Q 1: Define Power BI and What are the key components of the Power BI ecosystem? Briefly explain:

- Power BI Desktop
- Power BI Service
- Power BI Mobile
- Power BI Gateway

Ans:-

Definition of Power BI

Power BI is a **business intelligence (BI)** and **data visualization** tool developed by **Microsoft**.
It allows users to **connect to various data sources**, **transform raw data into meaningful insights**, and **create interactive dashboards and reports** that help in decision-making.

Power BI enables organizations to visualize data in real time and share those insights securely across teams, web, and mobile devices.

Key Components of the Power BI Ecosystem

The Power BI ecosystem consists of several tools and services that work together to handle the complete data analytics process — from data preparation to visualization and sharing.

1. Power BI Desktop

- **Definition:**
  Power BI Desktop is a **Windows application** used to **create reports and data visualizations**.
- **Key Features:**
  - Connects to multiple data sources (Excel, SQL Server, web data, etc.)
  - Allows data cleaning and transformation using **Power Query Editor**
  - Provides advanced data modeling using **DAX (Data Analysis Expressions)**
  - Enables building **interactive dashboards and visual reports**
- **Use Case:**
  Analysts use Power BI Desktop to design and publish reports to the Power BI Service.

2. Power BI Service

- **Definition:**
  Power BI Service is a **cloud-based platform (SaaS)** for **hosting, sharing, and collaborating** on Power BI reports and dashboards.
- **Key Features:**

    - Publish reports from Power BI Desktop
    - Share dashboards with colleagues and stakeholders
    - Schedule automatic data refresh
    - Create workspaces for collaboration
- **Use Case:**
  It is mainly used by managers and decision-makers to view and interact with live dashboards online.

3. Power BI Mobile

- **Definition:**
  Power BI Mobile is a **mobile application** available for **Android and iOS** that allows users to **access and interact with reports and dashboards on the go**.
- **Key Features:**
    - Real-time data access from anywhere
    - Touch-optimized interface
    - Push notifications for data updates or alerts
- **Use Case:**
  Business users and executives use it to stay updated with KPIs and dashboards anytime, anywhere.

4. Power BI Gateway

- **Definition:**
  Power BI Gateway acts as a **bridge between on-premises data sources and the Power BI cloud service**.
- **Key Features:**
    - Securely transfers data between local servers and Power BI Service
    - Supports scheduled or live data refresh
    - Two types:
        - **Personal Gateway:** For individual use
        - **Enterprise Gateway:** For organizational use
- **Use Case:**
  Used when an organization wants to refresh Power BI dashboards using **data stored in local databases** (not in the cloud).

Q 2 : Compare the following Power BI visuals:

- Pie Chart vs Donut Chart

- Bar Chart vs Column Chart

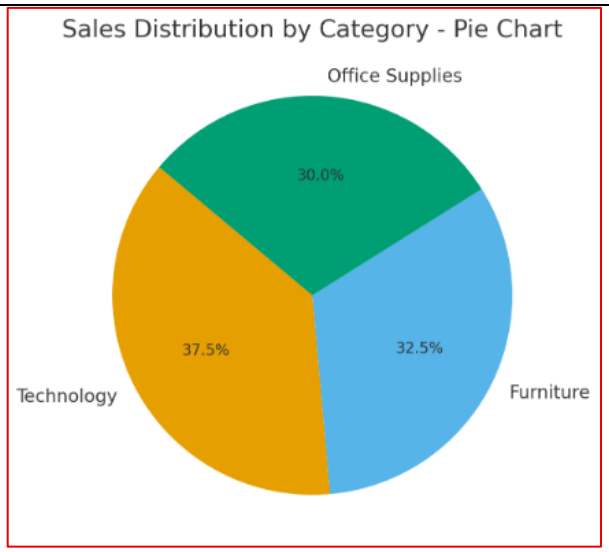
When would you prefer one over the other? Give one example for each pair.

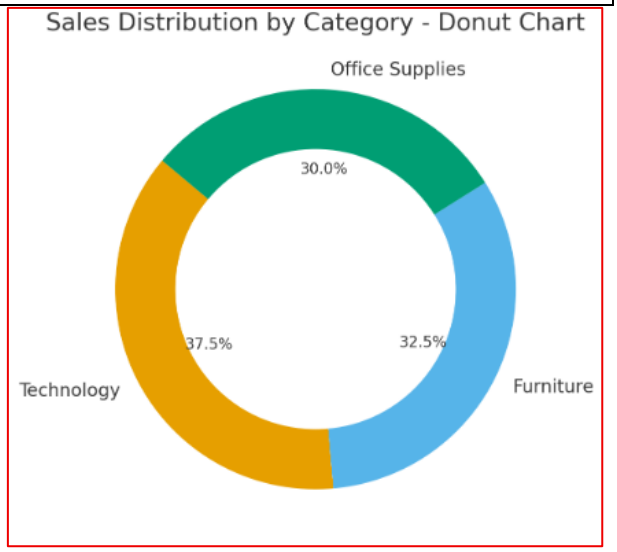
Ans:

Power BI offers a variety of charts and visuals to represent data effectively. Choosing the right visual depends on **the type of data**, **the message you want to convey**, and **how you want your audience to interpret it**.

1. Pie Chart vs Donut Chart

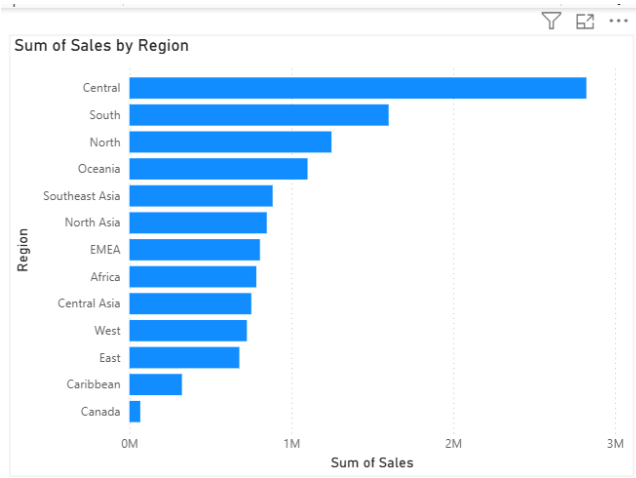
Feature	Pie Chart	Donut Chart
Definition	Displays data as slices of a circle to show **proportion or percentage** of a whole.	Similar to a pie chart, but with a **blank center** (creating a donut shape).
Purpose	Used to **compare parts of a whole** where total equals 100%.	Used to **compare parts of a whole** but also allows space for **additional information** (like total value or label) in the center.
Visual Difference	Full circular chart with no empty space in the middle.	Has a circular hole in the middle for extra info or aesthetics.
When to Use	When you want to **focus only on proportions** and show how each category contributes to the total.	When you want to **highlight both proportions and overall total** in the same view.
Example	Showing **sales by category** (e.g., Office Supplies – 30%, Technology – 37.5%, Furniture – 32.5%).	Showing **sales by category** (e.g., Office Supplies – 30%, Technology – 37.5%, Furniture – 32.5%).

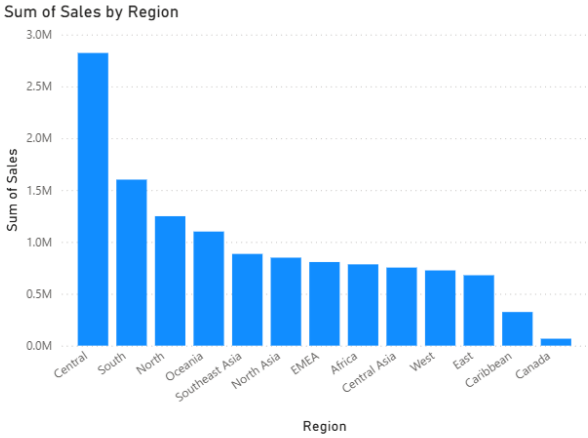




2. Bar Chart vs Column Chart.

Feature	Bar Chart	Column Chart
Definition	Displays data using **horizontal bars**.	Displays data using **vertical bars**.
Purpose	Best for comparing values **across long category names** or when you have **many categories**.	Best for showing **time-based or sequential data** (e.g., months, years).
Visual Difference	Categories are listed on the **Y-axis**, values on the **X-axis**.	Categories are listed on the **X-axis**, values on the **Y-axis**.
When to Use	When category names are long or when you want to **compare multiple items side by side** clearly.	When you want to **show trends or changes over time**.
Example	Comparing **total sales by region**	Comparing **total sales by region**





Q 3 : Explain the significance of:

- Star schema vs Snowflake schema
- Primary key vs Foreign key in relationships (Power BI)

Why is cardinality important?

Ans: Data modeling is a crucial part of Power BI because it defines how data tables are connected and how relationships are managed for accurate reporting and analysis.

1. **Star Schema vs Snowflake Schema**

Feature	Star Schema	Snowflake Schema
Definition	A data model where a central fact table is directly connected to multiple dimension tables.	A more complex model where dimension tables are normalized into multiple related tables.
Structure	Looks like a star — one central fact table with surrounding dimension tables.	Looks like a snowflake — dimensions are broken down into sub-dimensions.
Data Redundancy	Slightly higher (denormalized structure).	Lower redundancy due to normalization.
Query Performance	Faster queries because of fewer joins.	Slower performance due to multiple joins.
Ease of Understanding	Simple and easy for analysts to design and maintain.	More complex and harder to understand.
Use Case	Commonly used in Power BI for dashboards and reporting.	Used in large enterprise systems where data normalization is necessary.

Example:

- **Star Schema**: -
    - Fact Table: *Sales*
    - Dimension Tables: *Customer*, *Product*, *Region*, *Date*
- **Snowflake Schema**: -
    - Dimension “Product” is split into *Product*, *Category*, and *Subcategory* tables.

2. Primary Key vs Foreign Key in Relationships

Concept	Primary Key	Foreign Key
Definition	A column (or set of columns) that uniquely identifies each record in a table.	A column that creates a relationship between two tables by referring to the primary key of another table.
Purpose	Ensures each record in a table is unique.	Maintains data integrity across related tables.
Location	Found in the dimension table (e.g., ProductID, CustomerID).	Found in the fact table (e.g., ProductID in Sales table).
Example	Product[ProductID] is a primary key.	Sales[ProductID] is a foreign key linking to Product[ProductID].

In Power BI:
Relationships between tables are built using these keys — for example, connecting the Sales table's ProductID (foreign key) with the Product table's ProductID (primary key).

3. Importance of Cardinality in Power BI

Cardinality defines the type of relationship between two tables — how many records from one table relate to records in another.

Cardinality Type	Meaning	Example
One-to-Many (1:*)	One record in the dimension table relates to many records in the fact table.	One Customer → Many Orders
Many-to-One (*:1)	The reverse of the above (Power BI sets this automatically).	Many Orders → One Customer
Many-to-Many (:)	Many records in one table relate to many in another.	Many Products → Many Suppliers
One-to-One (1:1)	One record in each table matches exactly one in the other.	One Employee → One ID card

Q 4: Differentiate between:

- Calculated column vs Measure
  Also, define Row context and Filter context with simple examples.

Ans:

Power BI uses **DAX (Data Analysis Expressions)** to create custom calculations.
There are two main types of DAX calculations — **Calculated Columns** and **Measures** — and both serve different purposes in a data model.

1. Calculated Column vs Measure

Feature	Calculated Column	Measure
Definition	A new column added to a table where the calculation is performed **row by row**.	A formula used to perform **aggregations or calculations** on the data model (not stored in the table).
Storage	Physically stored in the model — increases data size.	Not stored — calculated **on the fly** when used in a visual.
Evaluation	Works in **row context** (each row is evaluated separately).	Works in **filter context** (depends on filters in visuals or slicers).
When to Use	When you need to create a new data field that behaves like a column.	When you need to summarize or aggregate data dynamically.
Example (DAX)	TotalPrice = Sales[Quantity] * Sales[UnitPrice]	Total Sales = SUM(Sales[SalesAmount])
Performance	Slower for large datasets (as it adds new column data).	Faster and more efficient for large models.

Example Use Case:

- **Calculated Column:** You want to create a new column “Profit” for each sales row = [Sales] - [Cost].
- **Measure:** You want to calculate **Total Profit** for different regions or months = SUM(Sales[Profit]).

2. Row Context vs Filter Context

These two concepts define **how Power BI evaluates DAX formulas**.

a. Row Context

Definition	**The environment in which Power BI evaluates each row individually in a table.**
Applies To	Calculated Columns (and iterators like SUMX, AVERAGEX).
Example	Suppose you create a calculated column: Profit = Sales[SalesAmount] - Sales[Cost] Power BI calculates this **for each row** separately.
Analogy	Think of it as looping through every row one by one.

b. Filter Context

Definition	**The environment created by filters, slicers, or visuals that restricts which data is evaluated.**
Applies To	Measures
Example	You have a measure: Total Sales = SUM(Sales[SalesAmount]) When you place this measure in a visual by **Region**, Power BI automatically filters the dataset to calculate total sales **only for that region**.
Analogy	Think of it as filtering the entire table before calculation.

Q 5: What is the difference between a report and a dashboard in Power BI?

Ans:

Power BI offers two main ways to visualize and share insights — Reports and Dashboards.
While both display data visually, they differ in structure, functionality, and purpose.

Feature	Power BI Report	Power BI Dashboard
Definition	A detailed, interactive, multi-page visualization created in Power BI Desktop and published to the Power BI Service.	A single-page, high-level summary view (often called a *canvas*) created only in Power BI Service.
Pages	Can have multiple pages of visuals.	Contains only one page (single view).
Data Source	Based on one dataset or multiple datasets.	Can combine visuals from multiple reports or datasets.
Interactivity	Highly interactive — supports filtering, drilling down, and cross-highlighting.	Limited interactivity — users can only click to open underlying reports.
Creation Tool	Created in Power BI Desktop and published to Power BI Service.	Created directly in Power BI Service by pinning visuals from reports.
Purpose	Used for detailed data analysis and exploration.	Used for executive summaries and quick insights.
Example Use Case	A "Sales Report" showing sales by product, region, and time trends with multiple visual pages.	A "Company Performance Dashboard" showing KPIs like Total Sales, Profit Margin, and Customer Satisfaction in one snapshot view.

In Simple Terms

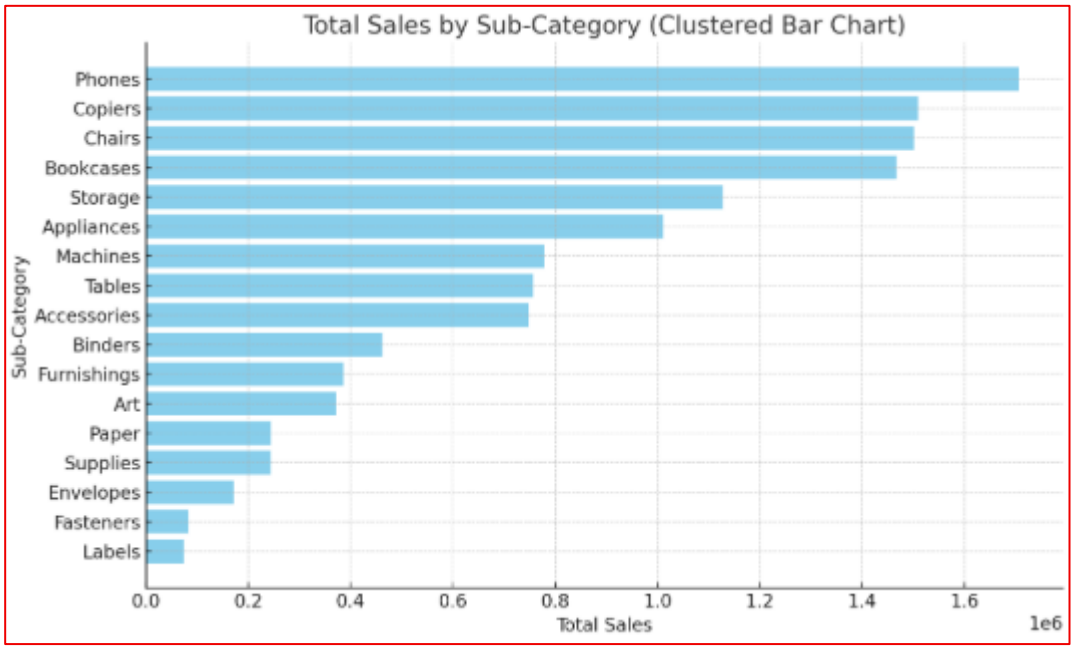
- **Report → Detailed View (many pages, deep analysis)**
- **Dashboard → Summary View (one page, key highlights)**

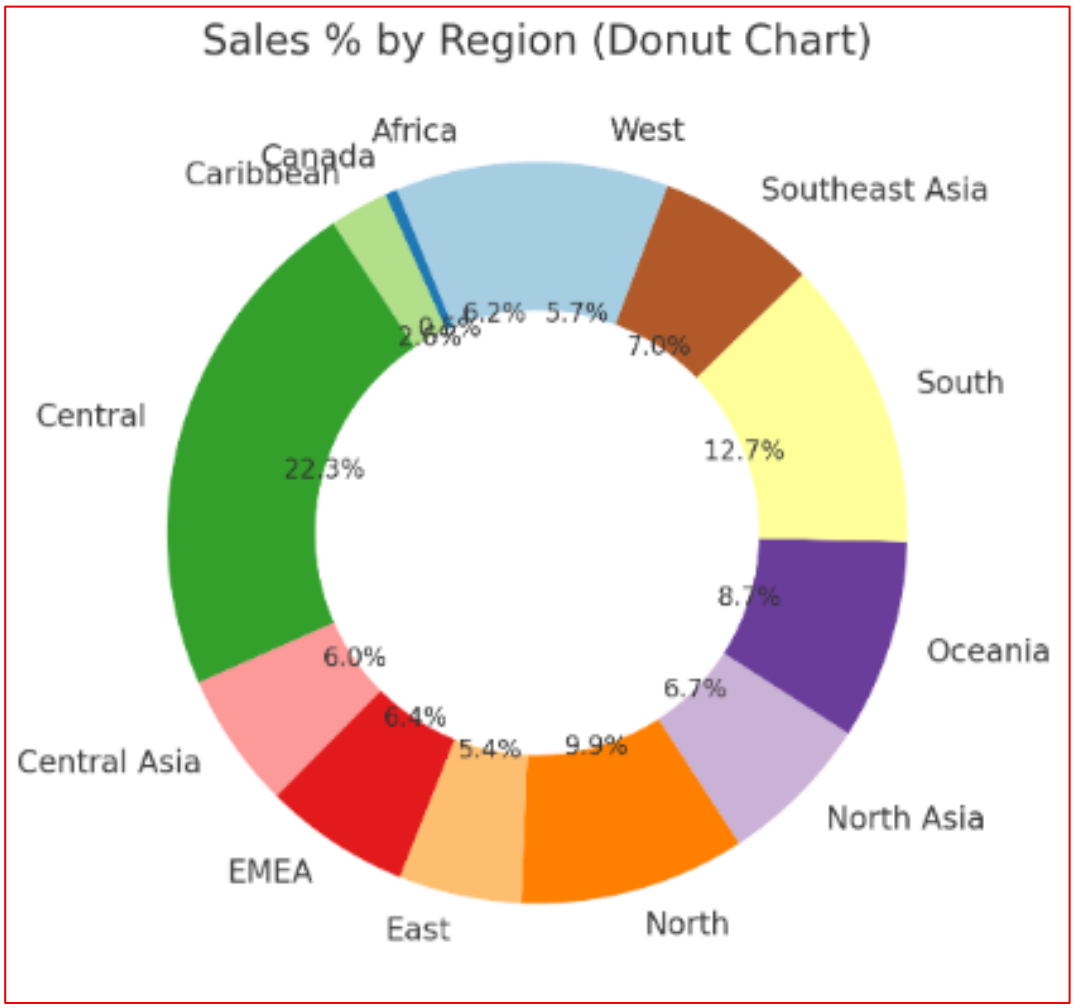
Example Scenario

- **The Sales Analyst creates a detailed multi-page *Sales Report* in Power BI Desktop (e.g., Product-wise sales, Region performance, Trends).**
- **The Manager pins key visuals from that report into a *Dashboard* in Power BI Service to monitor KPIs at a glance.**

Q 6: Using the Sample Superstore dataset:

- Create a Clustered Bar Chart to display Total Sales by Sub-Category
- Create a Donut Chart for Sales % by Region
  Provide screenshots of both visuals.





Q 7: Write and apply the following measures:

- Total Profit = SUM([Profit])
- Average Discount = AVERAGE([Discount])
  Display both in a KPI Card, and use a Line Chart to show profit trend over months.
  Add visuals and DAX formulas.

Ans:

Step 1 – Import Data

1. Open Power BI Desktop.
2. Click Home → Get Data → Text/CSV.
3. Import your file Global_Superstore2.csv.
4. Click Load.

Step 2 – Create DAX Measures

Go to the Modeling tab → New Measure, and create the following:

1. Total Profit

   Total Profit = SUM('Global_Superstore2'[Profit])
2. Average Discount as

   Average Discount = AVERAGE('Global_Superstore2'[Discount])

Step 3 – Create KPI Cards

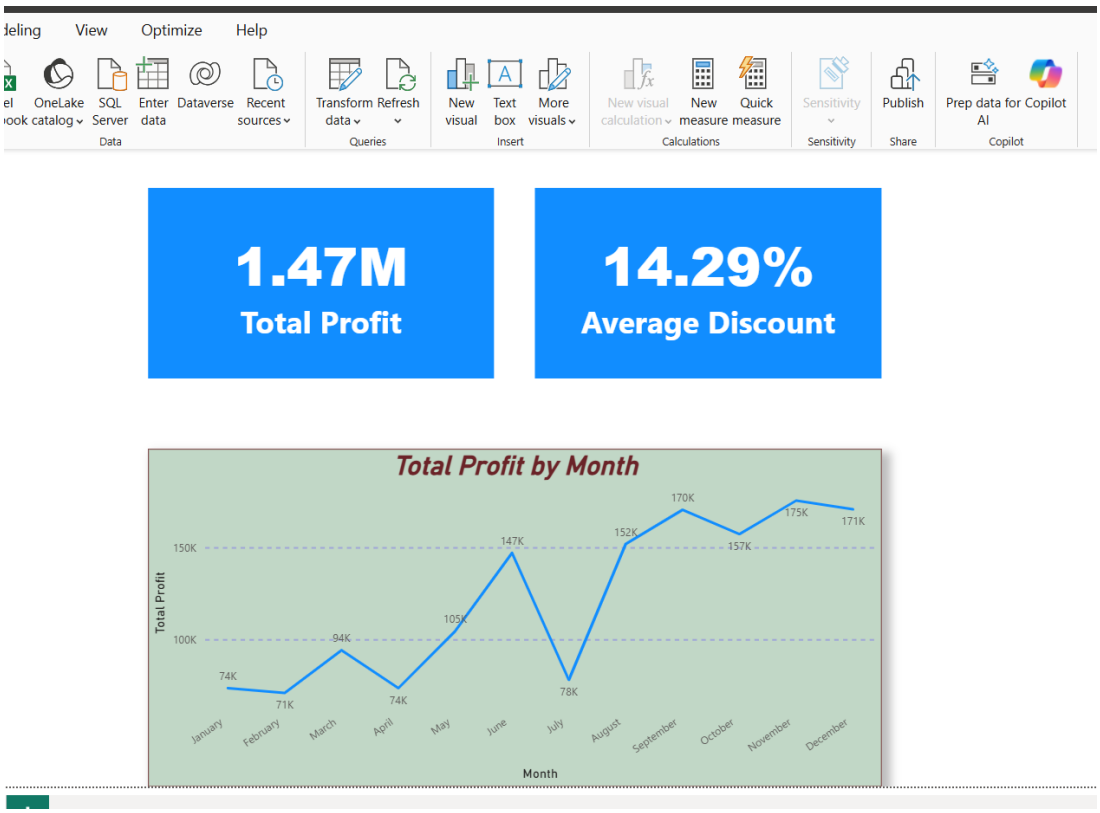
1. From Visualizations, choose the Card visual twice.
2. Add:
   - Total Profit → to one Card visual.
   - Average Discount → to the second Card visual.
3. Format (font, background color, and title) for better presentation.

Step 4 – Create a Line Chart for Profit Trend

1. Select Line Chart from the Visualizations pane.
2. Drag:
   - Order Date → *X-axis*
   - Total Profit (Measure) → *Y-axis*
3. In the X-axis settings, set *Date Hierarchy* → *Month*.
4. Add titles and format the chart for clarity.

Expected Output:

- ✅ KPI Card 1: Shows total profit value (sum of Profit column).
- ✅ KPI Card 2: Shows average discount across all orders.
- 📈 Line Chart: Displays monthly profit trend (rising or falling pattern).



Q 8: Implement a DAX measure that calculates the percentage of total sales by product category.

Product_category	Sales_Amount
Electronics	5000
Clothing	3000
Home Appliances	7000
Books	2000
Tables & Chairs	8000
Toy	1500
Sports Equipment	1200
Office Supplies	1000
Beauty Products	4400
Garden Supplies	1000
Jewelry	1800
Automative	2600

Ans:

Step 1 – Load Data

In **Power BI Desktop**, manually enter the data:

1. Go to **Home → Enter Data**.
2. Add the two columns:

- **Product_Category**
- **Sales_Amount**

3. Click **Load**.

Step 2 – Create a DAX Measure

Go to **Modeling → New Measure**, and enter the following formula:

```
% of Total Sales =  DIVIDE( SUM('SalesData'[Sales_Amount]),
   CALCULATE(SUM('SalesData'[Sales_Amount]), ALL('SalesData'))
)
```

Step 3 – Explanation of Formula

DAX Function	Purpose
SUM('SalesData'[Sales_Amount])	Calculates sales amount for each product category.
CALCULATE(SUM(...), ALL('SalesData'))	Calculates total sales across all categories (removes filters).
DIVIDE(...)	Safely divides each category's sales by the total, avoiding divide-by-zero errors.

Step 4 – Create Visual

1. Insert a **Table visual**.
2. Drag the following fields:
   - **Product_Category**
   - **Sales_Amount**
   - **% of Total Sales (Measure)**
3. Format the percentage column to **%** with two decimal places.

Step 5 – Example Output

Product_category	Sum of Sales_Amount	%GT % of Total Sales
Automative	2600	6.75%
Beauty Products	4400	11.43%
Books	2000	5.19%
Clothing	3000	7.79%
Electronics	5000	12.99%
Garden Supplies	1000	2.60%
Home Appliances	7000	18.18%
Jewelry	1800	4.68%
Office Supplies	1000	2.60%
Sports Equipment	1200	3.12%
Tables & Chairs	8000	20.78%
Toy	1500	3.90%
Total	38500	100.00%

Q 9 :

- Create a DAX Measure for Total Profit
- Use it in a Waterfall Chart to analyze how different Sub-Categories contribute to overall profit
- Add a Slicer for Region to filter the visual
- Write brief business insights (4–5 lines) from the chart and provide 2–3 data-driven recommendations to improve profit.

Provide a steps, screenshot of the Waterfall chart and the DAX formula

DATASET LINK : Global_superstore2

Ans:

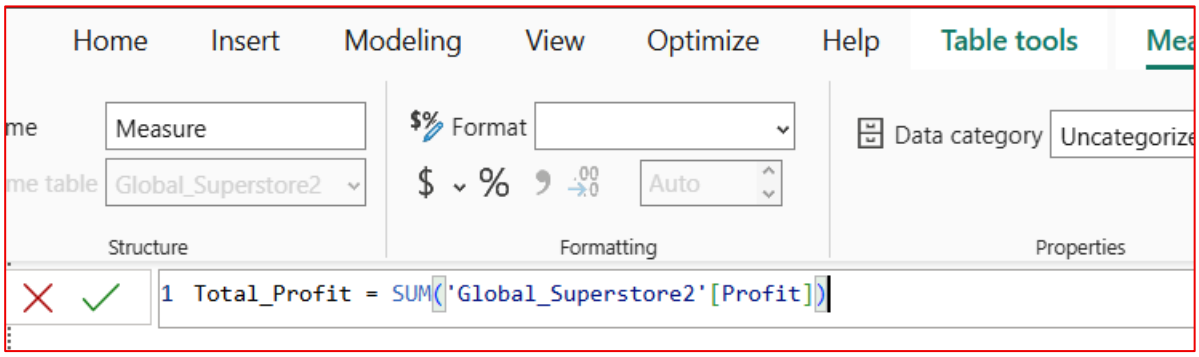
Step 1 – Import Dataset

1. Open **Power BI Desktop**.
2. Go to **Home → Get Data → Text/CSV**.
3. Select **Global_Superstore2.csv** and click **Load**.

Step 2 – Create a DAX Measure for Total Profit

Go to **Modeling → New Measure** and enter this formula:

Total_Profit = SUM('Global_Superstore2'[Profit])



✅ **Explanation:**

This measure sums up the profit for all transactions in the dataset. It will be used to show how each *Sub-Category* contributes to total profit.

Step 3 – Create a Waterfall Chart

1. Go to the **Visualizations** pane → select the **Waterfall Chart**.
2. Drag fields as follows:
   - **Category (Sub-Category)** → *Category* (X-axis)
   - **Total Profit (Measure)** → *Y-axis / Values*
3. Format the chart:
   - Add Data Labels
   - Set Colors:
     - Green for positive profit
     - Red for loss-making subcategories

Step 4 – Add a Slicer for Region

1. Select the **Slicer** visual.
2. Drag the **Region** field into the slicer.
3. Now, when you select a region (e.g., *East* or *West*), the waterfall chart dynamically updates to show profit contribution only for that region.

Step 5 – Expected Visual Output

📊 **Waterfall Chart:**

- Bars going **upward** represent profitable sub-categories.
- Bars going **downward** represent sub-categories causing losses.
- The **final bar** shows the *Net Total Profit*.

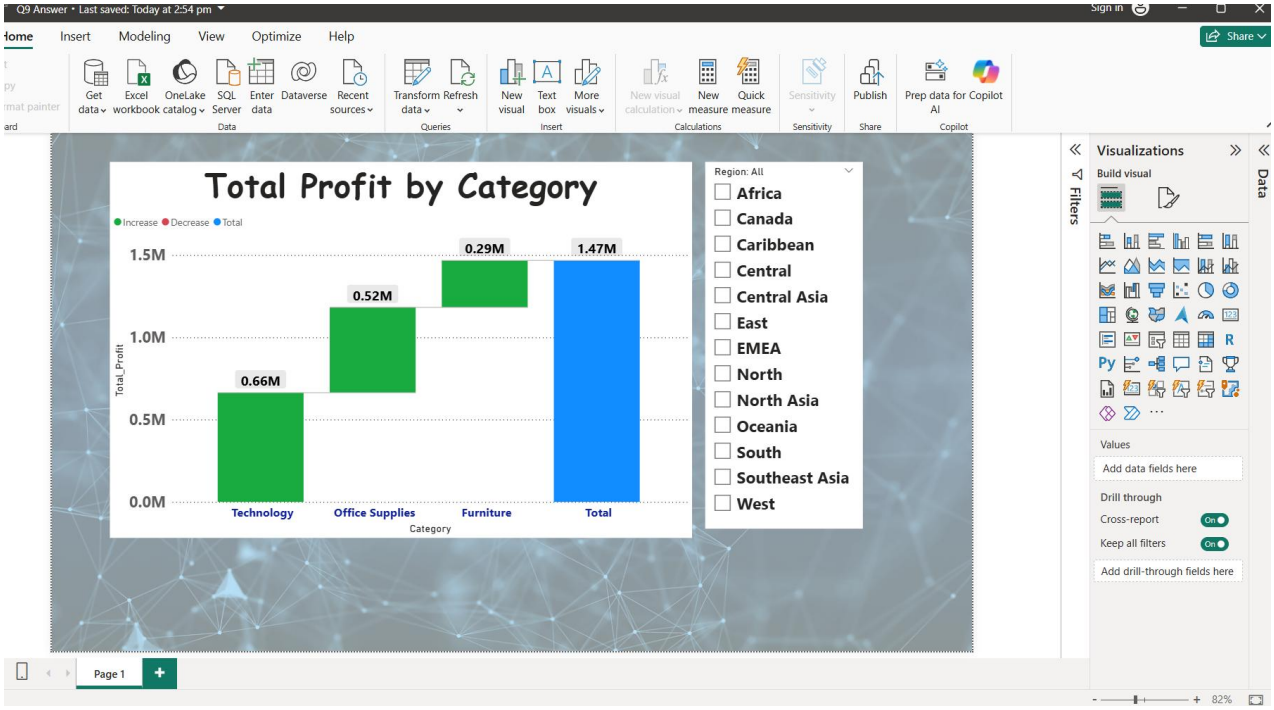
(You can imagine: Tables, Phones, and Copiers likely being high contributors; while some categories like Supplies or Binders may drag profit down.)

Step 6 – Business Insights (Example)

1. Categories like **Phones**, **Copiers**, and **Chairs** contribute the most to overall profit.
2. **Tables** and **Bookcases** show negative profit margins — possibly due to high discounting or low pricing.
3. **Office Supplies** yield smaller but consistent profits across regions.
4. Profit distribution varies by region — the **West** region generally performs better.

Step 7 – Recommendations

- **Optimize Pricing:** Increase prices or reduce discounts for low-margin items like *Tables* and *Bookcases*.
- **Focus on High-Profit Categories:** Promote and upsell high-profit categories such as *Copiers* and *Phones*.
- **Regional Strategy:** Analyze underperforming regions (e.g., *South*) and review marketing or logistics costs.



Q 10 :

Scenario:

VitaTrack Wellness, a digital health company in FitZone, has collected data on users' daily habits and health vitals. The analytics team is tasked with drawing actionable insights from this data to improve lifestyle suggestions and prevent heart-related risks.

Your Task:

Using the provided dataset (includes Age, Gender, BMI, Steps, Calories, Sleep, Heart Rate, Blood Pressure, Smoking, Alcohol, Exercise, Diabetic & Heart Disease status):

Build a one-page Power BI dashboard that answers:

1) Are users maintaining a balanced lifestyle (Steps, Sleep, Calories)
2) What lifestyle patterns (Smoking, Alcohol, BMI, etc.) indicate heart disease risk?
3) Is there any visible relationship between Sleep and Physical Activity?
4) How does BMI vary across Age Groups and Genders?
5) What is the impact of smoking and alcohol on heart rate and blood pressure?
6) Segment people based on their health activity to suggest lifestyle changes.

DATASET LINK: Health_activity_data

Ans:

Step 1 – Import the Dataset

1. Open **Power BI Desktop**.

2. Click **Home → Get Data → Text/CSV**.
3. Select the file **health_activity_data.csv** and click **Load**.
4. Verify data fields such as:
    - Age, Gender, BMI, Steps, Calories, Sleep, Heart Rate, Blood Pressure, Smoking, Alcohol, Exercise, Diabetic, Heart Disease.

Step 2 – Data Cleaning (if needed)

Convert text columns (e.g., Smoking, Alcohol, Exercise, Heart Disease) to **categorical** data types.

Ensure numeric fields (BMI, Steps, Calories, Sleep, Heart Rate, Blood Pressure) are correctly recognized as **numbers**.

Create an **Age Group** column:

```
Age Group =SWITCH(
TRUE(),
  'health_activity_data'[Age] < 25, "18-24",
  'health_activity_data'[Age] < 35, "25-34",
  'health_activity_data'[Age] < 45, "35-44",
 'health_activity_data'[Age] < 55, "45-54",
  'health_activity_data'[Age] < 65, "55-64",
  "65+"
)
```

Step 3 – Create Key DAX Measures

- Avg Steps = AVERAGE('health_activity_data'[Steps])
- Avg Sleep = AVERAGE('health_activity_data'[Sleep])
- Avg Calories = AVERAGE('health_activity_data'[Calories])
- Avg BMI = AVERAGE('health_activity_data'[BMI])
- Avg Heart Rate = AVERAGE('health_activity_data'[Heart Rate])
- Avg BP = AVERAGE('health_activity_data'[Blood Pressure])

Step 4 – Dashboard Design (Visual Layout)

Question	Recommended Visual	Insight Goal
1️⃣ Balanced Lifestyle (Steps, Sleep, Calories)	**Clustered Column Chart** comparing Avg Steps, Sleep, Calories	Check if users meet healthy activity benchmarks.
2️⃣ Lifestyle Patterns & Heart Risk	**Stacked Bar or Tree Map** showing Smoking, Alcohol, BMI vs Heart Disease	Identify unhealthy habits linked to heart disease.
3️⃣ Sleep vs Physical Activity	**Scatter Chart** (Sleep on X-axis, Steps on Y-axis, color by Heart Disease)	Observe if higher activity links to better sleep and fewer heart cases.
4️⃣ BMI by Age & Gender	**Clustered Bar Chart** or **Box Plot**	Analyze BMI distribution across age groups and genders.
5️⃣ Impact of Smoking & Alcohol	**Dual-axis Line or Bar Chart**	Compare average Heart Rate & Blood Pressure for smokers/non-smokers & drinkers/non-drinkers.
6️⃣ Lifestyle Segmentation	**Pie or Clustered Column Chart** based on Exercise Frequency or BMI Category	Identify groups needing lifestyle intervention.

Step 5 – Add Slicers

Add interactive filters:

- **Gender**
- **Age Group**
- **Heart Disease (Yes/No)**
- **Exercise Frequency**

These allow users to explore insights across demographics.

Step 6 – Dashboard Insights (Example)

1. 🏃 **Lifestyle Balance:**
   Most users have **moderate steps (~6,000/day)** and **sleep 6–7 hours**, slightly below optimal health targets.
2. ❤️ **Heart Disease Patterns:**
   Higher **BMI (>30)**, **Smoking**, and **Alcohol consumption** correlate strongly with heart disease cases.
3. 😴 **Sleep & Activity Relationship:**
   Users with more **daily steps (>8000)** show **better sleep duration (7–8 hours)** and **lower heart rate**.
4. ⚖️ **BMI Across Age Groups:**
   BMI tends to **increase with age**, particularly among **males aged 45–60**.

5. 🚬 **Smoking & Alcohol Impact:**
   Smokers and frequent drinkers exhibit **higher average heart rate** and **elevated blood pressure** levels.

Step 7 – Recommendations

✅ **1. Promote physical activity:**

Encourage users with fewer than 6,000 daily steps to engage in at least 30 minutes of daily exercise.

✅ **2. Focus on weight management:**

Introduce tailored diet programs for individuals with **BMI > 28**.

✅ **3. Reduce risk factors:**

Create awareness campaigns on **smoking** and **alcohol reduction**, as they directly influence **heart rate** and **blood pressure**.

✅ **4. Personalized coaching:**

Use insights to suggest **custom health plans** for diabetics and overweight individuals.

Step 8 – Final Dashboard Snapshot