produce from the source.	You can perform a what-if analysis and create a parameter from here. Values for the parameter can be set using a slicer.	You can create roles and define the row-level security in your Power BI model.	Once you properly configure the roles, you can check those and view them for validating the logic using this option.	Question and answer is an AI feature you will find in the Power BI to convert your data model and allow BI users to through questions in natural language and generate answers. You can configure the language support for Q&A.

# PowerBI Desktop View Tab

**A. Themes :**  
By changing the theme, you can add variety to your Power BI reports. There are pre-built themes you can quickly change.

**B. Page view :**  
This page view helps you to change into different reporting layouts.

**C. Mobile layout :**  
It will enable mobile view, then you can arrange visual elements the way you want to see in a mobile device.

**D. Gridlines :**  
This is a very helpful tool to show a grid in the report layer. After you create the visuals, this will help you to do the final touch by align the visuals perfectly.

**E. Snap to grid :**  
Once this is enabled, when you move, the visual objects are always aligned with one of the grid points.

**F. Lock objects :**  
Even if you perfectly align your visuals after everything is finished, there is a chance that these objects can move here and there by touching. So, by this option, you can ensure nothing is moved.

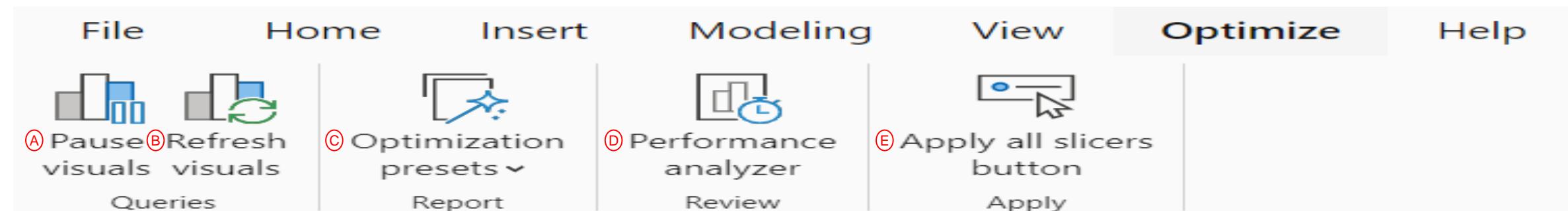
**G. Bookmarks :**  
This works like a snapshot; you can bookmark(save) your report objects with different filters and interactions.

**H. Selection :**  
Selection pane feature is also mostly used with bookmarking feature. You will be able to configure the visibility of objects for different bookmarks.

**I. Performance analyzer :**  
This enables you to trace time information for DAX queries and visual loading. Further, this allows you to identify the DAX query behind each visual, which you can use further performance tuning.

**J. Sync slicers :**  
This allows you to select one of many slicers in your report page and start to work those slices across multiple pages.

# PowerBI Desktop Optimize Tab

**A. Pause visuals :**

It allows you to add or remove field list items to a visual without the visual updating until you are ready.

**B. Refresh visuals :**

It allows you to update all visuals in a report page based on a constant interval such as one second or five minutes.

**C. Optimization presets :**

With this you can quickly choose and apply predefined combinations of settings tailored to your reporting scenario.

**D. Performance analyzer :**

This enables you to trace time information for DAX queries and visual loading. Further, this allows you to identify the DAX query behind each visual, which you can use further performance tuning.

**E. Apply all slicers button :**

When on the report page, will keep track of all the slicer selections and then apply them at once when you click the Apply all slicers button!

# PowerBI Desktop Help Tab



- A. About :** This will open small windows containing the section in the version, user ID, session ID.
- B. Guided Learning :** You will be redirected to the Power BI YouTube channel from this link.
- C. Training videos :** You will redirect to the Getting started article series. If you are good at learning this is for you.
- D. Documentation :** This will bring you to the Power BI up screen, you can access the page. This is the place for the latest you can ask products and for support. feature updates.
- E. Support :** This will bring you to the Power BI blog where all the experts are live in. That is the place you can raise concerns and ask questions.
- F. Power BI blog :** Like in the start- with a very large community base where all the peers and experts are live in. That is the place you can raise concerns and ask questions.
- G. Community :** You can connect to the Power BI blog where all the experts are live in. That is the place you can raise concerns and ask questions.
- H. Power BI for developers :** This will bring you to the developer page, expert or a newbie. where you can start to build custom visuals, always a installed embedding, and place locally and automate tasks like duplicate Power BI workspace idea, which automatically.
- I. Submit an idea :** It does not matter you are an access to developer page, expert or a newbie. where you can start to build custom visuals, always a installed embedding, and place locally and automate tasks like duplicate Power BI workspace idea, which automatically.
- J. External tools :** It provides easy access to developer page, expert or a newbie. where you can start to build custom visuals, always a installed embedding, and place locally and automate tasks like duplicate Power BI workspace idea, which automatically.
- K. Examples :** Another good way to start to look at existing work done by others. You can find download Power BI sample work files.
- L. Consulting services :** You will be redirected to paid and free consulting/ workshop for Power BI.

# PowerBI Desktop



## Report Tab :

In Report view you can create any number of report pages with visualizations. Report view in Power BI Desktop provides a similar design experience to the report's editing view in the Power BI service. You can move visualizations around, copy and paste, merge, and so on.



## Data Tab :

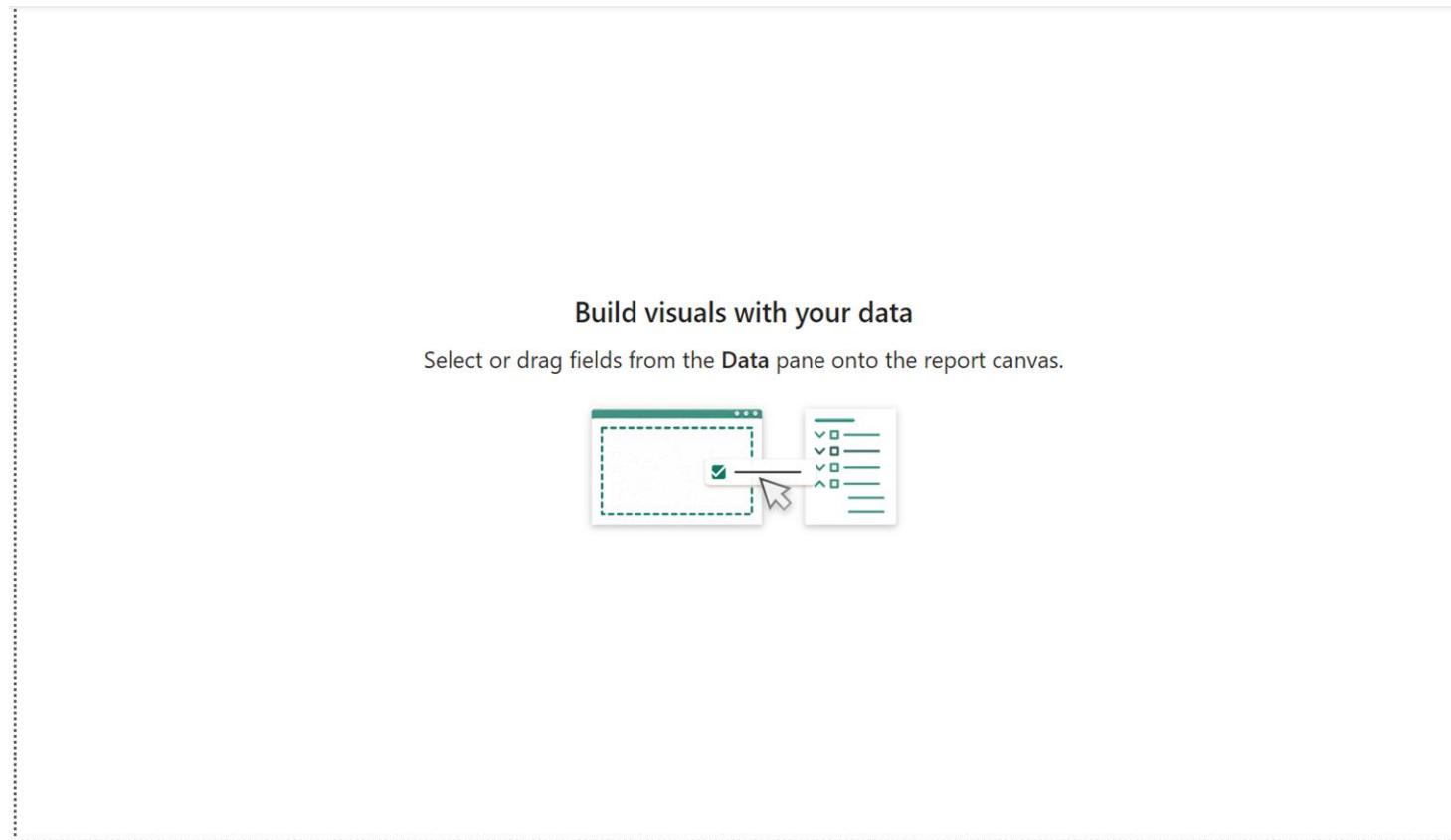
You can come to Data tab by clicking the table icon in the left vertical menu bar. This view is similar to the view we can see in Microsoft Excel. You can click the dropdown button in each field in order to do the data set, such as sorting, filtering, and so on.



## Model Tab :

The model tab is where you can view your Star schema or snowflake structure of the model. Further, this view helps you to build relationships, especially when you have many tables. This is also the ideal place to do advanced configurations like aggregations.

# PowerBI Desktop Report View



# PowerBI Desktop Data View

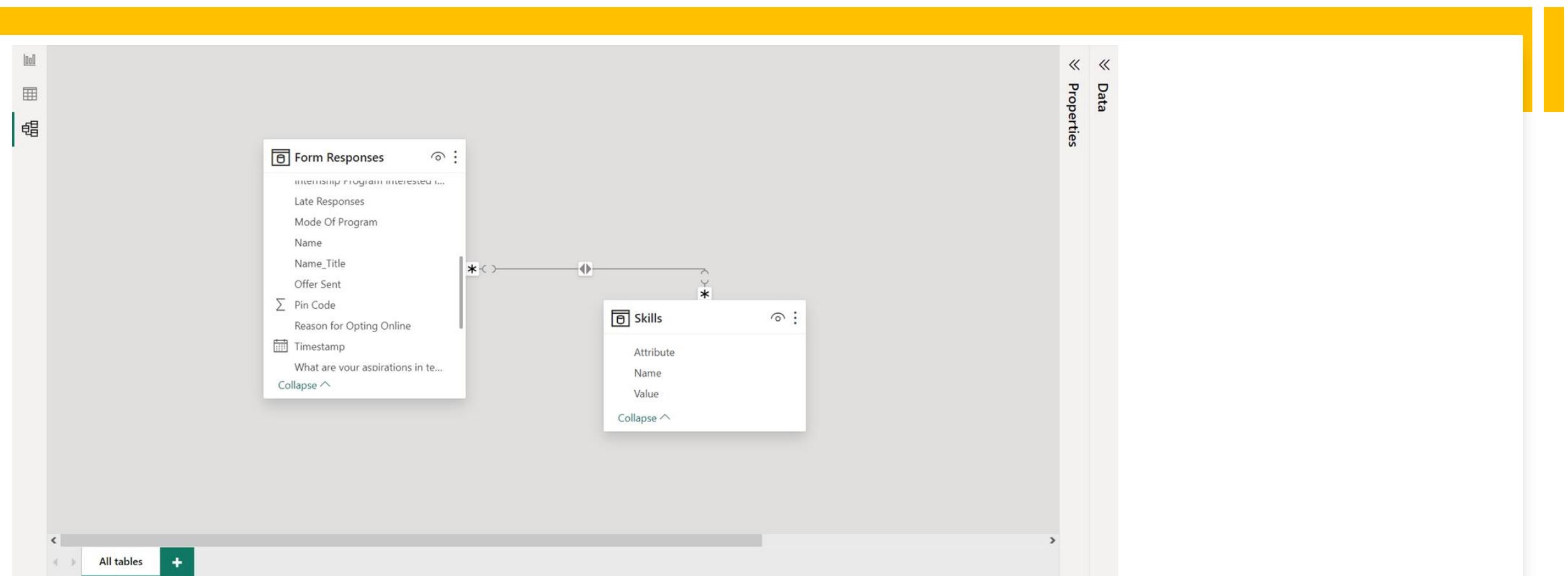
The screenshot shows a PowerBI Desktop interface with a yellow header bar. The main area displays a data view with the following columns:

expoure / knowledge?	What are your aspirations in terms of CTC/annual package?	Any specific companies you plan to target?	Late Responses	Degree
5L-10L			Late Response	B.Tech
>15L		Not decided yet	Late Response	BCA
Not Sure			Late Response	B.Tech
Not Sure			Late Response	B.Tech
BI/Tableau, Excel	5L-10L		Late Response	B.Tech
	5L-10L		Late Response	B.Tech
	5L-10L		Late Response	B.Tech
	5L-10L		Late Response	BCA
	5L-10L	IBM	Late Response	BCA
	5L-10L	no	Late Response	B.Tech
	5L-10L	no	Late Response	BCA
BI/Tableau	5L-10L	Data Science Oriented Companies	Late Response	BCA
	5L-10L	no	Late Response	MSc
bleau	5L-10L	I didn't have target	Late Response	BCA
	5L-10L		Late Response	B.Tech
	5L-10L	no	Late Response	B.Tech
	5L-10L	MNC company's	Late Response	B.Tech
BI/Tableau, Excel	>15L		Late Response	B.Tech
BI/Tableau, Excel	5L-10L	no	Late Response	B.Tech
BI/Tableau	5L-10L	no	Late Response	BCA
	5L-10L		Late Response	B.Tech
	5L-10L		Late Response	B.Tech
BI/Tableau, Excel	5L-10L	Accenture	Late Response	B.Tech
BI/Tableau, Excel	10L-15L		Late Response	B.Tech
BI/Tableau, Excel	5L-10L		Late Response	B.Tech
	10L-15L		Late Response	B.Tech
BI/Tableau, Excel	5L-10L	Microsoft   Amazon   Accenture   Infosys	Late Response	MCA

On the right side, there is a "Data" pane with a search bar and a list of items:

- Dim\_Location
- Form Responses
- Pincode to lat long
- Skills

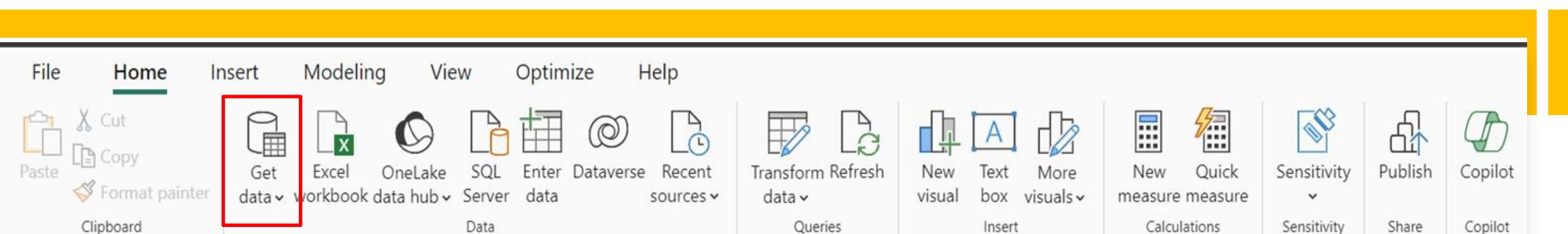
# PowerBI Desktop Model View



# Data Transformation and Modelling

---

# Getting Data - Basic Import from Excel



In order to connect with any data source, first, you need to click the **Get Data** in the **Home** tab in the top menu bar in the Power BI desktop.

Once you click the **Get Data** button, it will open the **Get Data** window, which gives huge list of sources:



**Get Data**

The screenshot shows the 'Get Data' dialog box from Power BI. On the left, there's a sidebar with a search bar and categories: All, File, Database, Microsoft Fabric, Power Platform, Azure, Online Services, and Other. The 'All' category is selected. The main pane lists data sources under 'All', including Excel Workbook, Text/CSV, XML, JSON, Folder, PDF, Parquet, SharePoint folder, SQL Server database, Access database, SQL Server Analysis Services database, Oracle database, IBM Db2 database, IBM Informix database (Beta), IBM Netezza, and MySQL database. At the bottom, there are buttons for 'Certified Connectors' and 'Template Apps', and two main action buttons: 'Connect' (green) and 'Cancel'.

Get Data

All

Excel Workbook

Text/CSV

XML

JSON

Folder

PDF

Parquet

SharePoint folder

SQL Server database

Access database

SQL Server Analysis Services database

Oracle database

IBM Db2 database

IBM Informix database (Beta)

IBM Netezza

MySQL database

Certified Connectors

Template Apps

Connect

Cancel

# Get Data

# Power BI Connectors

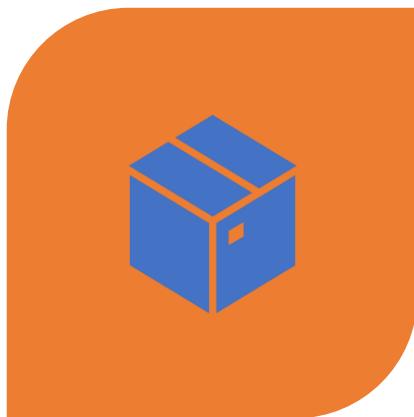
Microsoft Power BI offers a range of connectors, starting from text documents/CSV files to Big Data sources like **Hadoop File Systems (HDFS)**. The Power BI team keeps adding new data connectors to the Power BI.

The following diagram demonstrates the list of connectors available in the Power BI. However, this is just few data sources.

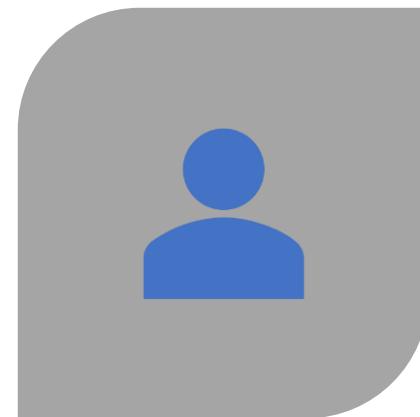
<b>File</b>	<b>Database</b>	<b>Microsoft Fabric</b>	<b>Power Platform</b>
 Excel Workbook	 SQL Server database	 Power BI semantic models	 Power BI dataflows (Legacy)
 Text/CSV	 Access database	 Dataflows	 Common Data Service (Legacy)
 XML	 SQL Server Analysis Services database	 Datamarts (preview)	 Dataverse
 JSON	<b>Azure</b>	<b>Online Services</b>	<b>Other</b>
 Folder	 Azure SQL database	 SharePoint Online List	 Web
 PDF	 Azure Synapse Analytics SQL	 Microsoft Exchange Online	 SharePoint list
 Parquet	 Azure Analysis Services database	 Dynamics 365 Online (legacy)	 OData Feed

# Connection Types

However, the way of connecting data is different from connector to connector. Predominantly, we can group them into three connection types.



**IMPORT**



**DIRECT QUERY**



**LIVE CONNECTION**

# Import

## Load

-  Form Responses  
Waiting for other queries...
-  Dim\_Location  
Evaluating...
-  Pincode to lat long  
Waiting for other queries...



In the **Import** connection type, when you connect any data source, the data will be imported into the Power BI. You can experience this when you try to connect to text/csv or Excel data sources or even in Microsoft SQL Server connector. While you are importing data into Power BI, it will show up a window with the load status along with the list of data sets.

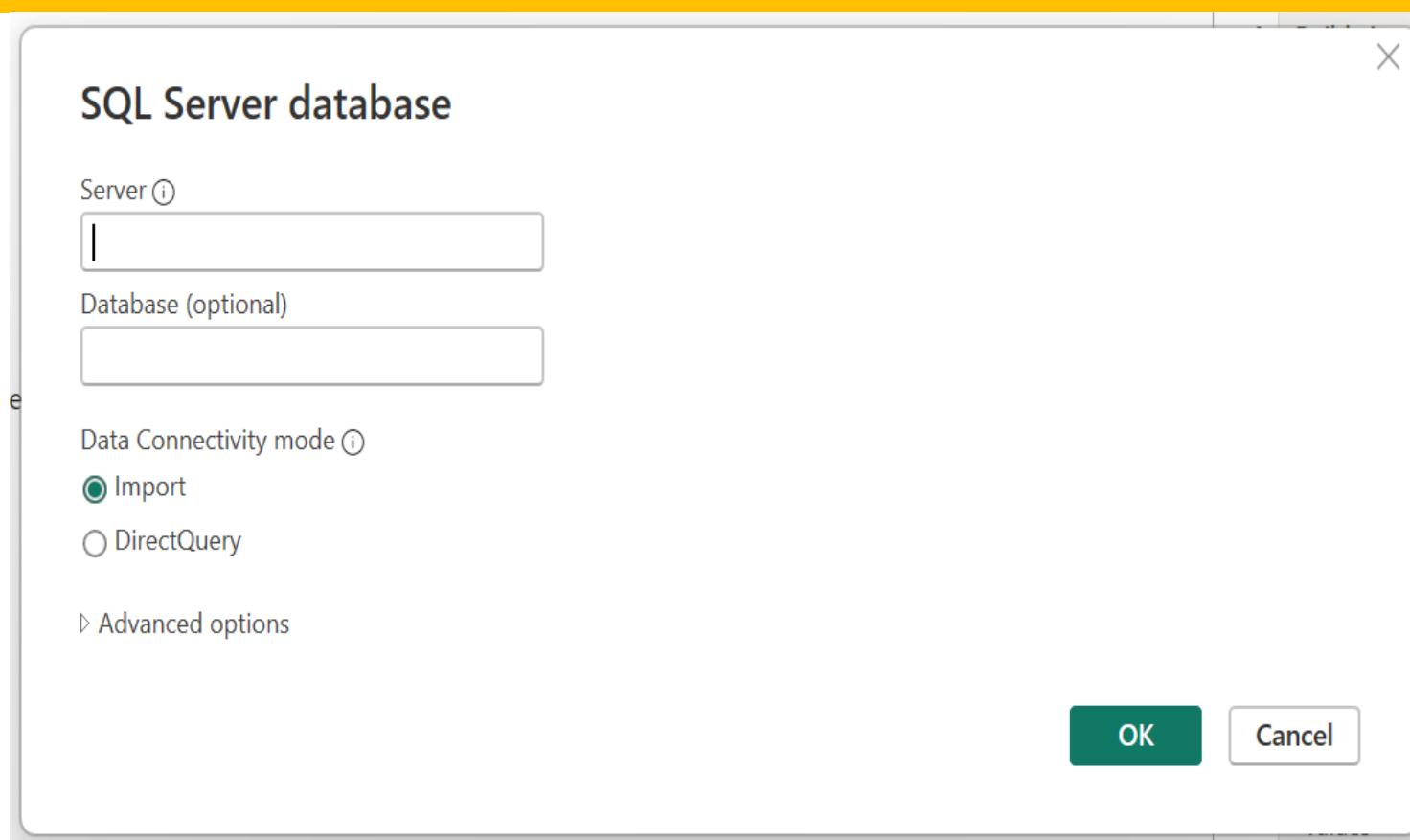
 Cancel

When you import data from the Power BI storage engine, Vertipaq (Compression Engine of Microsoft Power BI), compresses and store your data in-memory. That memory is machine you are holding the Power BI file. Once you publish the report into

The screenshot shows the Windows Task Manager interface. At the top, it displays 'Task Manager' and a search bar with the placeholder 'Type a name, publisher, or PID to search'. Below this, the title 'Processes' is visible. A header row provides performance metrics: Name, Status, CPU, Memory, Disk, and Network. The CPU column is currently sorted, showing values of 17%, 45%, 1%, and 0% respectively. The Memory column is also sorted, with values of 0%, 2.9 MB, 0 MB/s, and 0 Mbps. The task details are as follows:

Name	Status	CPU	Memory	Disk	Network
Microsoft OneDriveFile Co-Au...		0%	2.9 MB	0 MB/s	0 Mbps
Microsoft SharePoint		0%	4.7 MB	0 MB/s	0 Mbps
Microsoft SQL Server Analysis ...		0%	536.9 MB	0 MB/s	0 Mbps

# Direct Query



- **DirectQuery Definition:**
  - DirectQuery is a connection option in Power BI that allows real time data querying from the source without importing the data into Power BI.
  
- **Real – Time Data Visualization :**
  - DirectQuery ensures you are always seeing the latest data by live querying the backend source.



## ➤ **Performance Considerations:**

- DirectQuery may be slower than Import mode due to factors like data source performance and network latency.
- Import mode compresses and stores data in-memory (e.g., Vertipaq), offering better performance for visualization.

## ➤ **Use Cases:**

- Choose DirectQuery for scenarios where real-time data is crucial.
- Opt for Import mode when performance and speed are more critical, and the data doesn't need to be real-time.

## ➤ **SQL Server Database Connection:**

- Both Import and DirectQuery options are available when connecting to SQL Server databases.
- Select the appropriate option based on your specific needs and performance requirements.

➤ **Proof of Concept (PoC):**

- If data quality is unknown and requires transformation, use Import mode to create a data model.
- For data warehouses with pre-cleaned and transformed data, DirectQuery is preferable as importing an entire data warehouse might not be feasible.

➤ **DirectQuery Specifics:**

- **Connection:** Instead of loading data, DirectQuery creates connections to the data source.



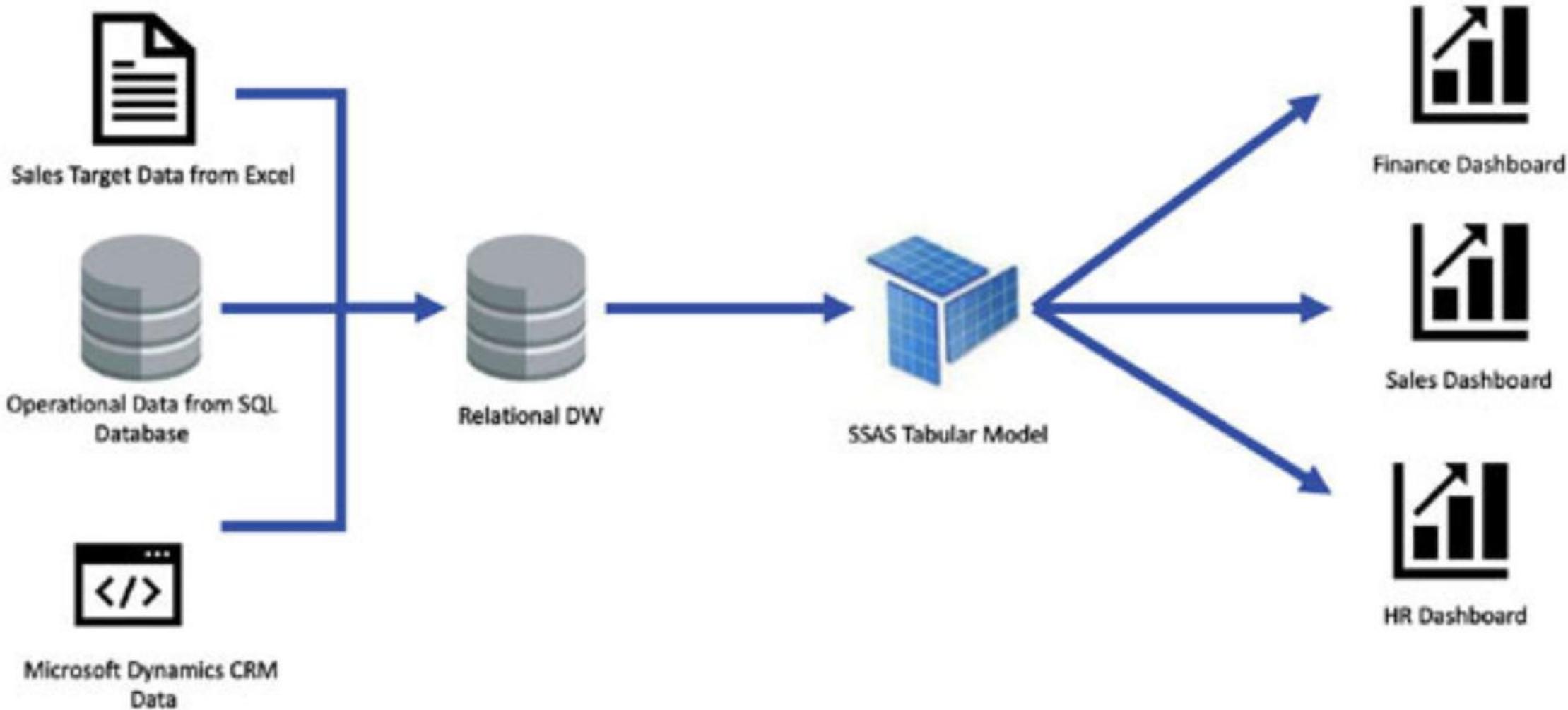
# Live Connection

## ➤ Import Mode Limitation:

- In Import mode, there's a limitation on data import, especially when publishing a report to Power BI service with data volume exceeding 1 GB.
- Power BI doesn't limit data analysis; SQL Server Analysis Services (SSAS) Tabular database handles large data volumes.
- Both Microsoft Power BI and SSAS Tabular use the same storage engine technology, Vertipaq.

## ➤ Corporate BI Projects:

- As a BI Engineer, implementing data warehouses for clients is common.
- After loading data into a relational data warehouse, you can:
  - Connect Power BI directly for analysis.
  - Create a semantic data model using the SSAS database (tabular database).
- Implementing SSAS database improves user experience for Business Intelligence users.
- Similar functionalities are available in both Power BI and SSAS Tabular.



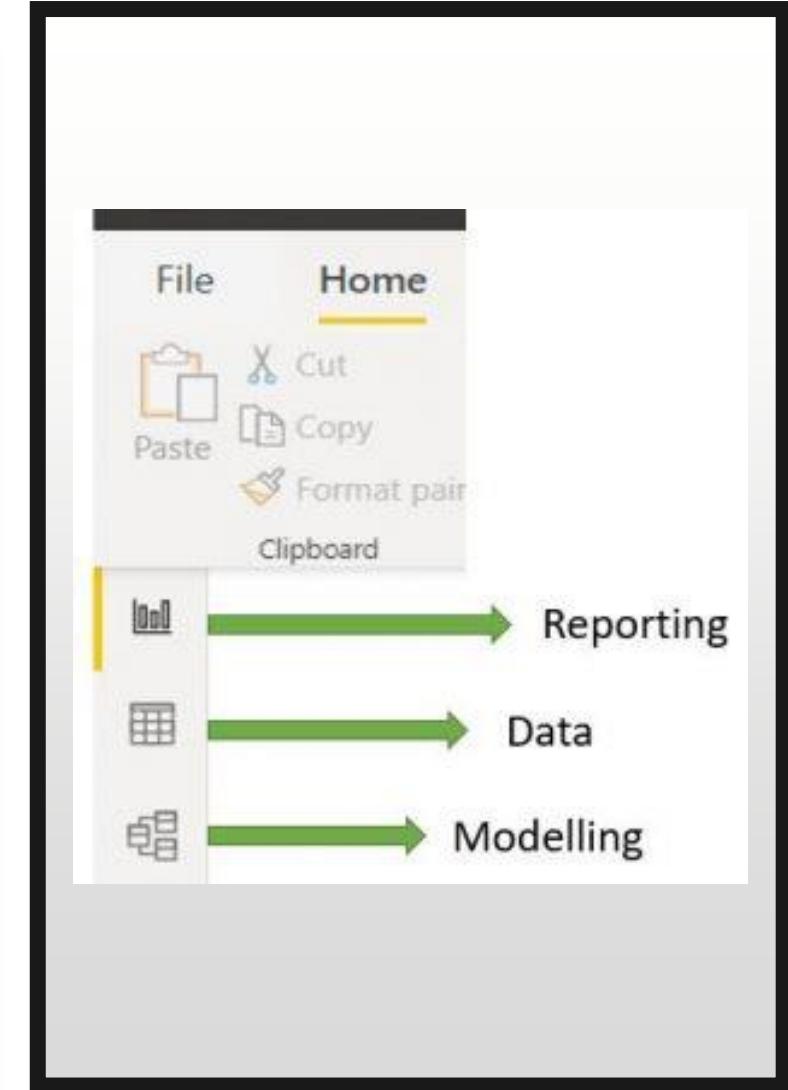
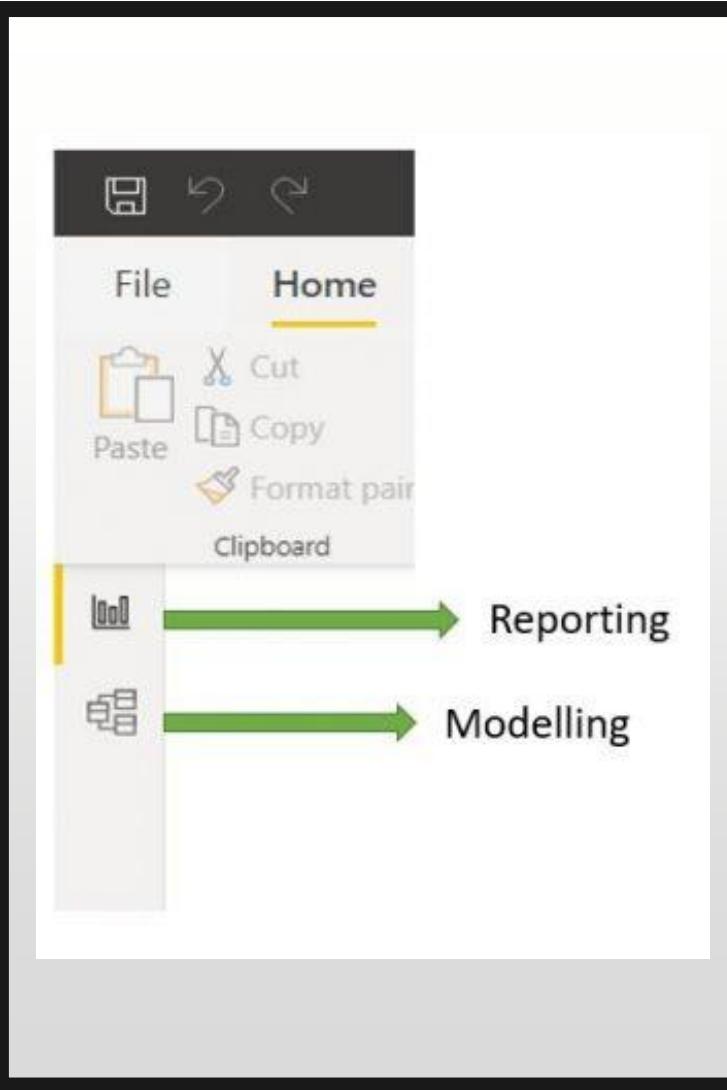
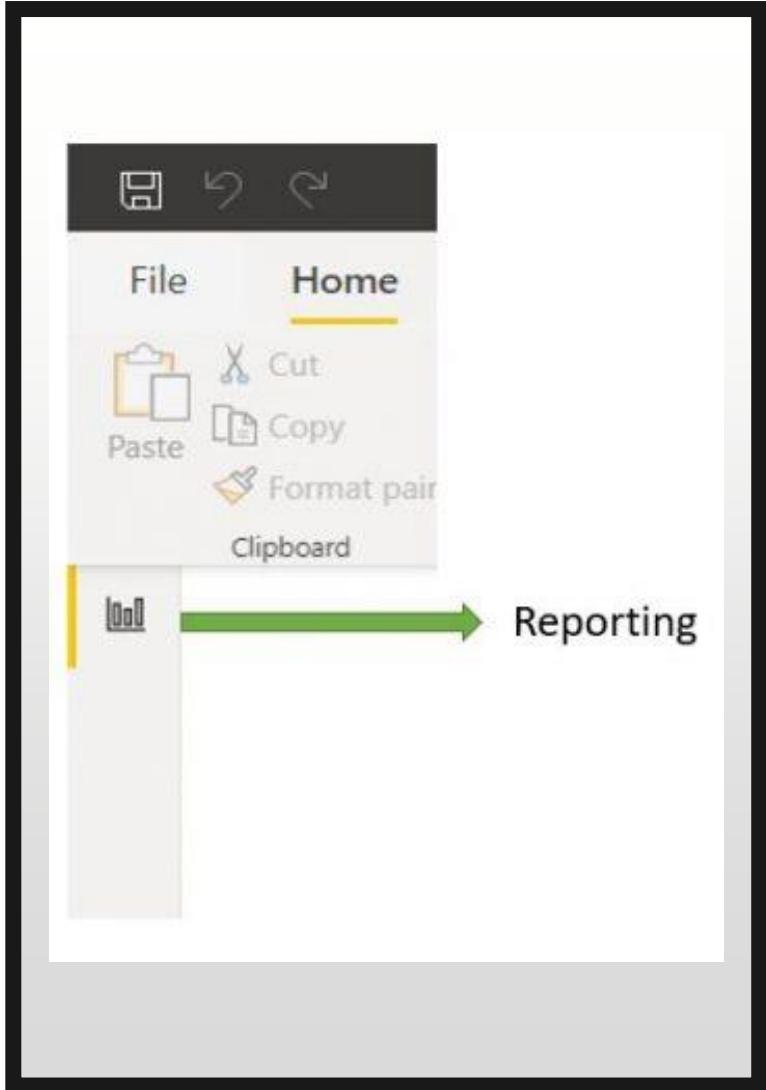
## ➤ **SSAS Tabular Data Model:**

- Implemented after setting up the data warehouse using SQL Server.
- SSAS Tabular is an in-memory database consuming deployed server memory.
- No upper limit for memory expansion, surpassing the 1 GB data limit in Import mode.
- In this solution, Power BI functions as a reporting tool, not for data transformation/modeling.
- Steps for data transformation/modeling are handled in the relational data warehouse and SSAS Tabular model.

# Guess the connection type ?



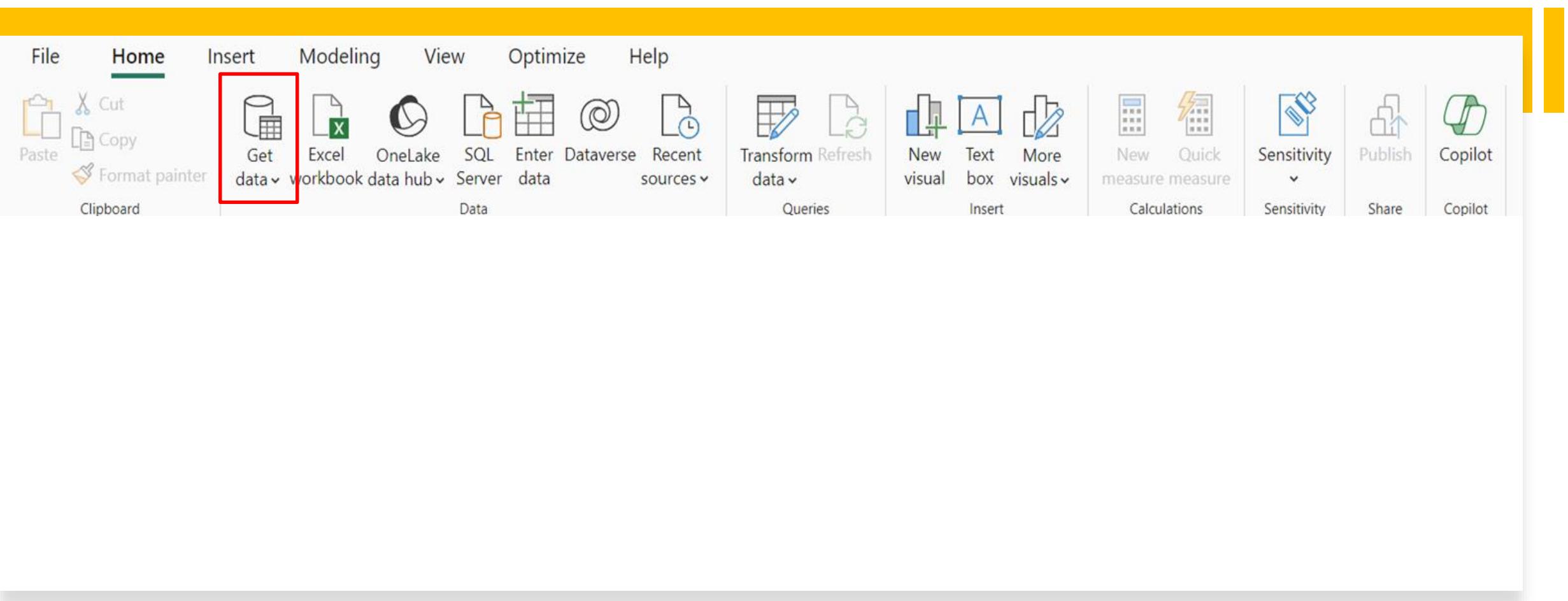
DAY



# Creating Reports and Visualizations

---

# Load data into Power BI



The screenshot shows the Microsoft Power BI ribbon interface. The 'Home' tab is currently selected, indicated by a green underline. The 'Insert' tab is also visible, with its icon highlighted by a red box. Other tabs like 'Modeling', 'View', 'Optimize', and 'Help' are present but not selected.

The 'Insert' tab contains several data source icons: 'Get data' (highlighted), 'Excel', 'OneLake', 'SQL Server', 'Enter data', 'Dataverse', and 'Recent sources'. Below these are 'Transform data' and 'Refresh' buttons. The 'Queries' section includes 'New visual', 'Text box', 'More visuals', 'New measure', 'Quick measure', and 'Sensitivity'. The 'Insert' section has 'Calculations', 'Sensitivity', 'Share', and 'Copilot' options. The 'Clipboard' section includes 'Paste', 'Cut', 'Copy', 'Format painter', and 'Clipboard' buttons.

# Data cleansing and blending

## ➤ **Data Transformation:**

- Data transformation converts data into a required structure through cleansing activities and applying business logic.

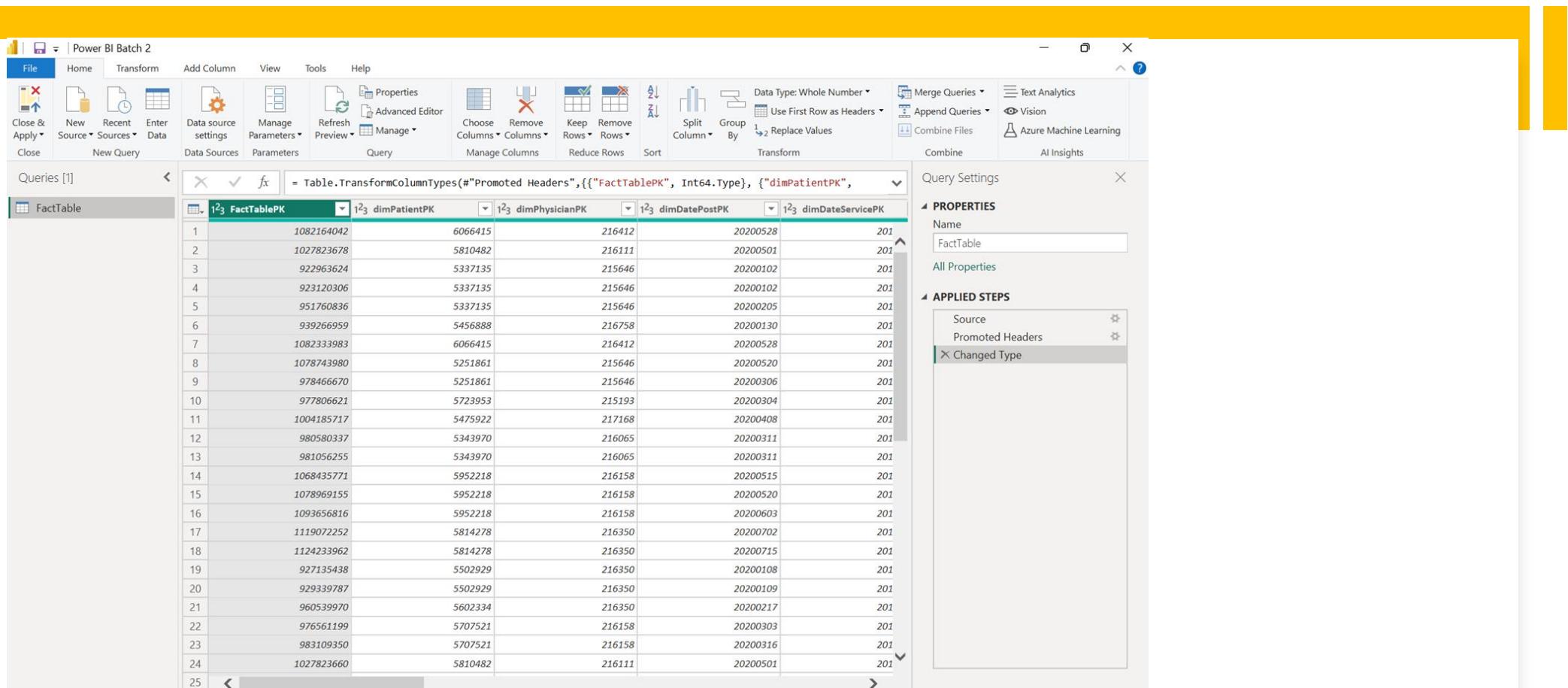
## ➤ **Power Query:**

- Utilizes ETL (Extract, Transform, Load) technology within Power BI for data transformations.
- Accessed via the "Edit Queries" button in the top menu bar, which opens the Power Query Editor window.
- Provides various tools for performing data transformations.

## ➤ **Power Query Features:**

- Microsoft's data connectivity and data preparation technology.
- Allows business users to engage with data import and preparation tasks seamlessly.
- Code-free graphical user experience speeds up data preparation.
- Typically, 80% of the time is spent connecting and preparing data, which is streamlined by Power Query.

# Power Query Editor



The screenshot shows the Microsoft Power BI Power Query Editor interface. The title bar reads "Power BI Batch 2". The ribbon menu includes File, Home, Transform, Add Column, View, Tools, and Help. The Home tab is selected. The ribbon also contains sections for Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, and various Transform tools like Choose Columns, Remove Columns, Keep Rows, Remove Rows, Split Column, Group By, Replace Values, Merge Queries, Append Queries, Combine Files, and AI Insights.

The main area displays a table titled "FactTable" with 24 rows of data. The columns are labeled "FactTablePK", "dimPatientPK", "dimPhysicianPK", "dimDatePostPK", and "dimDateServicePK". The data consists of various numerical values across these columns.

On the right side, there are two panes: "Query Settings" and "PROPERTIES". The "PROPERTIES" pane shows the "Name" is set to "FactTable". The "APPLIED STEPS" pane lists the steps taken: "Source", "Promoted Headers", and "Changed Type".

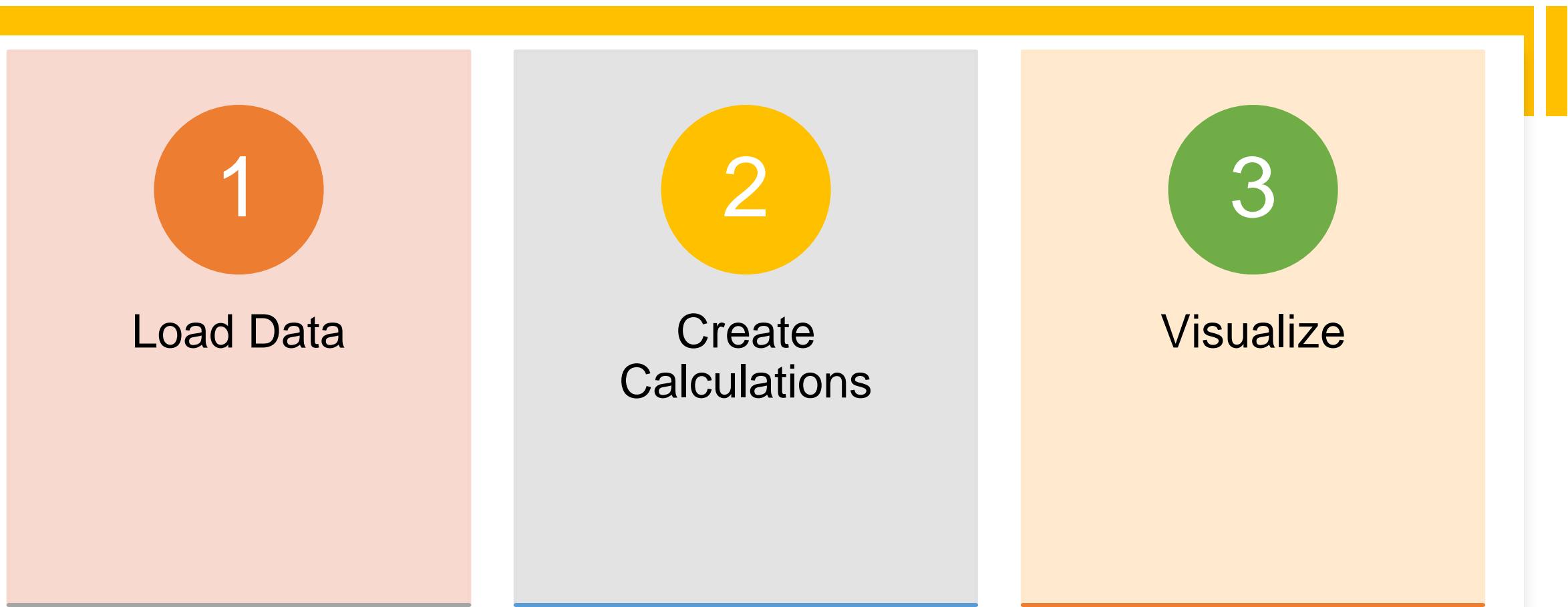
# Dimensional Modeling Concepts

---

# What is Data Modelling and Why?

- Data modeling involves designing a logical structure to store data in a database.
- Considers aspects like accessibility and optimization.
- Dimensional modeling is a subset, specifically for report databases or data warehouses.
- Discuss classical approach and data modeling approach in detail and their benefits.

# Classic BI Approach

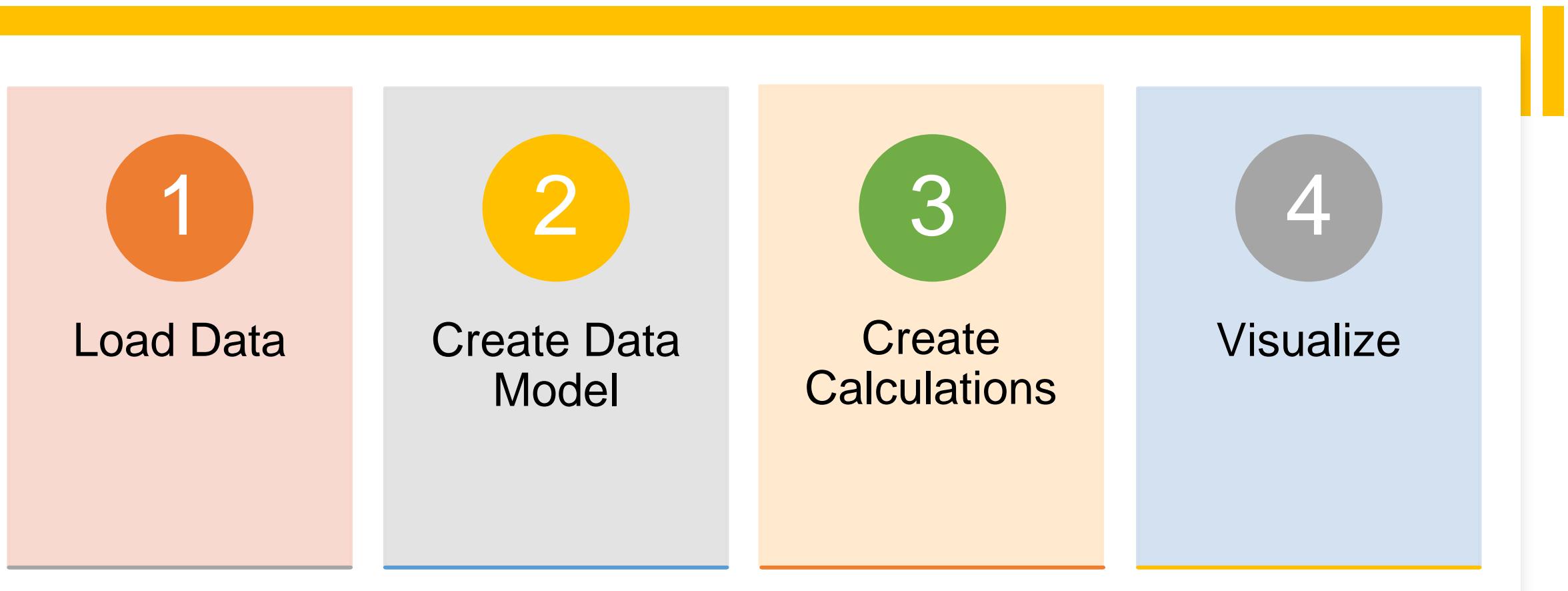


- Deal with multiple data source systems (e.g., sales, marketing, finance, HR, healthcare, logistics).
- Systems known as Online Transactional Processing (OLTP) systems.
- Highly normalized and optimized for transaction recording.
- Classic BI approach includes:
  - Connecting to data source.
  - Defining metrics.
  - Creating interactive visuals.
- Traditional report development approach: Connect to data source, define metrics, create interactive visuals..

# Classic BI Approach Limitation

- Works fine for simple datasets with few tables and low data volume.
- Difficult to analyze with large datasets (e.g., 50+ tables) or large data volume (approx. 1 GB).
- OLTP systems are optimized for transaction processing, not for reporting.
- Connecting and analyzing multiple tables can become a tedious task.

# Modern Self – Service BI Approach





- Overcomes limitations of traditional BI approach.
- Adds an intermediate layer between imported data and defined metrics: Data model or dimension model.
- Data modeling creates a simplified data structure optimized for analytics.
- Steps include: Load Data, Create Data Model, Create Calculations, Visualize.

# End – to – End BI with Microsoft Power BI

- Traditional corporate BI stack replaced with simplified, self-service approach using Power BI.
- Provides tools for data extract/load, data modeling, metric definition, reporting, and collaboration.
- Integration of SSIS, SSAS Tabular, and Power BI services for data warehousing and BI stack.
- Enables end-to-end BI solutions for large data volumes.

# Data Warehouse

- **Typical OLTP Systems:**

- Highly normalized and optimized for transaction recording.
- Challenges in reporting: Slow query execution, difficulty in analytics.

- **Data Warehouse Tables:**

- **Fact Tables:**

- Store business processes, events, or transactions.

- **Dimension Tables:**

- Store textual data and entities (e.g., product, category, salesperson, customer, date, time).

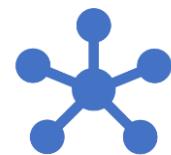
- **Dimensional Model:**

- Identified as the data warehouse structure.
- Creates a data warehouse inside Power BI.
- During implementation, map business processes to dimensions in the design phase.

# Difference Between Data Warehouse and Databases

Data Warehouse	Databases
Store large volume of data	Store small volume of data
Designed for read heavy operations	Designed for write heavy operations
High Latency	Low Latency
Denormalized (Data Redundancy High)	Highly Normalized (Data redundancy low)
Columnar Storage	Row based
Parallel Processing of request	Not optimized for parallel processing
OLAP (Online Analytics Processing)	OLTP (Online Transactional Processing)

# Characteristics of Data Warehouse



Integrated



Subject  
Oriented



Time Variant



Non Volatile

# Fact Table and Dimensional Table

- Fact Table : Measurement
- Dimension Table : Context

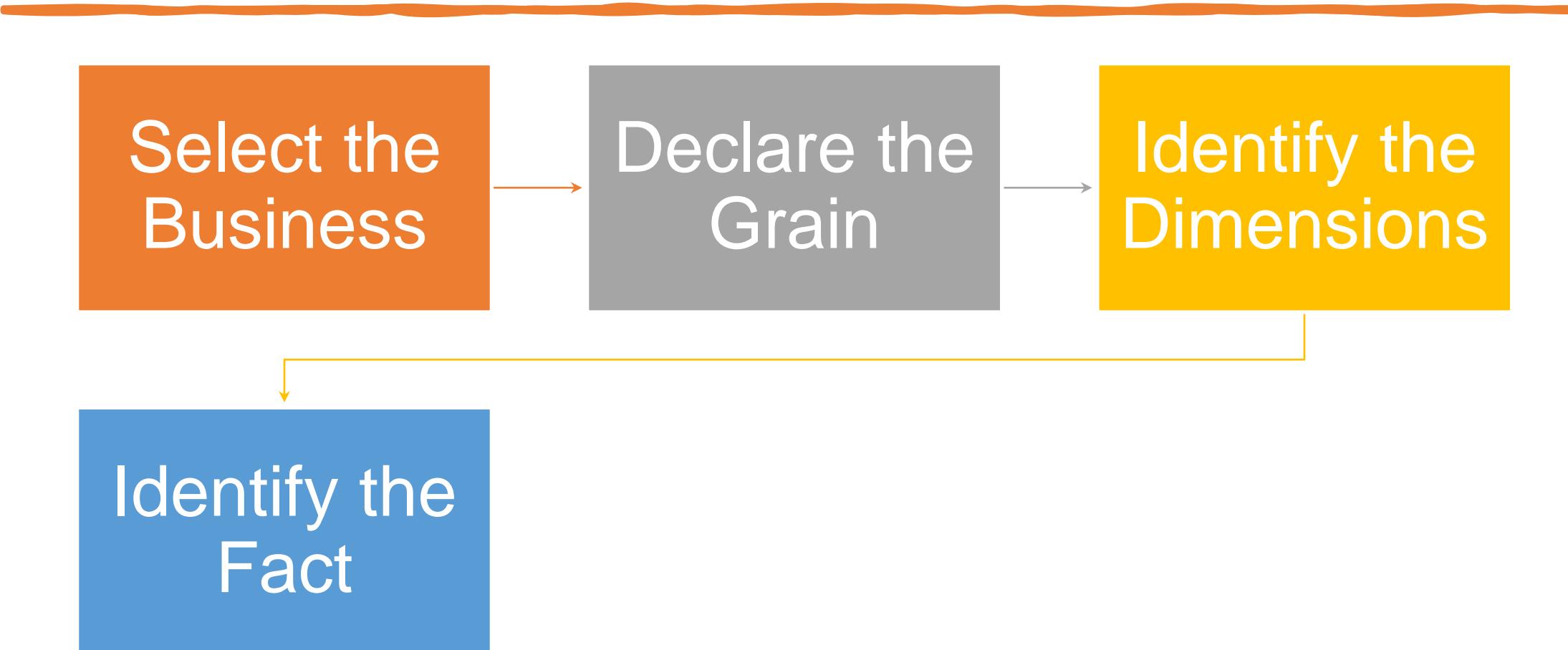
Trxn_id	Sales_quantity	Regular_unit_price	Discount_unit_price	Net_unit_price	Sales_amount	Discount_amount
TXN001	5	20	18	15	90	10
TXN002	3	120	96	100	288	72
TXN003	7	85	68	60	476	119
TXN004	1	24	21.6	20	21.6	2.4
TXN005	1	150	135	150	135	15
TXN006	2	200	160	130	320	80
TXN007	6	5	5	4	30	0

# Fact Table and Dimensional Table

R	Trxn_id
e	Sales_quantity
t	Regular_unit_price
a	Discount_unit_Price
s	Net_Unit_Price
l	Sales_amount
	Discount_amount

Pr	product_id
o	Product_description
d	Brand_description
u	Category_description
c	Package_size
m	Weight

# Fundamental of Dimensional Modelling



# Select the Business

- What does your business do?
- What measurement you want to analyze?
- How does your current Operational dataset look like?

# Declare the grain

- Granularity is the number of dimensions linked to a fact
- Grain is level of details available with the fact table. Basically this means what single record in the fact table shows.

# Identify Dimension and Fact Table

- Who, What, When, Where, How of data in the fact table
- Promotion, Date, store, Product, sales.

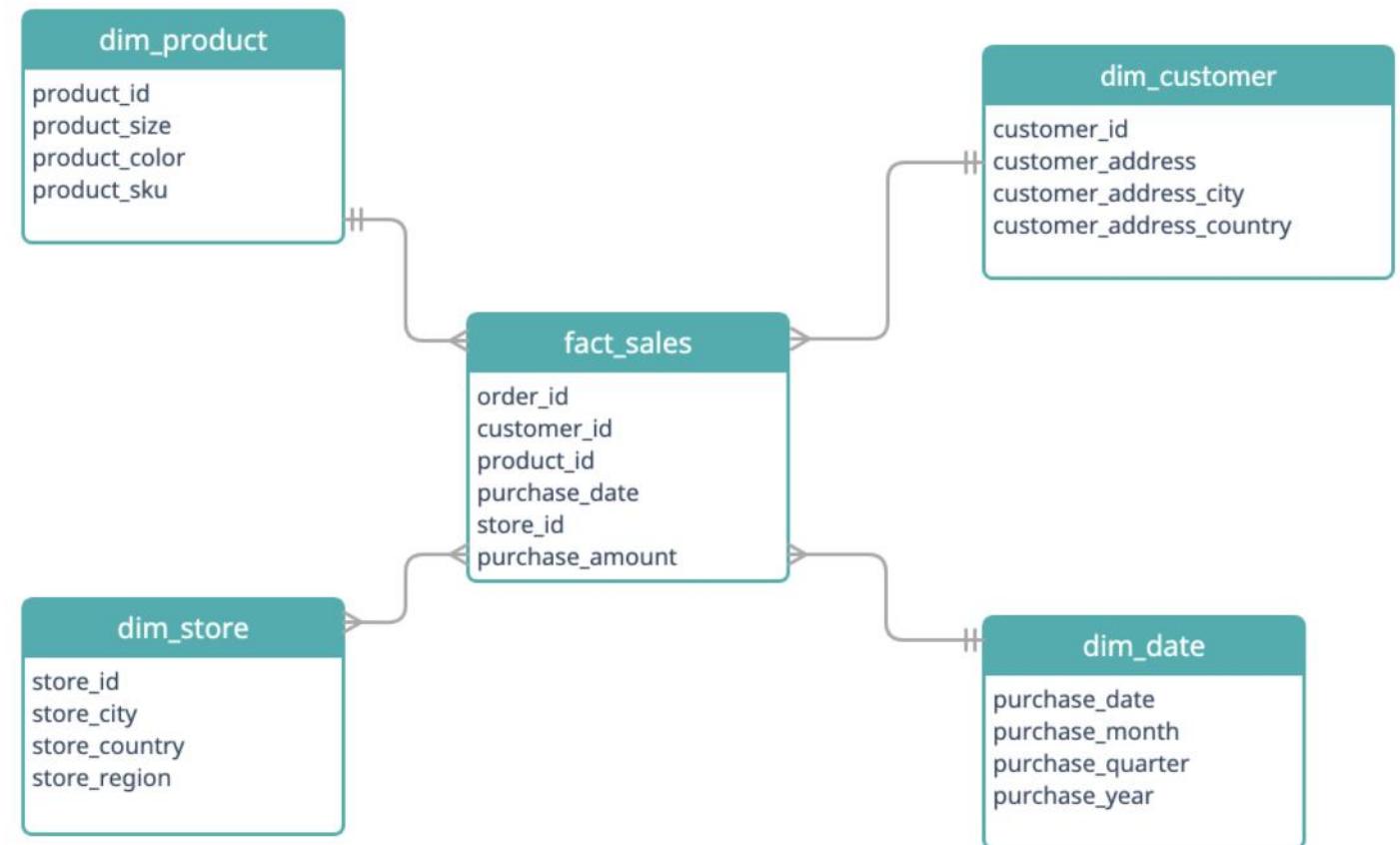
# Bus Matrix Architecture

BUSINESS PROCESSES	Dimensions					
	Date	Product	Customer	Sales Person	Location	Vendor
Sales Order	X	X	X	X		
Product Inventory	X	X			X	
Purchase Order	X	X				X

# Star Schema

- A star schema is a data model that stores information in multiple table types: a single fact table and multiple dimensional tables.
- In contrast to the classical database design of normalizing tables, star schemas connect dimensional data with fact data in a shape resembling a star (hence the name)

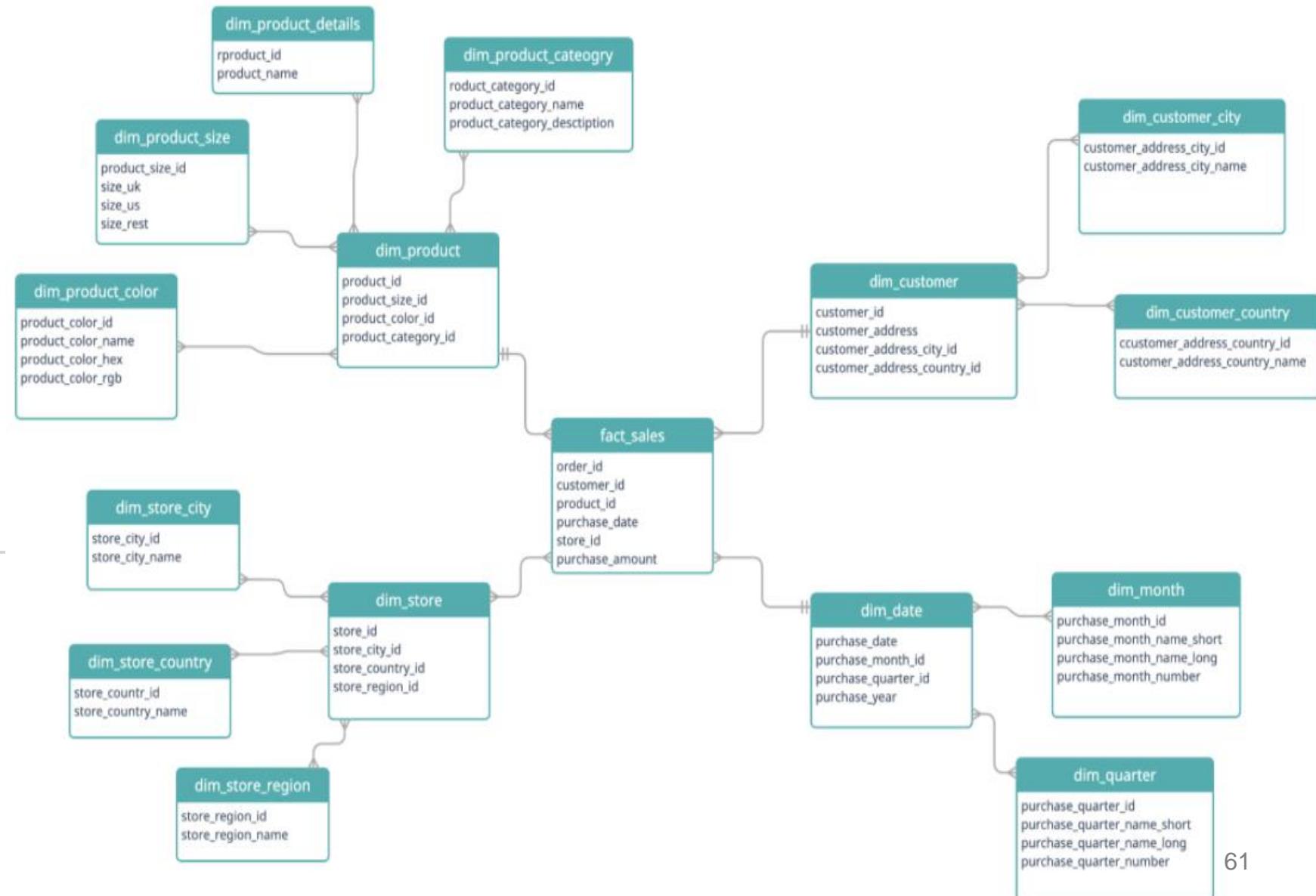
# Star Schema



# Snowflake Schema

- A snowflake schema is very similar to the simple star schema above. The main difference is that snowflake schemas split dimensional tables into further dimensional tables (also called lookup tables).

# Snowflake Schema



# Differences in Star Schema and Snowflake Schema

	<b>Star schema</b>	<b>Snowflake schema</b>
Normalization of dimension tables	normalized	denormalized
Data redundancy	stores it	avoids it
Query complexity	simple	complex
Query performance	faster	slower
Disk space	more	less
Data integrity	higher risk	lower risk
Set up and maintenance	easier to set up / harder to maintain	harder to set up / easier to maintain

# Cardinality



One-to-One (1:1)



One-to-Many  
(1:N)



Many-to-Many  
(N:N)

# One-to-One (1:1)

This cardinality indicates that each value in the column of one table is related to a unique value in the column of another table. It implies a direct and singular relationship between the two tables.

# One-to-Many (1:N)

This cardinality indicates that each value in the column of one table can be related to multiple values in the column of another table. It represents a hierarchical or parent-child relationship, where one record in the parent table can have multiple related records in the child table.

# Many-to-Many (N:N)

This cardinality indicates that multiple values in the column of one table can be related to multiple values in the column of another table. It implies a complex relationship where multiple records in one table can have associations with multiple records in another table.

# Joins

**LEFT OUTER**



All rows from the left table, matching rows from the right table

**RIGHT OUTER**



All rows from the right table, matching rows from the left table

**FULL OUTER**



All rows from both tables

**INNER**



Only matching rows from both tables

**LEFT ANTI**



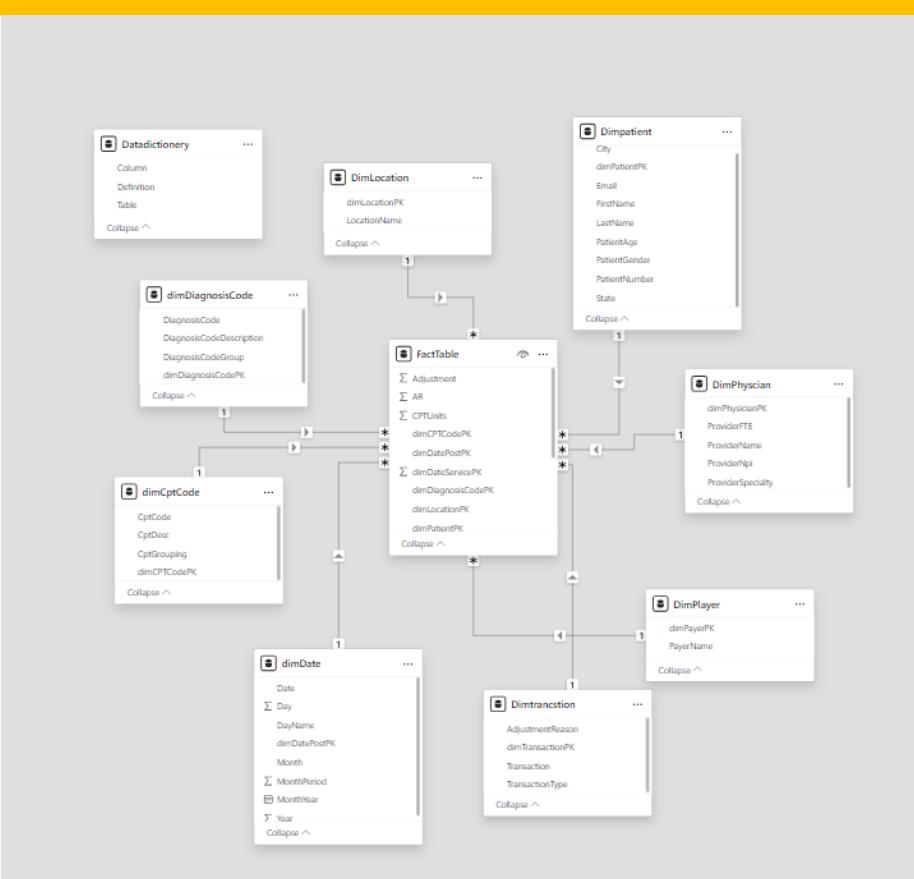
Only rows from the left table

**RIGHT ANTI**



Only rows from the right table

# Dimensional Modelling in Project





**DATA**PLAY

# Let's Create a Dashboard

# Hospital Management

- Overview Page
- Patient Analysis
- Physician Performance
- Financial Performance
- Transaction Details

# Hospital Management

## Page 1: Overview Dashboard

**Purpose:** Provide a high-level summary of key metrics.

- **Key Visuals:**

- **KPIs:** Total Gross Charges, Total Payments, Total Adjustments, Total Accounts Receivable (AR)
- **Bar Chart:** Gross Charges by Month
- **Pie Chart:** Payments by Payer Category
- **Donut Chart:** Distribution of Transaction Types

- **Interactivity:**

- Filters: Date range, Payer Category, Transaction Type

# Hospital Management

## Page 2: Patient Analysis

**Purpose:** Analyze patient demographics and activity.

- **Key Visuals:**

- **Table:** Patient List with columns (Patient Number, First Name, Last Name, Email, Gender, Age, City, State)
- **Bar Chart:** Number of Patients by Age Group
- **Map:** Patient Distribution by City and State
- **Pie Chart:** Gender Distribution of Patients

- **Interactivity:**

- Filters: Gender, Age Group, City, State

# Hospital Management

## Page 3: Physician Performance

**Purpose:** Evaluate physician performance and workload.

- **Key Visuals:**

- **Table:** Physician List with columns (Provider NPI, Provider Name, Specialty, FTE)
- **Bar Chart:** Gross Charges by Physician
- **Stacked Bar Chart:** CPT Units by Physician and Specialty
- **Line Chart:** Physician Workload (FTE) over Time

- **Interactivity:**

- Filters: Specialty, Physician Name

# Hospital Management

## Page 4: Financial Performance

**Purpose:** Monitor financial performance indicators.

- **Key Visuals:**

- **Line Chart:** Monthly Gross Charges vs. Payments
- **Stacked Bar Chart:** Adjustments by Reason
- **Gauge Chart:** Outstanding AR as a Percentage of Gross Charges
- **Heatmap:** Payments by Payer and Month

- **Interactivity:**

- Filters: Date range, Payer, Adjustment Reason

# Hospital Management

## Page 5: Transaction Details

**Purpose:** Dive into the details of individual transactions.

- **Key Visuals:**

- **Table:** Transaction List with columns (Transaction Type, Transaction Description, Adjustment Reason, CPT Code, CPT Description, Gross Charge, Payment, Adjustment, AR)
- **Bar Chart:** Transactions by Type
- **Treemap:** Adjustments by Reason
- **Line Chart:** AR Trends Over Time

- **Interactivity:**

- Filters: Transaction Type, Adjustment Reason, CPT Code

# Drill - Through

- Filters
- Bookmarks
- Interactivity of slicers

# RLS (Row Level Security)

Row-level security (RLS) with Power BI can be used to restrict data access for given users. Filters restrict data access at the row level, and you can define filters within roles.

# Types of RLS

- Static RLS
- Dynamic RLS

# DAX for creating tables and columns

DAX FUNCTIONS IN POWER BI



# DAX stands for data analysis expressions

- DAX is a formula expression language used in multiple Microsoft analytics tools



- DAX formulas include functions, operators and values to perform advanced calculations
- DAX formulas are used in:
  - **Measures**
  - **Calculated columns**
  - **Calculated tables**
  - Row-level security



# Measures vs calculated columns

## Calculated Columns:

- Calculated on data import
- Visible in data & report Pane

```
COST = Orders[Sales] - Orders[Profit]
```

Order_ID	Sales	Pofit	Cost
3151	\$77.88	\$3.89	\$73.99
3152	\$6.63	\$1.79	\$4.84
3153	\$22.72	\$10.22	\$12.50
3154	\$45.36	\$21.77	\$23.59

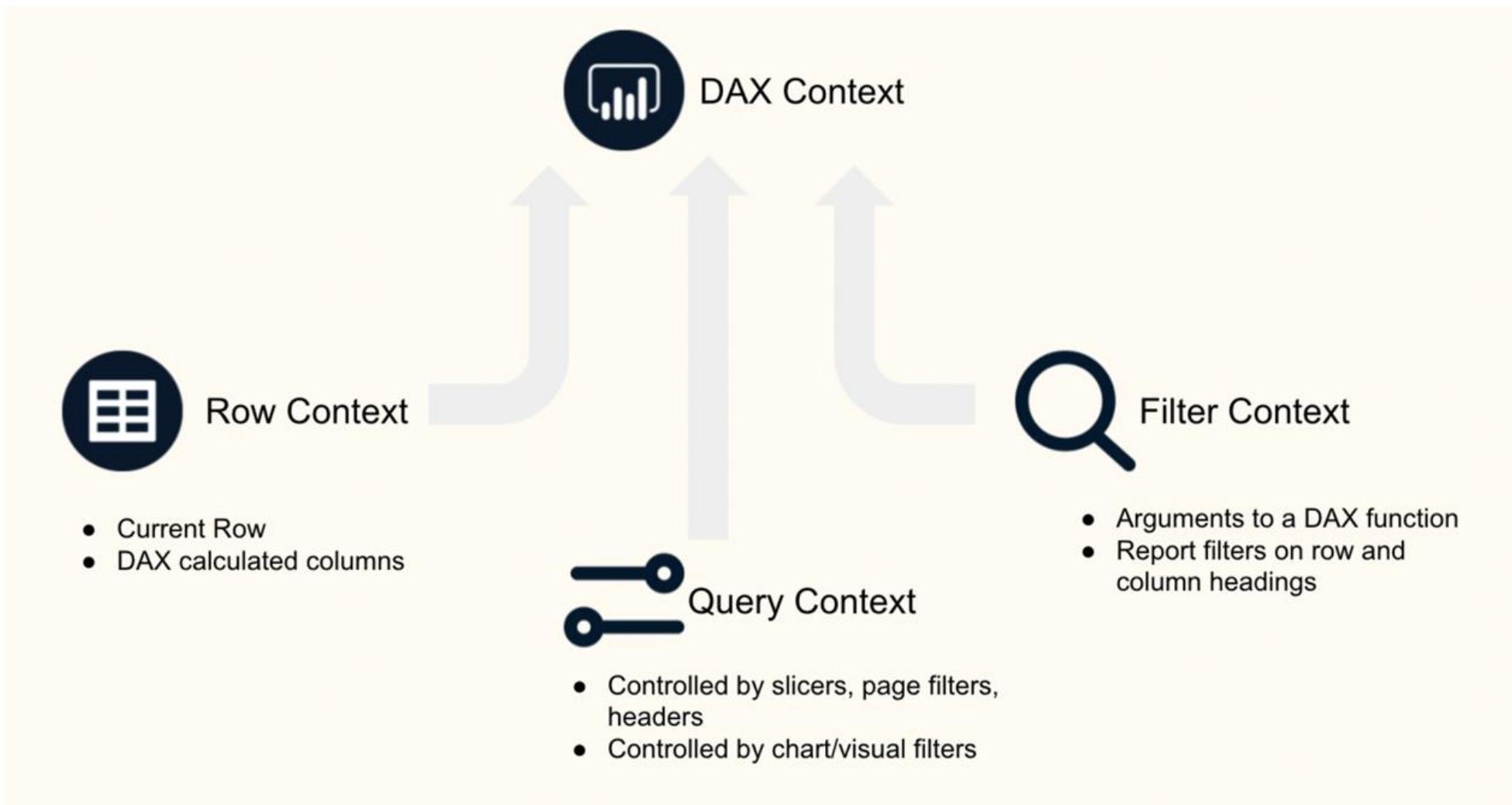
## Measures:

- Calculated at query run-time
- Visible only in report pane

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

Region	Total Sales
Central	\$501,239.89
East	\$678,781.24
West	\$391,721.91
South	\$725,457.82
Total	\$2,297,200.86

# Context in a nutshell



# Let's practice!

DAX FUNCTIONS IN POWER BI

# Implicit vs explicit measures

## Implicit

- Automatically created by Power BI
- Comes directly from the Database
- E.g.: If we drag `Sales` to values of a table, Power BI will automatically sum it
- Using a dropdown menu we can define the aggregation: sum, average, count, ...

## Explicit

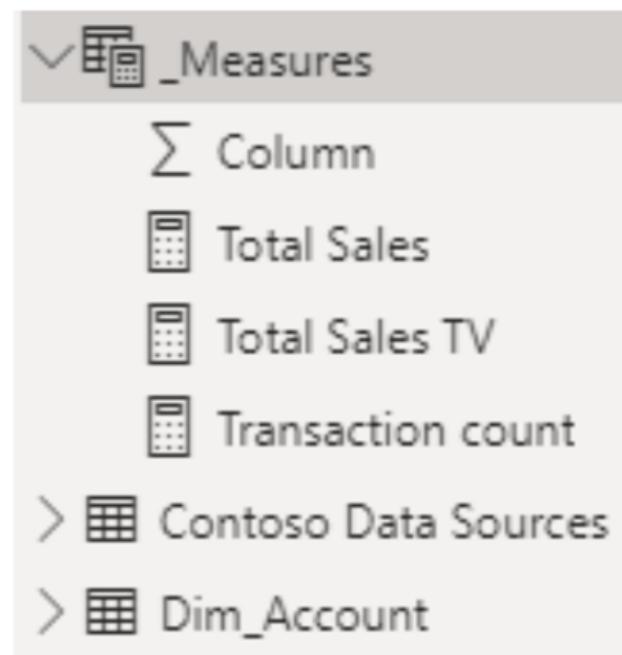
- Writing measures in an explicit way
- E.g.: `Total Sales = SUM(Orders[Sales])`
- Offer flexibility

# Why explicit measures are preferred

- Reduces confusion of what a measure is or does
  - Total Sales = SUM(Orders[Sales])
  - Total Sales is more clear than Sales (SUM, AVG, MIN, ... ?)
- Reusable within other measures
  - Total Sales East = CALCULATE([Total Sales], Orders[Region] = 'East')
- Can be given a custom name to explain its functionality
- Makes maintenance of complex models more sustainable

# Best practices

- Keep DAX measures grouped together:
  - Measures are free to move to any table
  - This is in contrast with calculated columns, which belong to a specific table
- Format and comment with DAX:
  - Use indentations to increase understanding
    - Shift Enter to start a new line
    - Tab to indent
  - Add comments after a //



# Use variables to improve your formulas

- Stores the result of an expression as a named variable
- Can be used as an argument to other measure expressions
- Four main advantages:
  - Improve performance
  - Improve readability
  - Simplify debugging
  - Reduce complexity

## Syntax:

- `VAR <name> = <expression>`
  - Name = The name of the variable
  - A DAX expression which returns a scalar or table value
  - Followed by a `RETURN` statement

# Use variables to improve your formulas - example

- Calculate the sales from last year and store it as a variable

```
VAR  
SALESPRIORYEAR = CALCULATE([SALES],SAMEPERIODLASTYEAR('DATE'))  
RETURN
```

- Use the variable in a formula

```
Sales growth = [Sales] - SALESPRIORYEAR
```

# Use variables to improve your formulas - example

- All together it would look like this:

```
Sales growth =
```

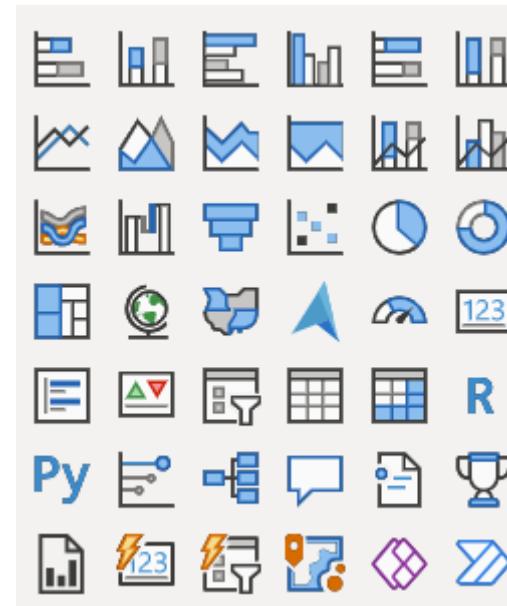
```
VAR
```

```
SALESPRIORYEAR = CALCULATE([SALES],SAMEPERIODLASTYEAR('DATE'))
```

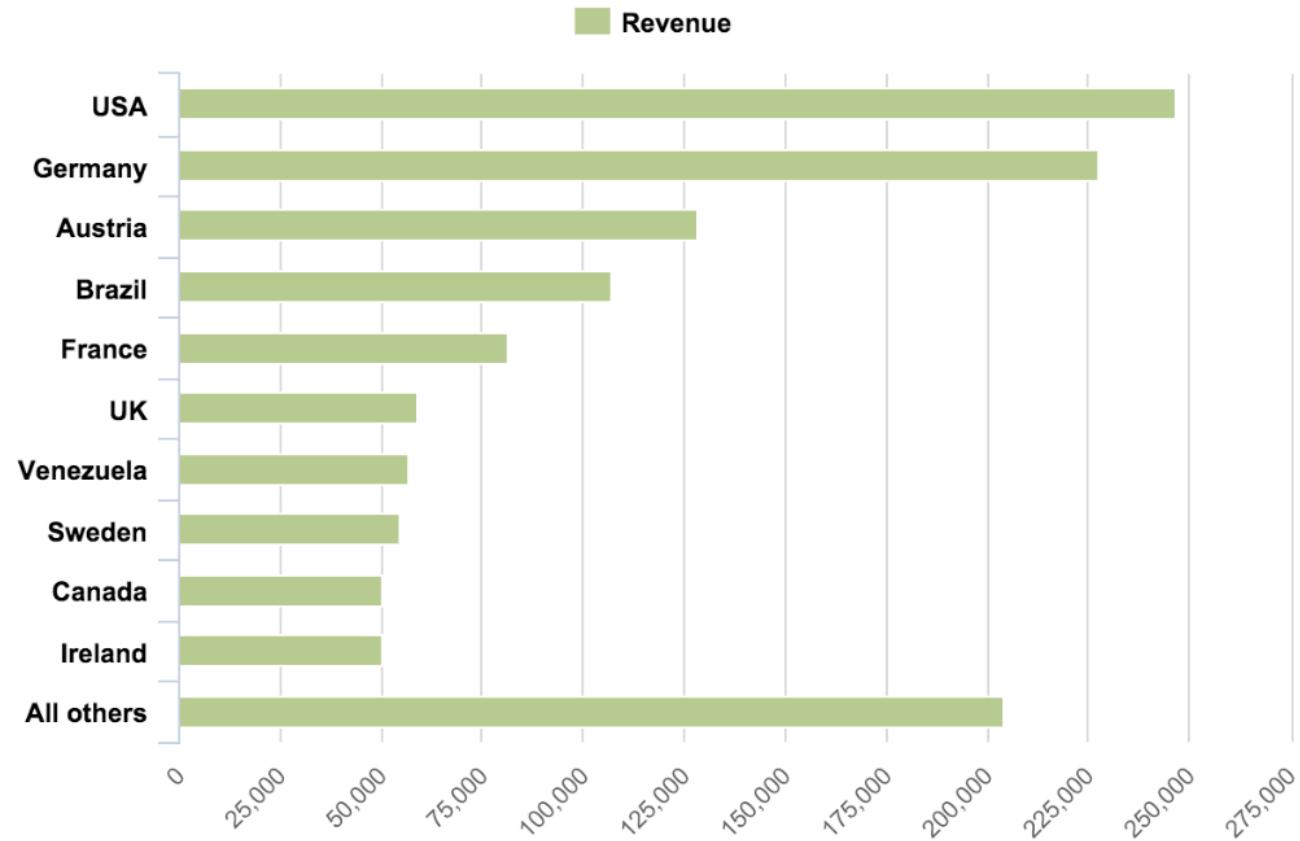
```
RETURN
```

```
Sales growth = [Sales] - SALESPRIORYEAR
```

# Choosing the Right Chart

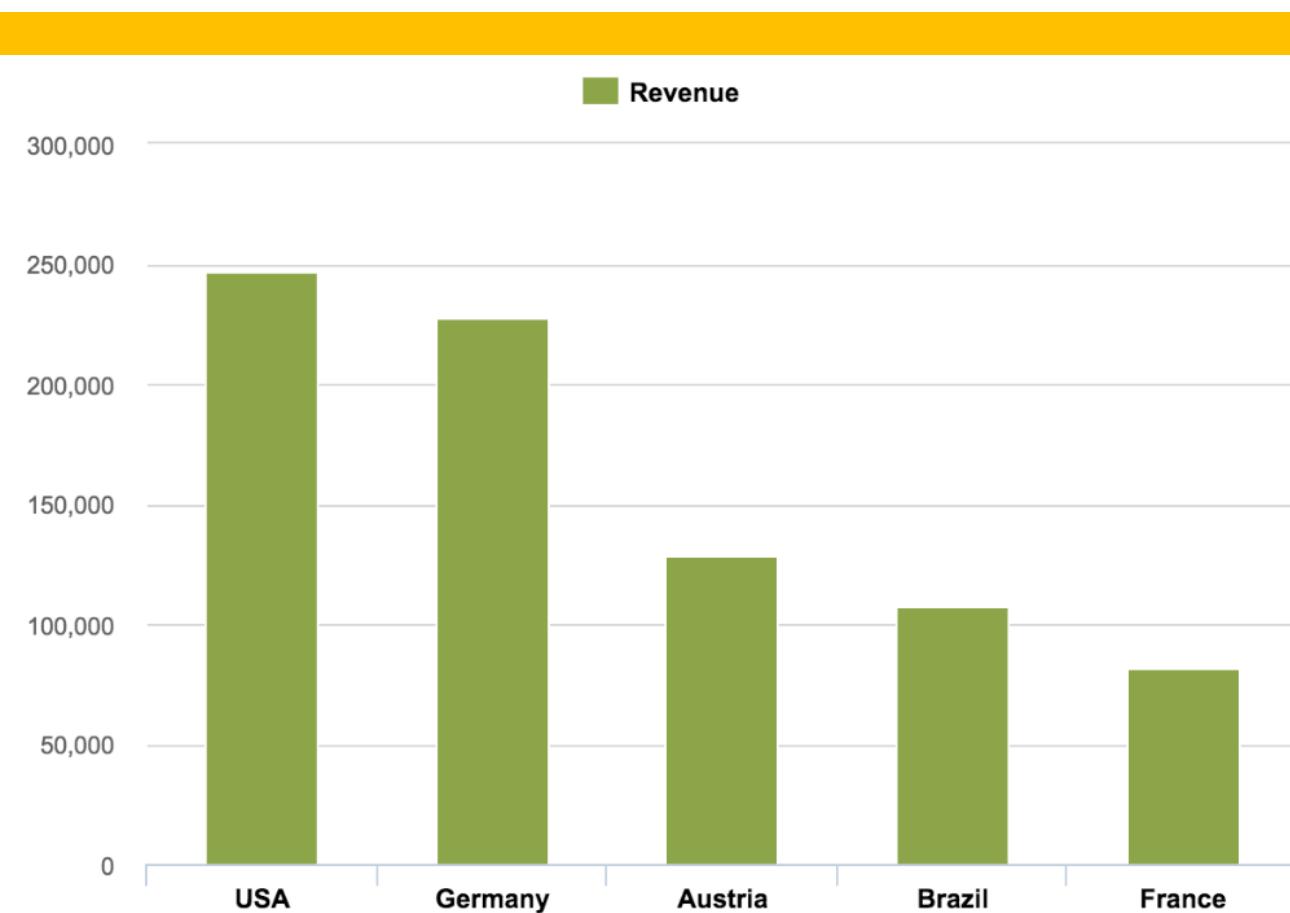


# Bar Chart



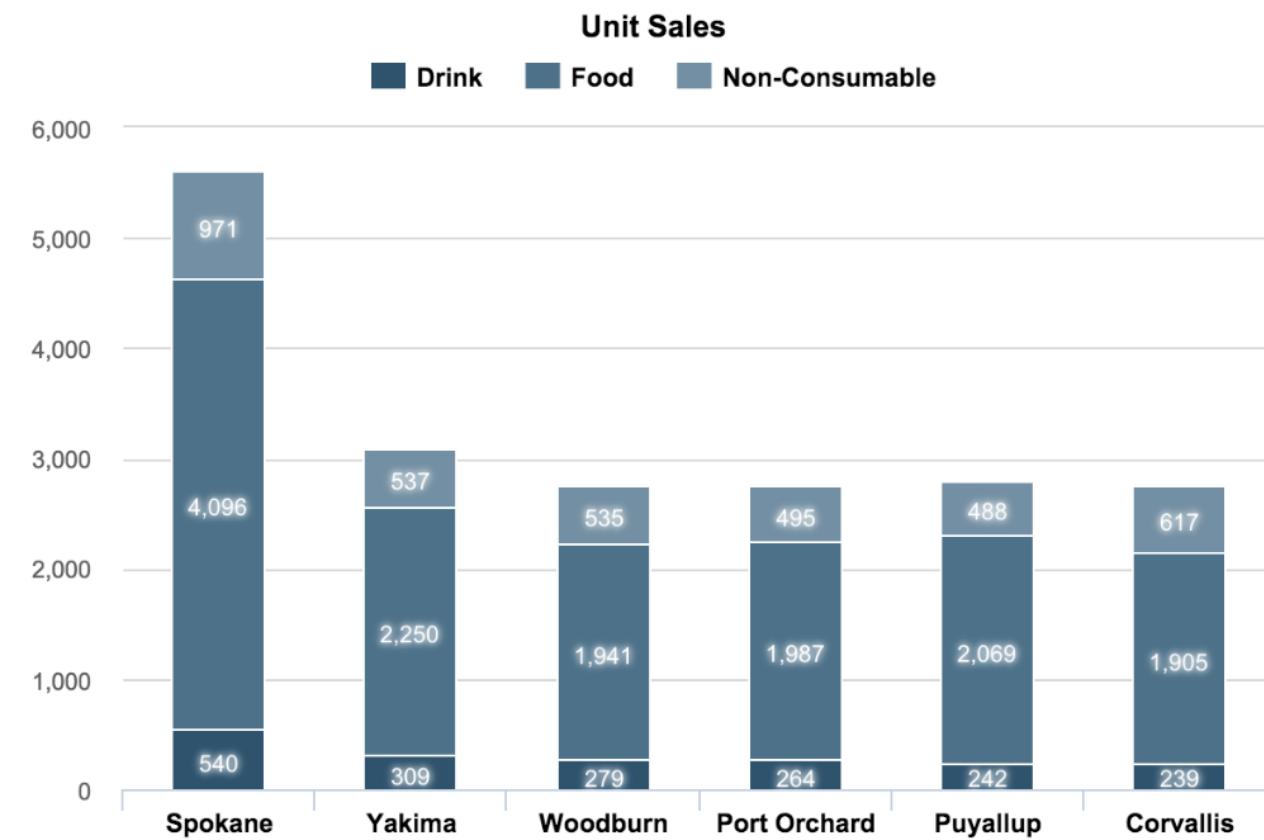
1. Categorical and Quantitative Data
2. Use a bar chart to compare data across categories

# Column Chart



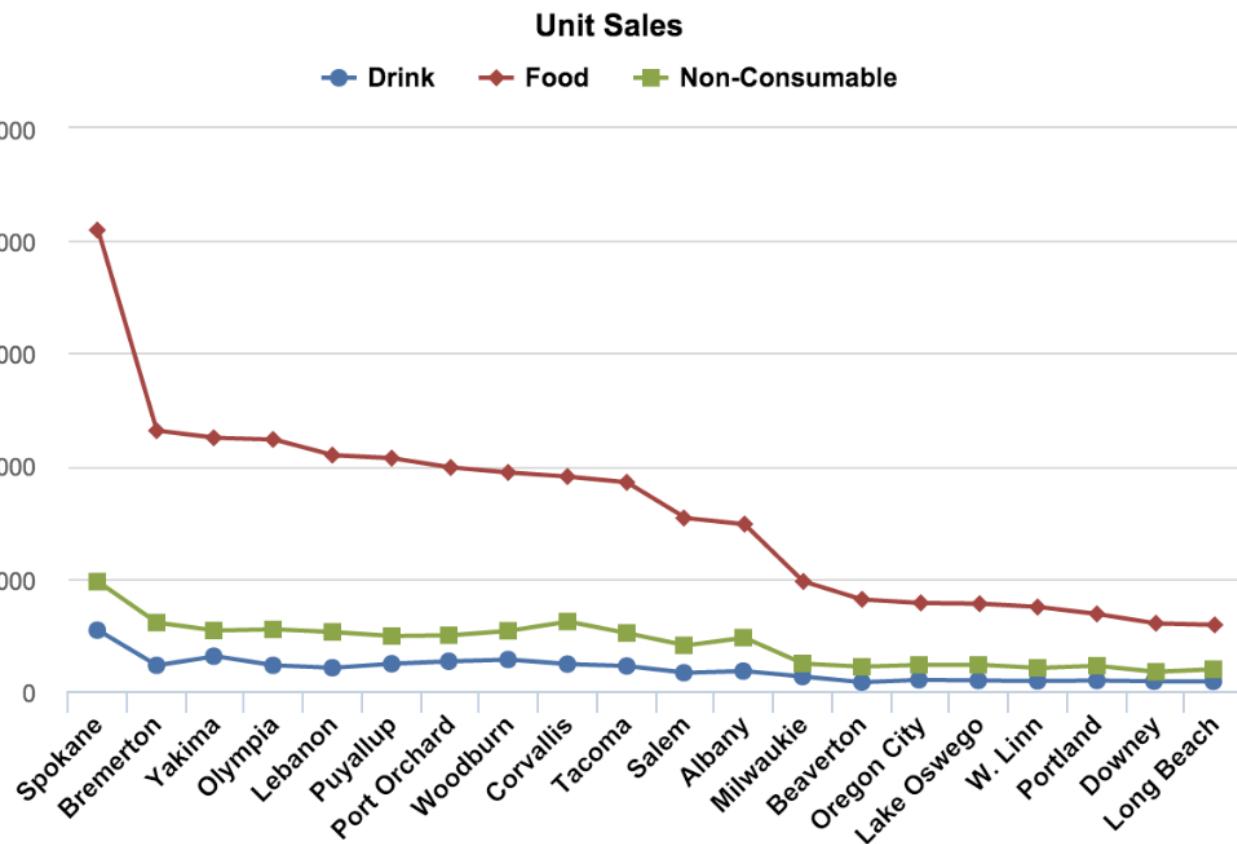
1. Continuous
2. Use the column chart to display data distribution

# Stacked Bar Chart and Column Chart



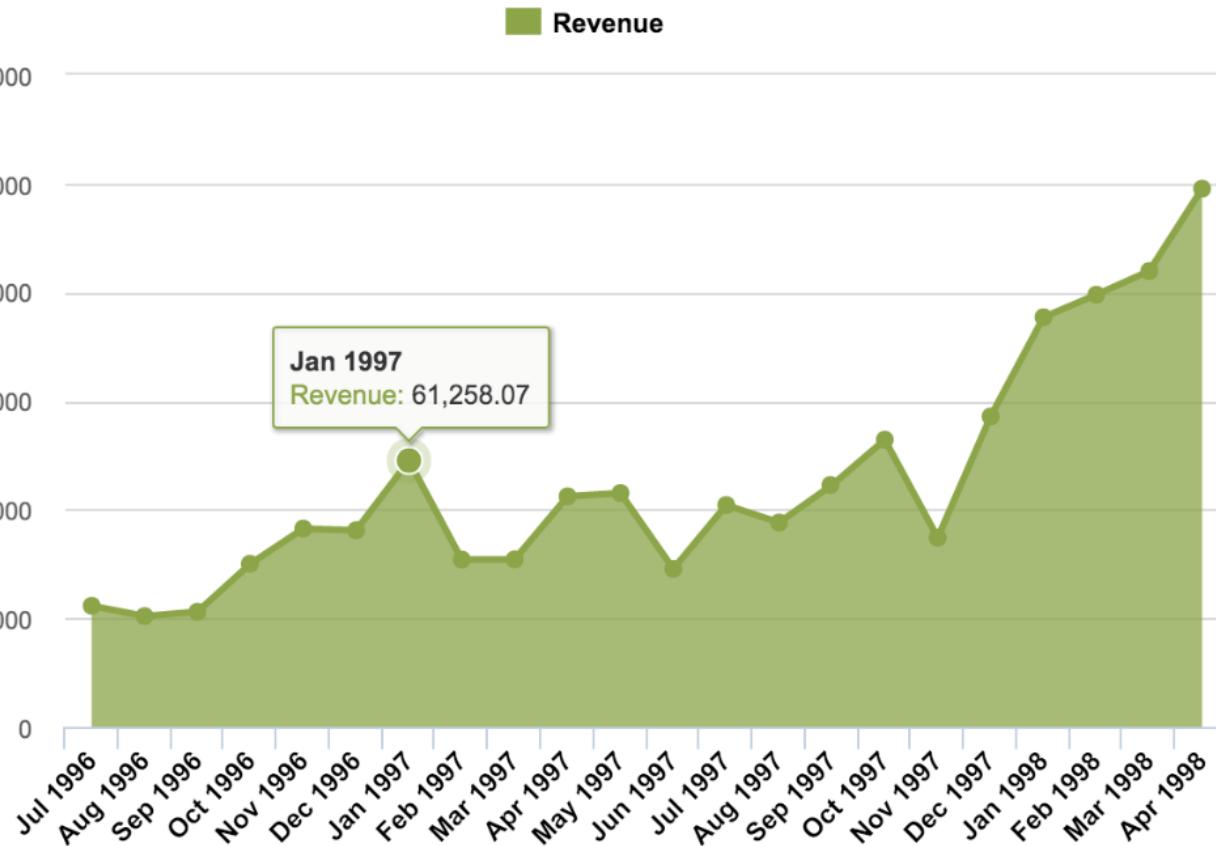
1. Continuous or Categorical and quantitative data
2. Use Stacked chart for showing composition of a column or a bar.

# Line Chart



1. Continuous and time - series
2. Use Line Chart to show trends over time.

# Area Chart



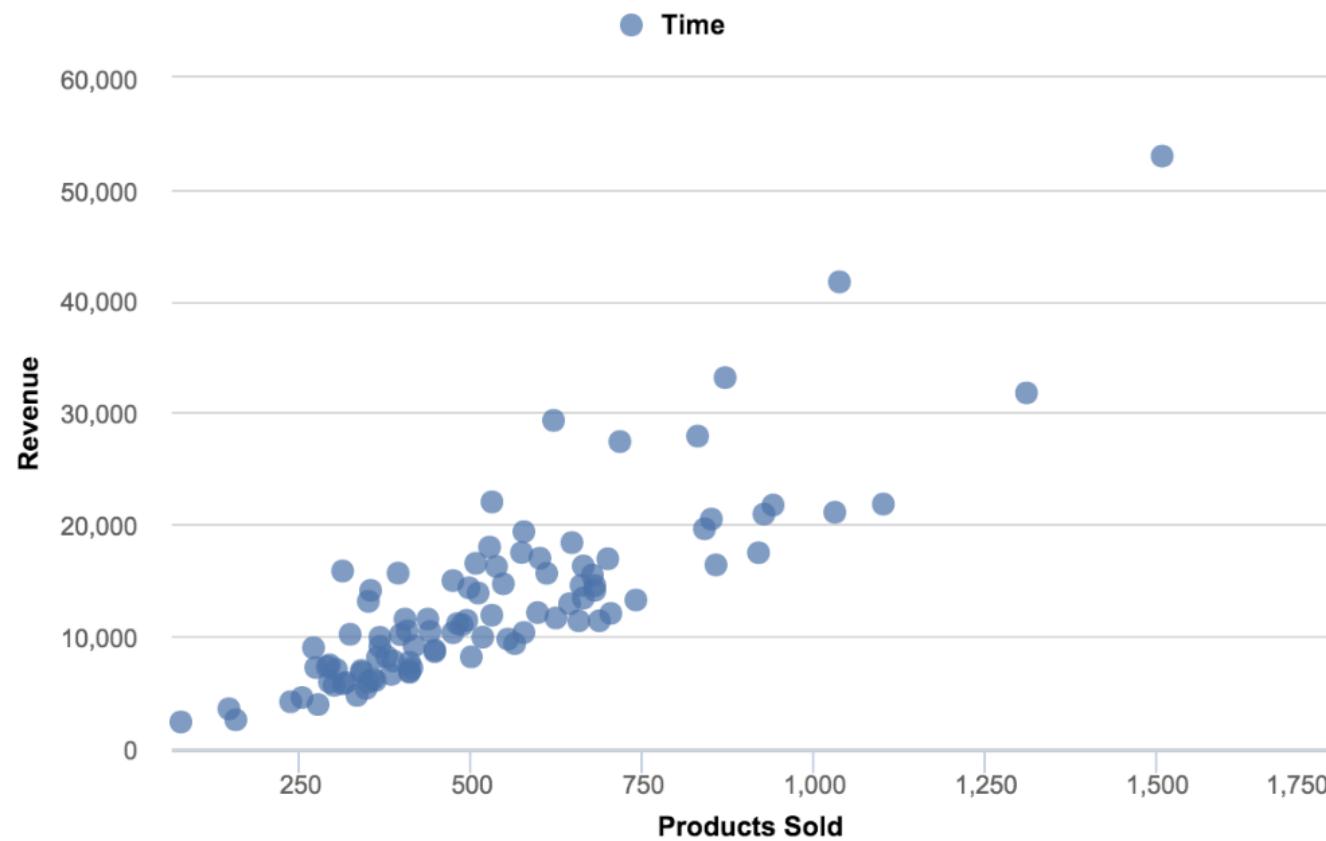
1. Continuous and time - series
2. Use and Area Chart to show the volume or magnitude of data over time.

# Waterfall Chart



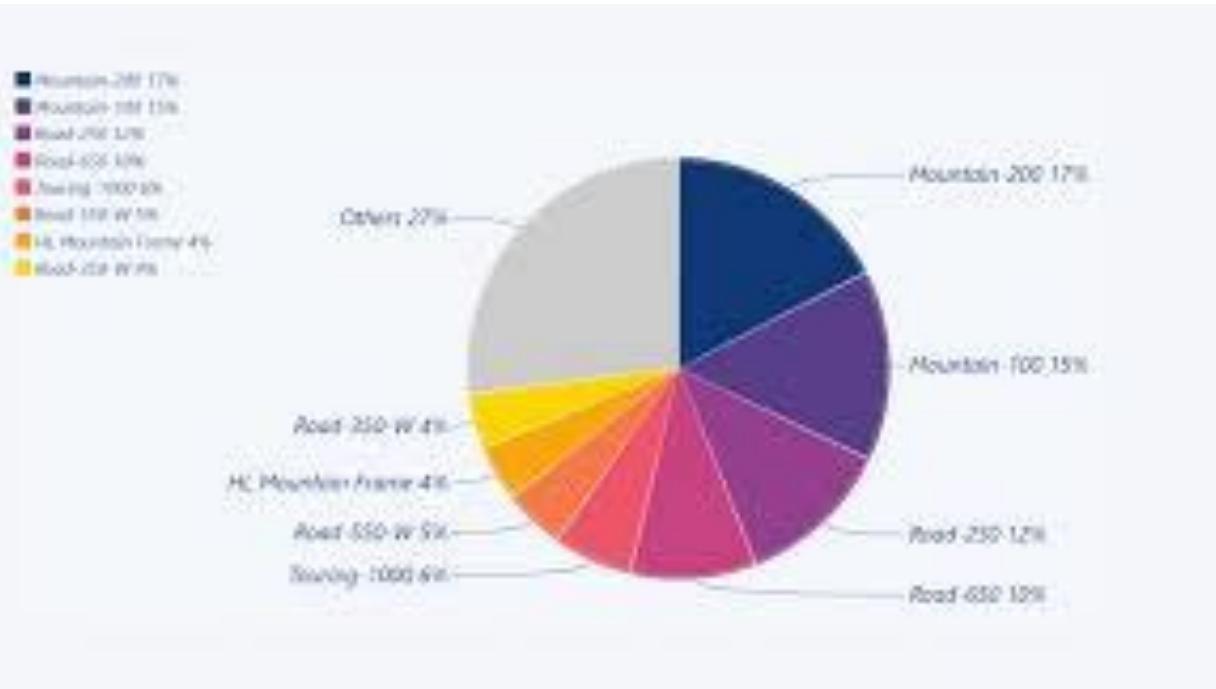
1. Quantitative and sequential
2. Use waterfall chart to visualize the cumulative effect of sequential data, such as financial or inventory changes.

# Scatter Plot



1. Continuous and Bivariate
2. Use a scatter plot to display the relationship between two variables

# Donut Chart or Pie Chart



1. Categorical and Proportional
2. Use a donut chart to show the proportion of each category.

# Tree Map



1. Categorical and Hierarchical
2. Use Tree Map to display hierarchical data or to show the proportion of each category as a whole.

