National University of Computer and Emerging Sciences



Laboratory Manuals

*for*

**Data Warehousing & Business Intelligence Lab**

(DL - 3003)

|  |  |
| --- | --- |
| Course Instructor | Mr. Ishaq Raza |
| Lab Instructor | Mr. Durraiz Waseem |
| Lab Demonstrator | Mr. Shahzeb Mubashar |
| Section | BDS-5A |

*Department of Computer Science FAST-NU, Lahore, Pakistan*

**Lab Manual 01**

**Objectives:**

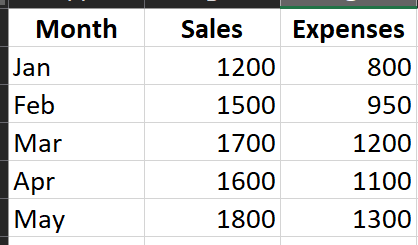
* Introduce students to the concepts of DW & BI.
* Understand plotting in Excel
* Understand the role of BI dashboards in decision-making.
* Import structured data into Power BI.
* Perform basic transformations (Power Query).
* Create simple visualizations (bar chart, pie chart, card).
* Publish insights in a storytelling dashboard.

**Plotting in Excel**

# Preparing Your Data

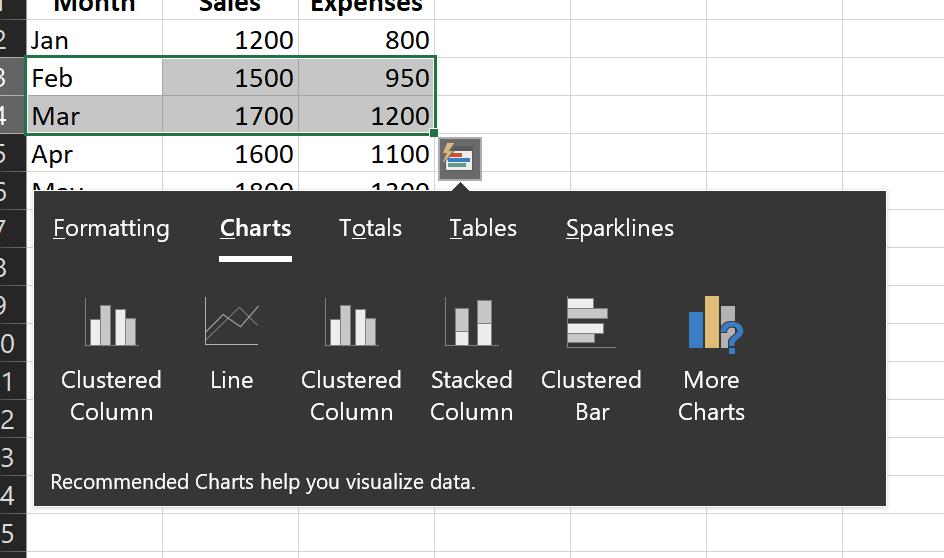
Before plotting, make sure your data is **organized in a table**:

* + **Columns = Variables/Attributes**
  + **Rows = Observations/Records**

****

# Selecting Data for Plotting

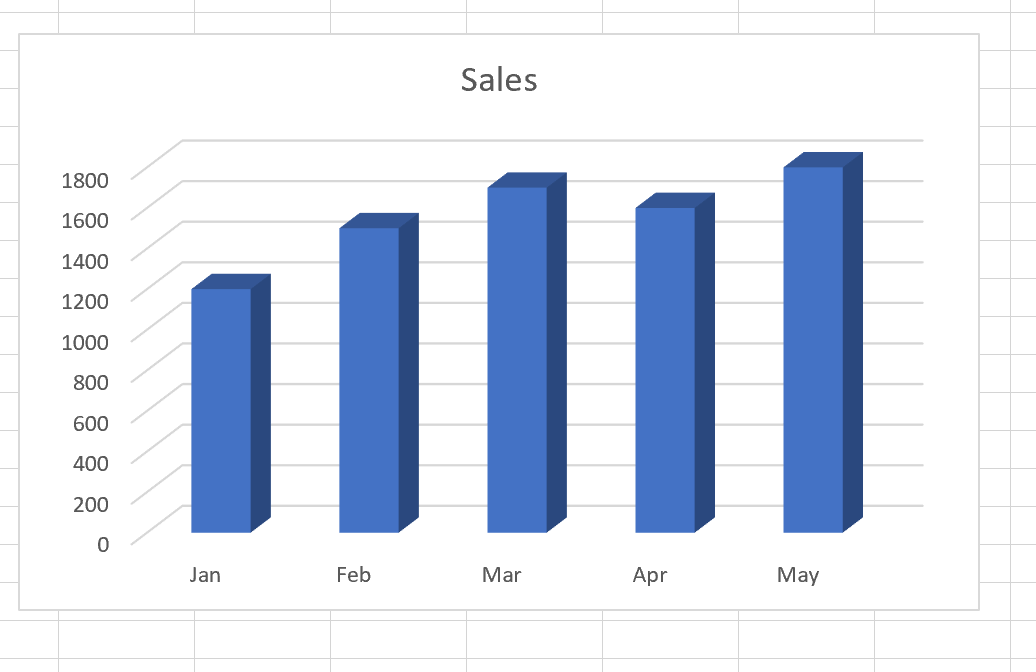
1. Highlight the data you want to plot (e.g., Month and Sales).
2. Go to the Insert tab on the Ribbon.
3. Look at the Charts group (Bar, Column, Line, Pie, Scatter, Area, etc.).



# Types of Plots in Excel

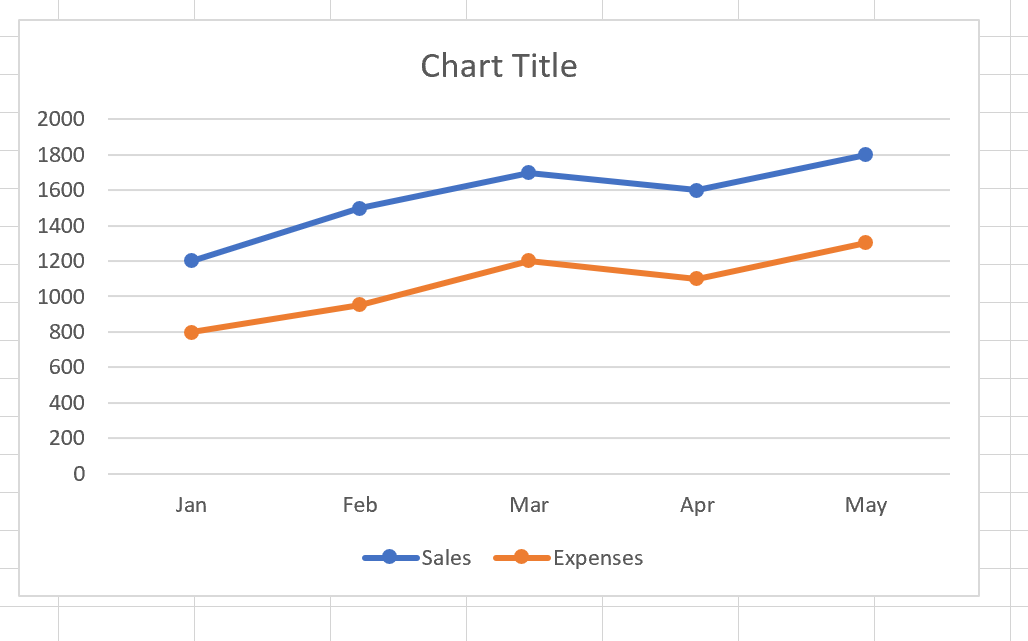
## Column/Bar Chart

* + Best for comparing categories. Steps:
  1. Select Month + Sales.
  2. Insert → Column Chart → “Clustered Column”.
  3. Customize:
     + Add Chart Title → “Monthly Sales”.
     + Right-click → Add Data Labels.
     + Change colors via Format.



## Line Chart

* Best for showing trends over time. Steps:
  1. Select Month, Sales, and Expenses.
  2. Insert → Line Chart → “Line with Markers”.
  3. Customize:
     + Add Legend (Sales vs Expenses).
     + Add Axis Titles → X-axis = Month, Y-axis = Amount ($)



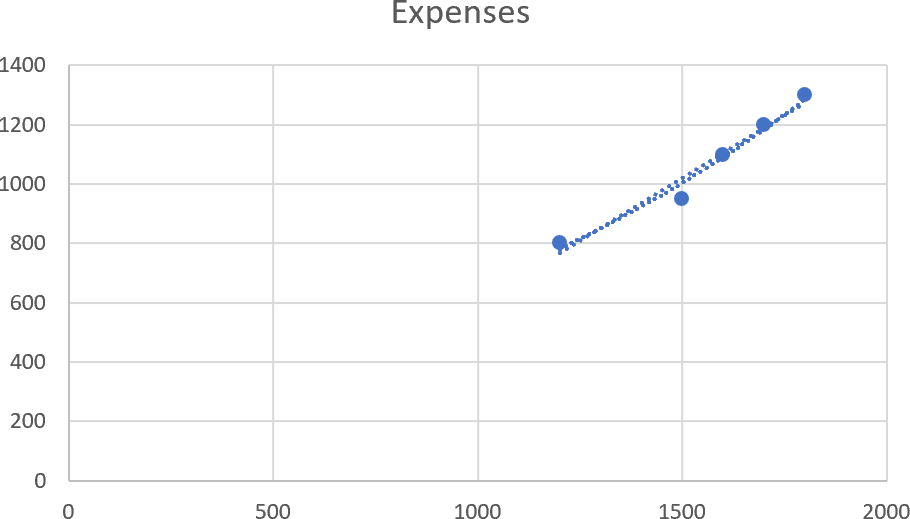
## Pie Chart

* Best for showing proportions/percentages. Steps:
  1. Select Month + Sales (or just one month’s breakdown).
  2. Insert → Pie Chart → “2D Pie”.
  3. Customize:
     + Right-click → Add Data Labels → Show Percentage.
     + Explode slices by dragging out.



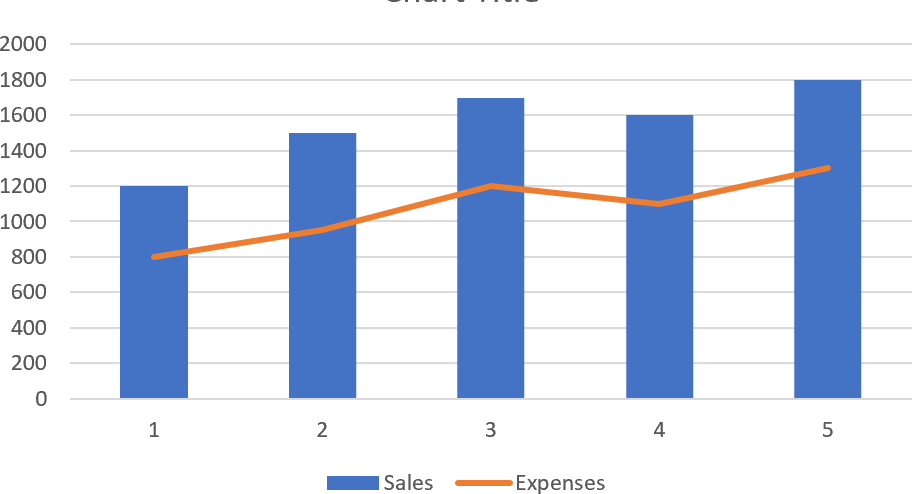
## Scatter Plot

* Best for relationships between two numeric variables. Steps:
  1. Select Sales and Expenses.
  2. Insert → Scatter → “Scatter with Markers”.
  3. Customize:
     + Add Trendline (Right-click → Add Trendline → Linear/Exponential).
     + Show Equation on Chart.



## Combo Chart

* Best when comparing different scales. Example: Sales (bars) and Expenses (line). Steps:
  1. Select all data.
  2. Insert → Combo Chart → “Custom Combo Chart”.
  3. Choose Sales = Column, Expenses = Line.
  4. Enable Secondary Axis if needed.



# Customizing Plots

### Chart Elements (+ button on chart):

* + - Add Axis Titles, Data Labels, Legends, Gridlines.

### Format Tab:

* + - Change line styles, bar colors, markers.

### Chart Design Tab:

* + - Quick Layouts (predefined styles).
    - Change chart type if needed.

# Advanced Features

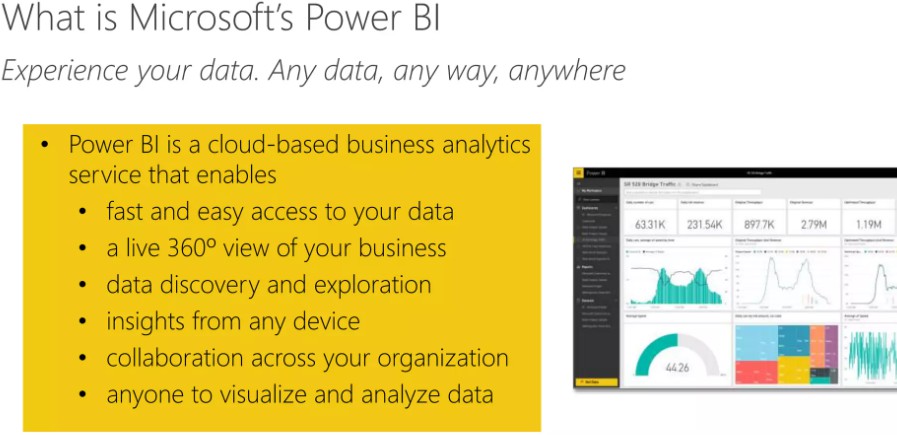
* + **Slicers (if using PivotTables):** Interactive filters for charts.
  + **Secondary Axis:** Useful when values differ in scale.
  + **Dynamic Charts:** Use tables so chart updates automatically when new rows are added.
  + **Trendlines:** Analyze growth patterns (linear, exponential, polynomial).

# Printing & Exporting Charts

* + Click on chart → Ctrl + C → Paste into Word/PowerPoint.
  + Or right-click → Save as Picture.

**Introduction to Power BI**

**What is Power BI?**

****

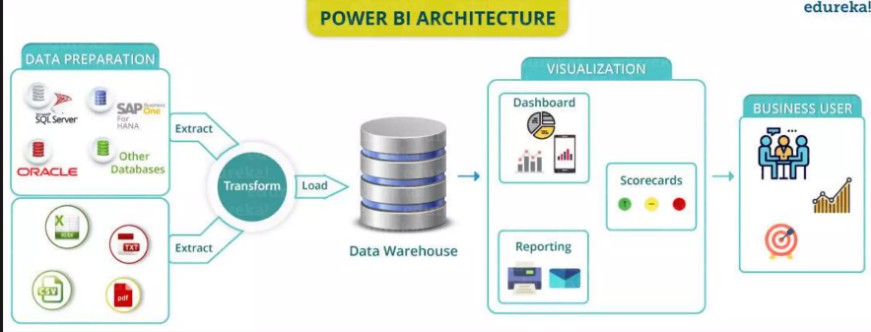
**How Does Power BI Work?**

****

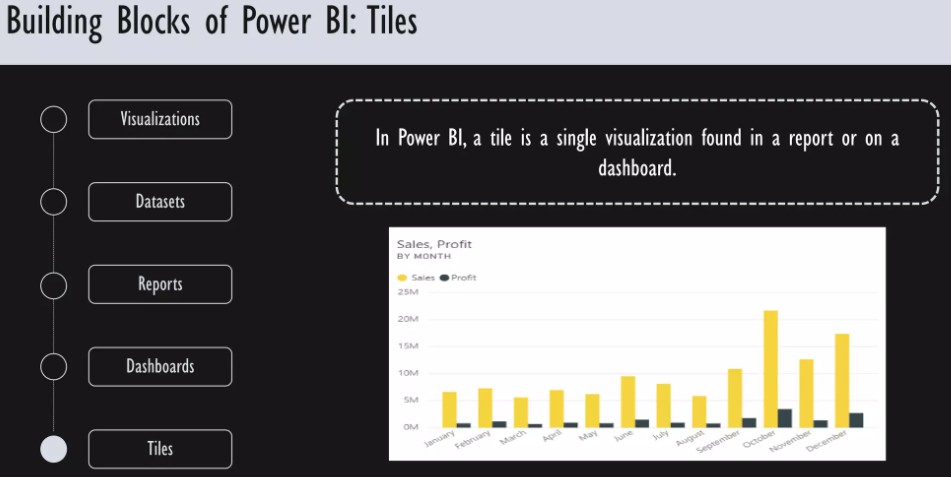
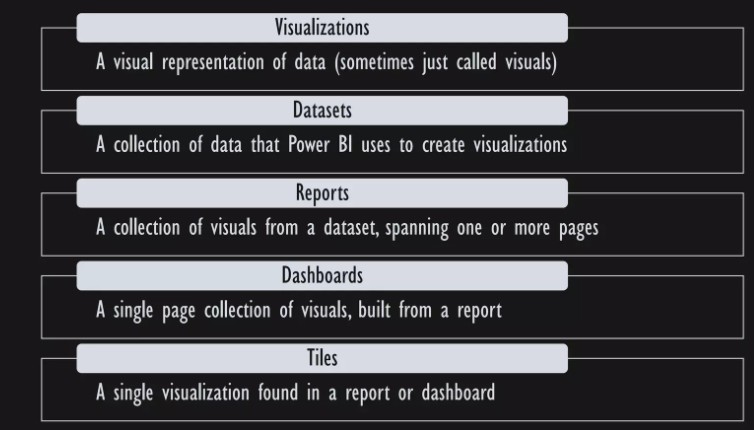
**Components of Power BI**

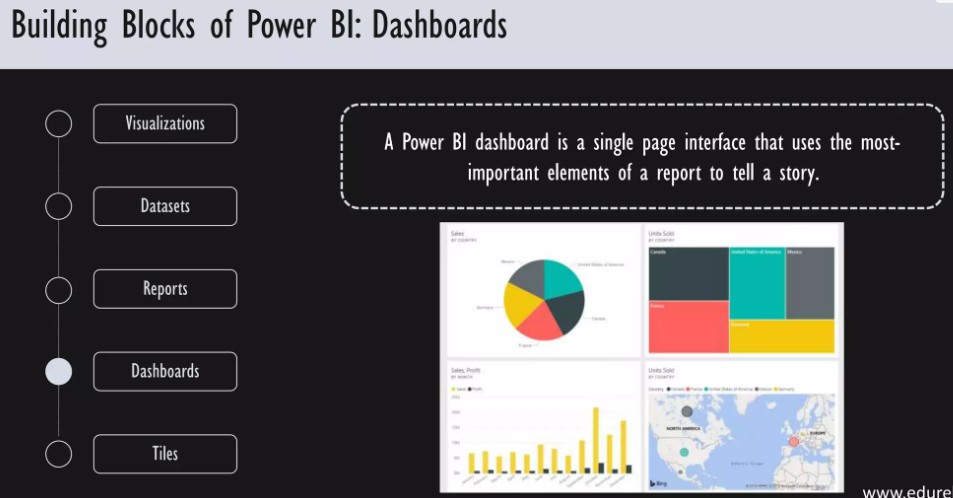
****

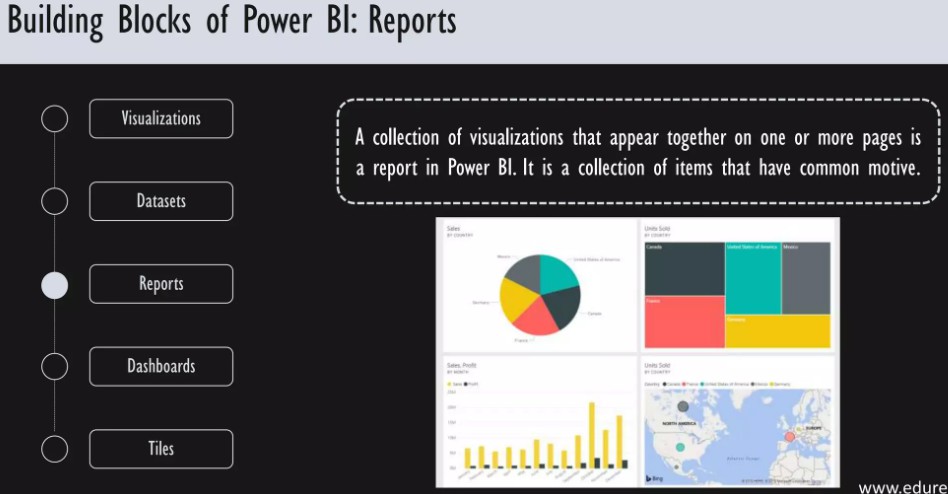
**Architecture of Power BI**

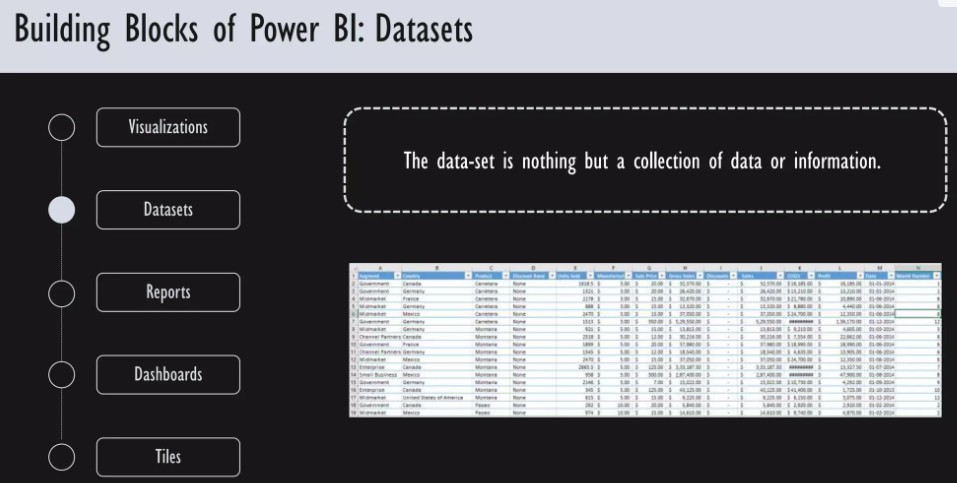
****

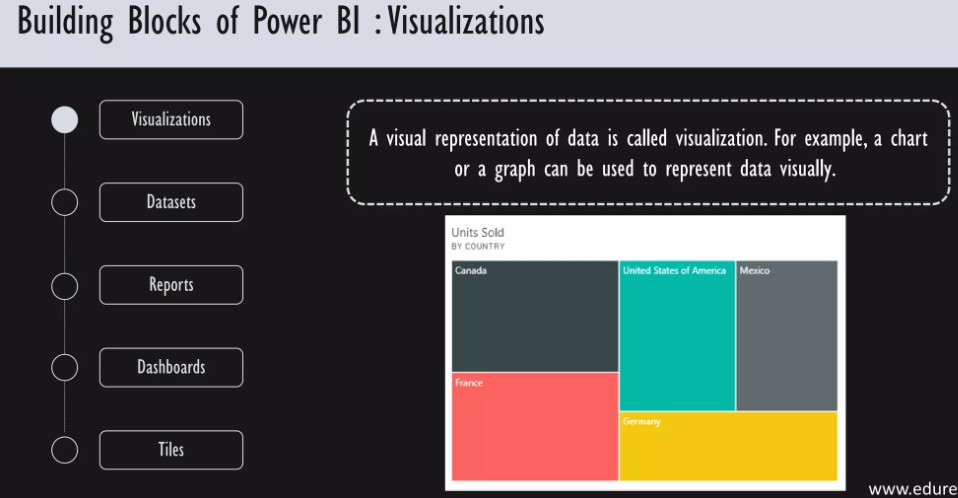
**Building Blocks of Power BI**

****

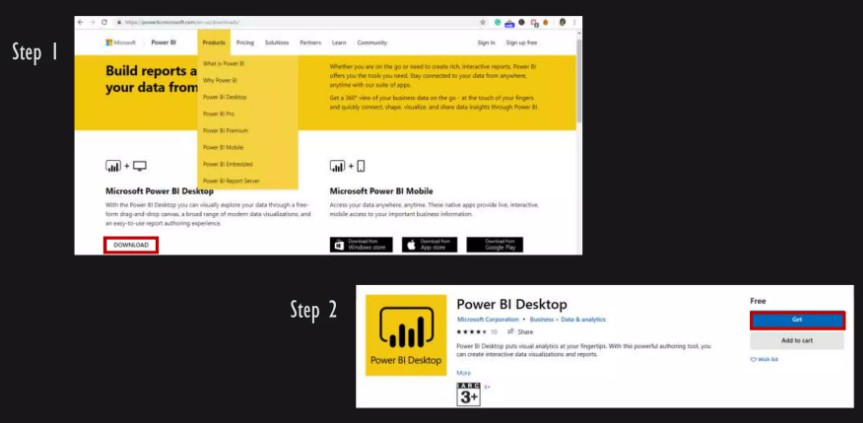


****

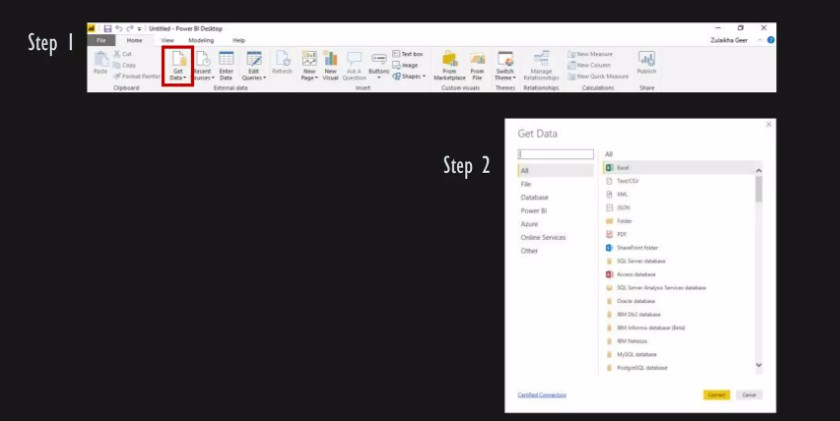


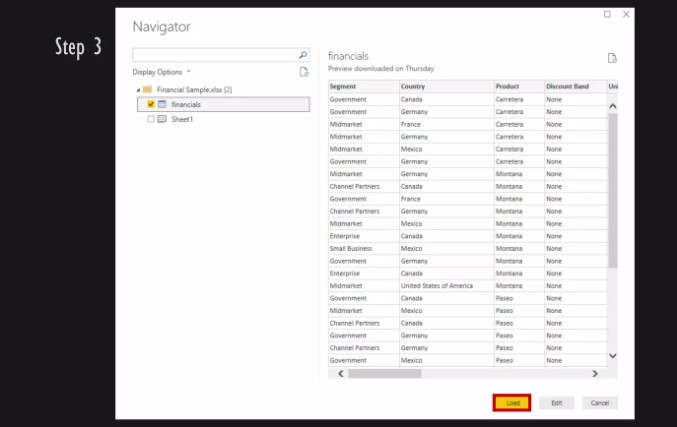
****

* 1. **- Installation of Power BI**

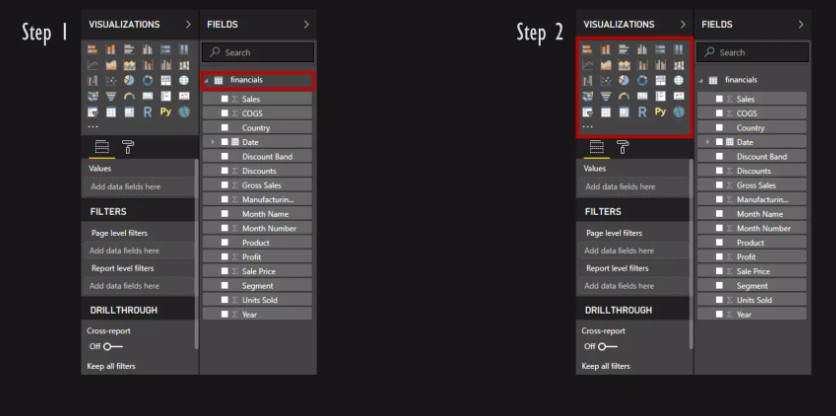
****

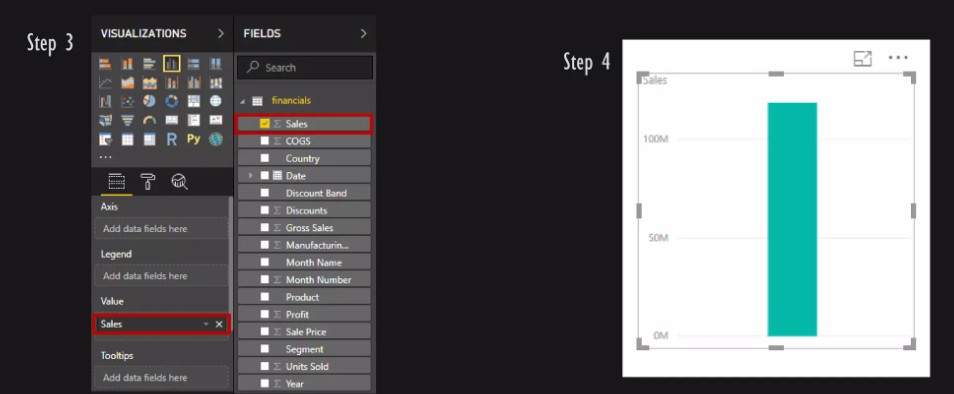
* 1. **- Importing Data in Power BI**

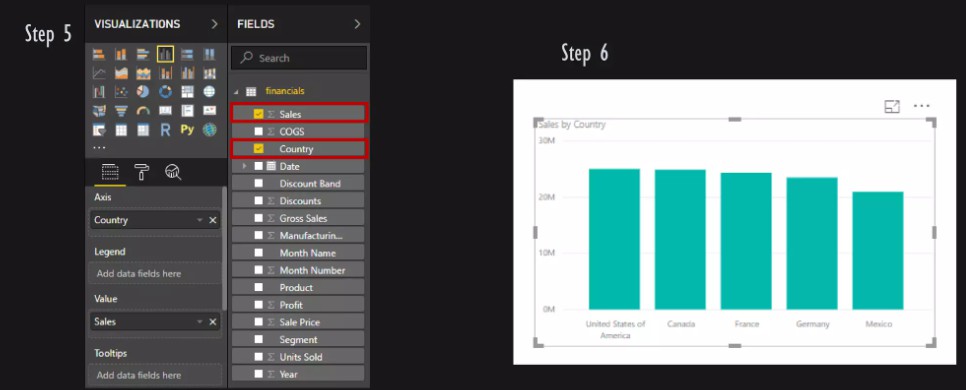
****



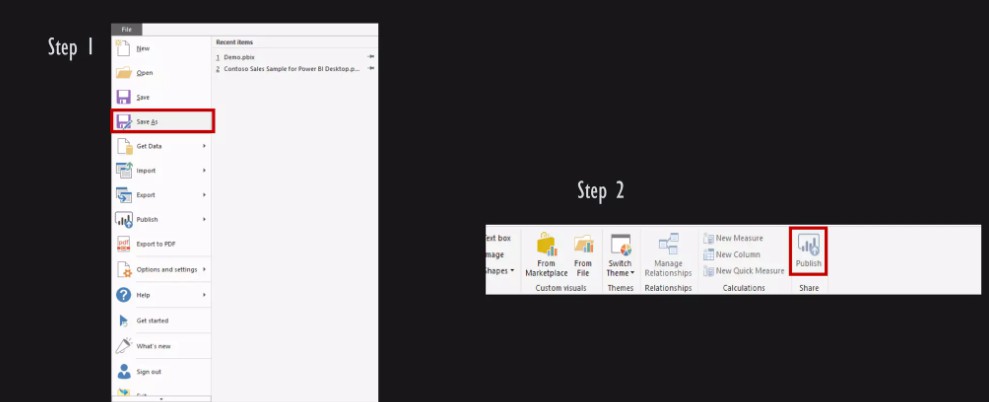
* 1. **- Working with Power BI**

****

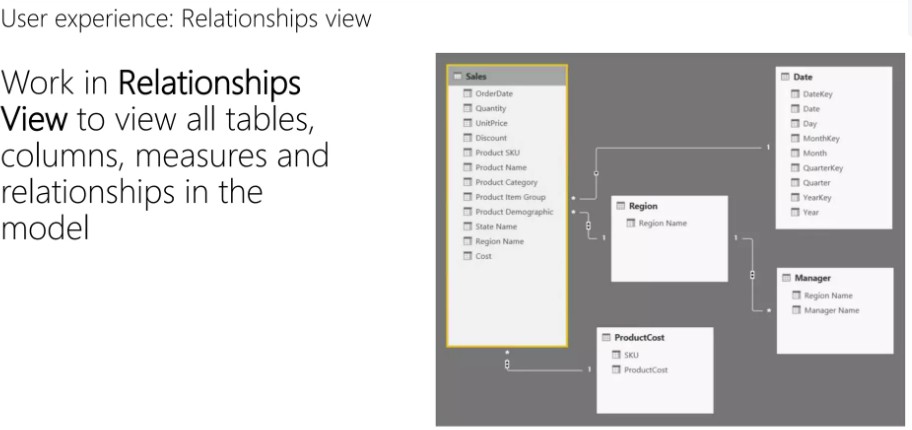
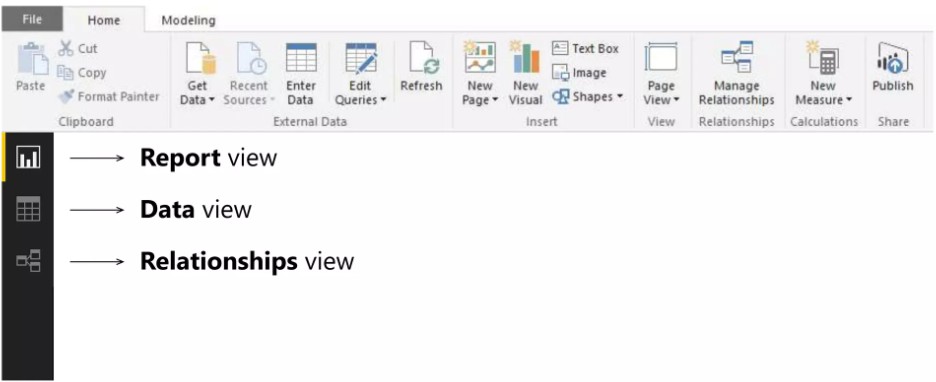


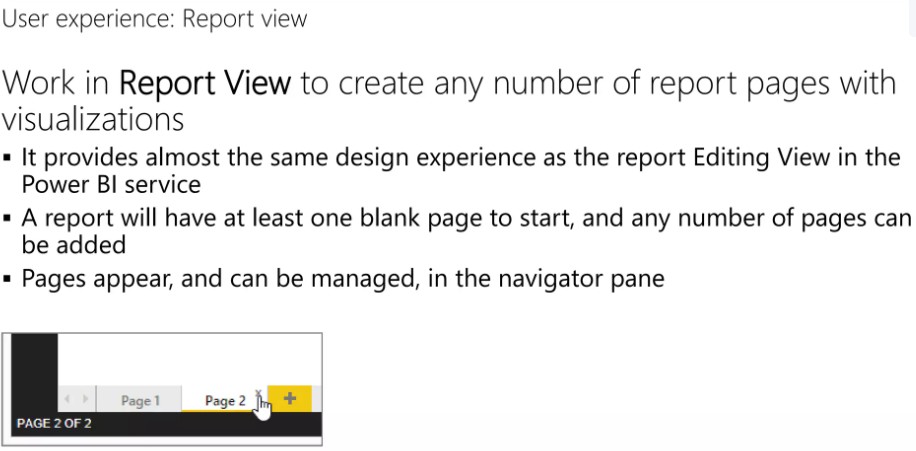
****

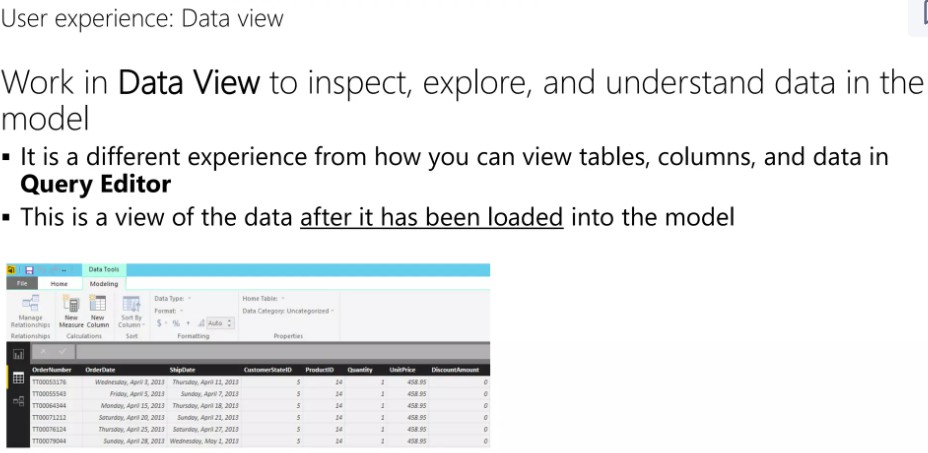
* 1. **- Save & Publish**

****

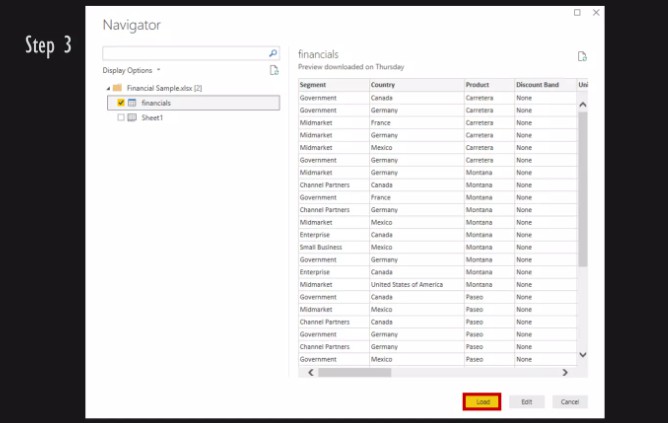
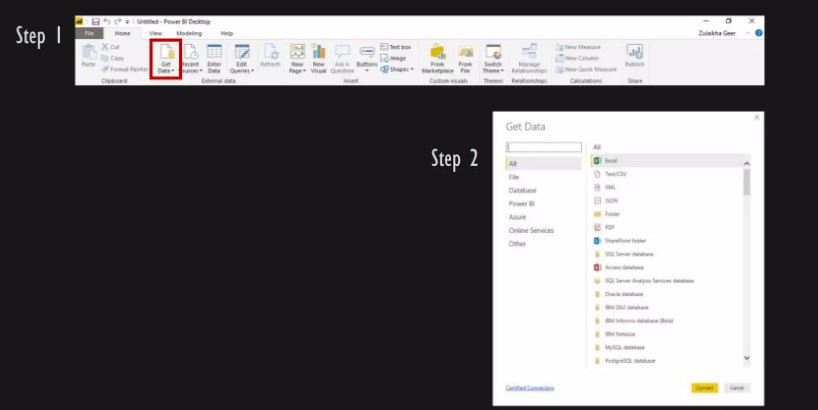
**Views of Power BI**

****

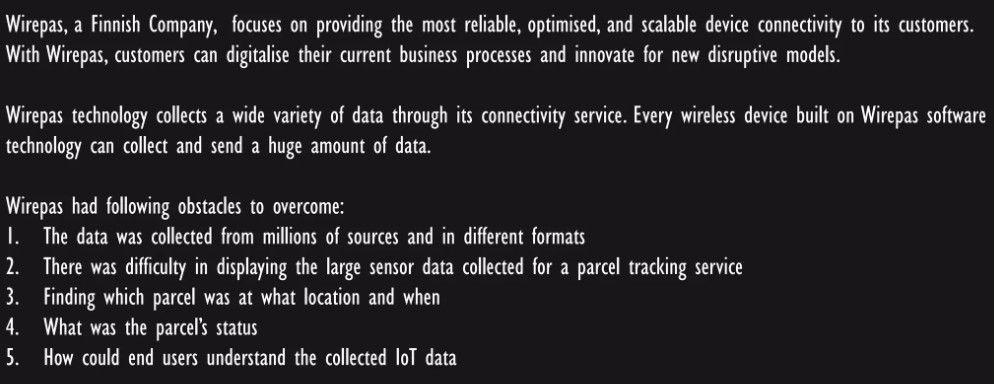


****

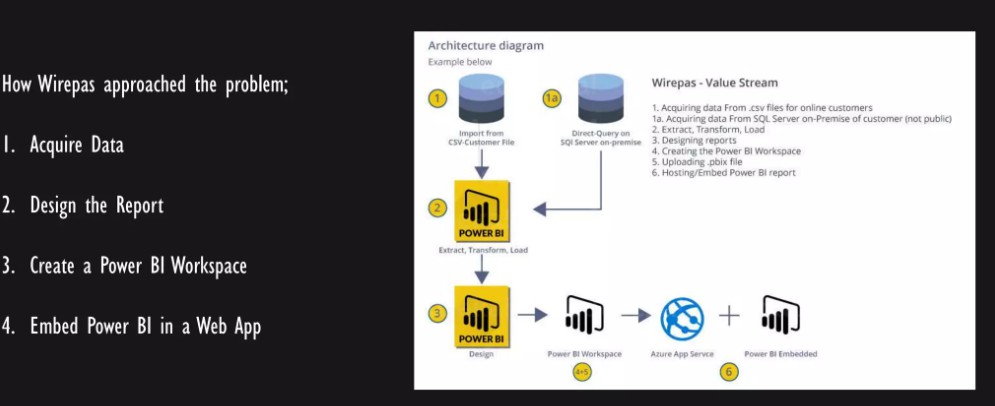
**2.0 Adding & Joining Another Data Source**

****

**Practice Question:**

****

**Solution:**

****



**Context**

Doraemon, the robotic cat from the future, has one true weakness: **Dora Cakes** 🍰.

Every month, he spends a large portion of his allowance (and sometimes even Nobita’s pocket money) to satisfy his craving.

However, as Doraemon’s appetite keeps increasing, Nobita’s mother has started complaining about **rising expenses**. To keep things under control, Doraemon decides to maintain a **monthly expense tracker** of all Dora Cakes he eats, how much they cost, and how much Nobita contributes.

Now Doraemon has asked *you* to act as a **Business Intelligence (BI) Analyst** and build insights from his Dora Cake consumption. He wants to know:

1. **Consumption Trend** → How many Dora Cakes he eats month by month.
2. **Expense Trend** → How much money he spends overall, and whether Nobita’s contribution is enough.
3. **Seasonal Pattern** → Are there months when Doraemon eats more cakes? (e.g., birthdays, holidays).
4. **Cost Analysis** → Does the price per cake rising affect his spending?
5. **Decision-Making** → Should Doraemon reduce his cake consumption or ask Nobita for more money?

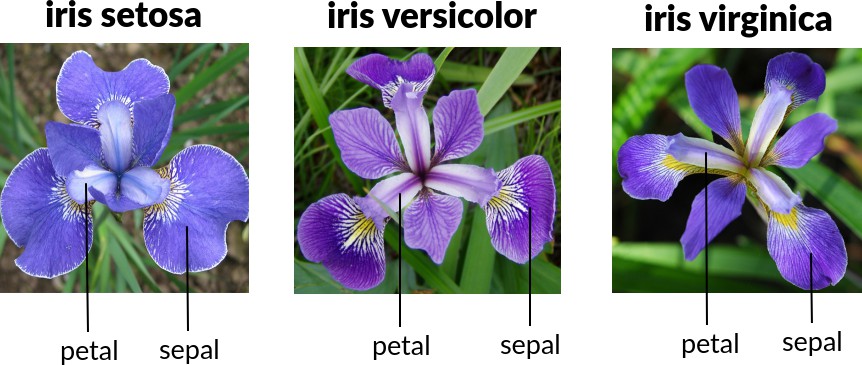
**Lab Tasks & Requirements**

**Question 1.** Using Doraemon’s Dora Cake Expense dataset:

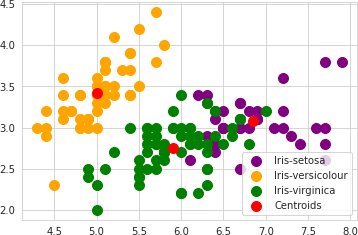
* + Import the data into **Excel**.
  + Create at least **3 different charts** (Column, Line, Pie, or Combo).
  + Highlight insights such as:
    - Which month Doraemon spent the most.
    - Which month Nobita contributed the highest share.
    - The overall cake-eating trend.
  + Finally, submit your Excel file with the charts and a **short paragraph (2–3 sentences)** explaining what Doraemon should do to manage his Dora Cake addiction better.

**Question 2:** You are working as a Data Analyst for a botanical research lab. The lab has collected flower measurements of Iris species (Setosa, Versicolor, Virginica). Your job is to analyze, visualize, and provide insights using Power BI dashboards.

1. Iris dataset columns:
   1. SepalLengthCm
   2. SepalWidthCm
   3. PetalLengthCm
   4. PetalWidthCm
   5. Species (Setosa, Versicolor, Virginica)



1. Data Loading & Preparation
   1. Import the Iris dataset (CSV/Excel) into Power BI.
   2. Verify datatypes (numerical vs categorical).
   3. Create a new calculated column:
   4. FlowerSize = SepalLengthCm + SepalWidthCm + PetalLengthCm + PetalWidthCm



1. Basic Visualizations
   1. Create a Bar Chart showing the count of flowers per Species.
   2. Create a Histogram (or column chart) of Sepal Length distribution.
   3. Build a Pie Chart for percentage distribution of each species..
2. Further Visualization & Decision-Making Questions (Analysis Part)
   1. Which species has the largest overall flower size?
   2. Can we differentiate species by looking at Petal Length vs Petal Width?
   3. Which feature (Sepal or Petal) is more discriminative between species?
   4. If you had to build a classification rule, which visual pattern would you rely on?
3. Submission
   1. Please insert the screenshots of Power BI to the same Excel and write your analysis in the same Excel

**Question 3:** Download the file “Student\_Performance.xlsx” (provided in the lab resources). The dataset contains the following columns: Student\_ID, Name, Department, Gender, Exam\_Score, Attendance (%), Semester.

1. Your task is to:
   * Open Power BI Desktop and connect to the Excel file.
   * Use the Power Query Editor to perform the following cleaning steps:
     + Remove duplicate rows.
     + Ensure that all numeric columns (Exam\_Score, Attendance) are in the correct data type.
     + Rename the column “Exam\_Score” to “Final\_Score”.
   * Load the cleaned dataset into Power BI.
2. Create the following visualizations:
   * A bar chart showing Average Final Score by Department.
   * A pie chart showing the distribution of students by Gender.
   * A line chart showing Average Attendance by Semester.
3. Add a slicer that allows filtering the dashboard by Department.

**Question 4:** In this question, you will learn how to create a database in PostgreSQL, design a table schema, insert sample data, and then connect it to Power BI for visualization and analytics.

1. Your task is to:
   * Create a Database
     + Open PostgreSQL and create a new database named `salesdb`.
   * Design Table Schema
     + Inside the `salesdb` database, create a table named `sales\_data` with the following columns:



* + Insert Data
    - Insert at least 10 rows of data into the table using SQL INSERT commands.
  + Connect PostgreSQL to Power BI
    - Open Power BI Desktop.
    - Go to Get Data → Database → PostgreSQL database.
    - Enter your server name, database name (salesdb), and login credentials.
    - Load the sales\_data table into Power BI.

1. Perform Analytics in Power BI

* Create a visualization showing Total Sales Revenue by Region.
* (Hint: Revenue = Quantity × Unit Price)
* Create a line chart showing Sales Trends over Time using sale\_date.
* Create a bar chart comparing Total Quantity Sold by Category.

**Submission Guidelines**

1. submit following the only file strictly following the naming convention
   1. l231234.xlsx

**Good Luck!**