7. Creating Reports & Visualizations-Different types of charts, Formatting charts with Title, Colors.

Experiments to be executed.

1. Number of customers by each city using Pie Chart.

2. Number of customers by each state using Donut Chart.

3. Number of customers by each city and customer segment using Clustered Column Chart.

4. Distribution of Sales over time using Line Chart.

5. Revenue by Payment Method and City Visualization using Horizontal Clustered Bar Chart

6. Revenue by Store Visualization using Donut Chart.

7. Map Visual for Revenue Distribution using Map.

Data visualization is the graphical representation of data using visual elements like charts, graphs, maps, and plots. Its main purpose is to make complex data easier to understand, analyze, and interpret by displaying it visually, allowing patterns, trends, and insights to be quickly identified.

Tools: Common tools for creating data visualizations include Excel, Power BI, Tableau, and programming libraries like Matplotlib and Seaborn in Python.

<u>**1. Number of customers by each city**</u>
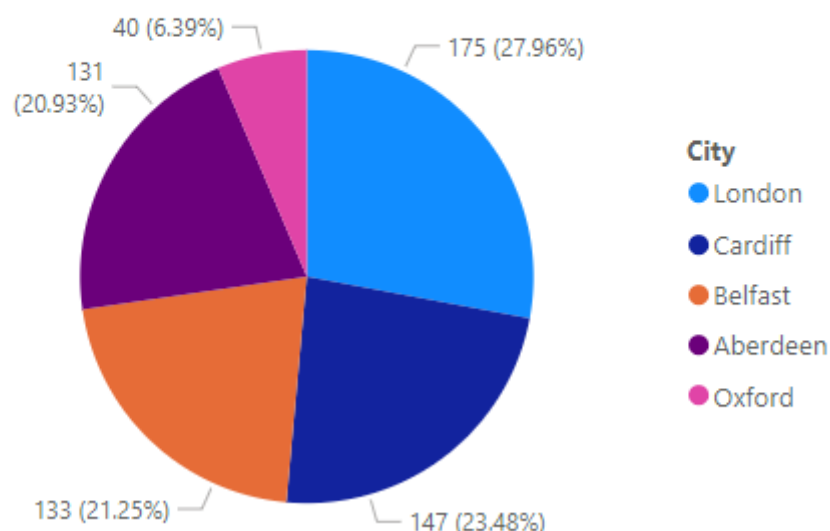
- Here we want to see the proportion of customers from a particular city compared to the total number of customers.
- To see how much of a something compared to the whole, use a pie chart or a donut chart,

Get the Pie chart to the canvas

- A pie chart is a circular statistical graphic that is divided into slices to illustrate numerical proportions. Each slice represents a category, and the size of each slice is proportional to the percentage or fraction of the whole it represents. It is commonly used to visualize data where the total can be split into discrete parts, and you want to show how each part contributes to the whole.

Fields of Pie Chart and Donut chart are same. That is Legend, Values

Pie and doughnut charts allow users to visualize parts of a whole. The key fields for these charts include Legend, Values, Details, and Tooltips, and they work similarly for both chart types.

To display the number of customers by city, the Customer ID is added to the Values field, and the City is added to the Legend field. Selecting "distinct count" shows the unique number of customers per city.

Customer ID in the Values Field: The *Values* field in a chart is where you place the data you want to measure or count. In this case, the *Customer ID* is added here because each customer has a unique ID. This allows Power BI to count how many customers there are.

- **Distinct Count**: By selecting "distinct count," Power BI ensures that only unique customers are counted. This is important because, without this option, Power BI might count duplicate customer IDs, which could result in an inaccurate total. Selecting "distinct count" ensures the number of unique customers per city is displayed correctly.
- City in the Legend Field: The *Legend* field is used to categorize the data. By adding the *City* field here, Power BI will group the customers based on the city they belong to. This means the chart will break down the total customer count by city.

1. <u>**Number of customers by each state**</u>

A **pie chart** is a full circle divided into slices representing proportions of a whole, while a **doughnut chart** has a hole in the center, creating a ring-like appearance. Pie charts are best for fewer categories, focusing on slice size to show proportions. Doughnut charts, ideal for 5 or more categories, offer a cleaner look and reduce clutter. Both charts display similar data but differ in visual design and space efficiency.
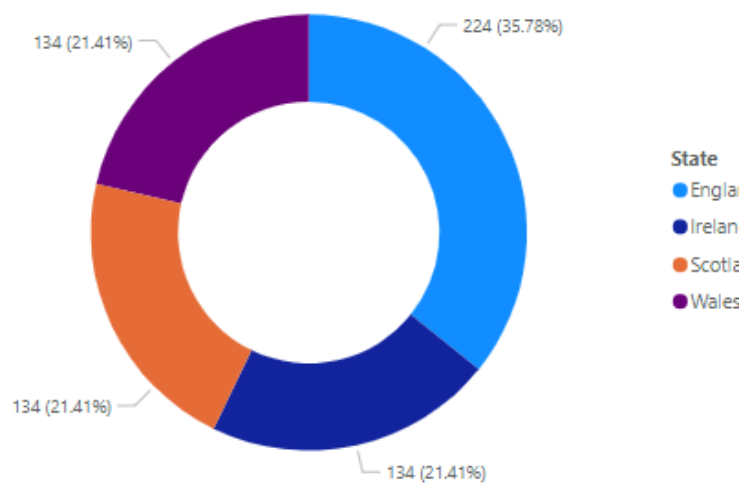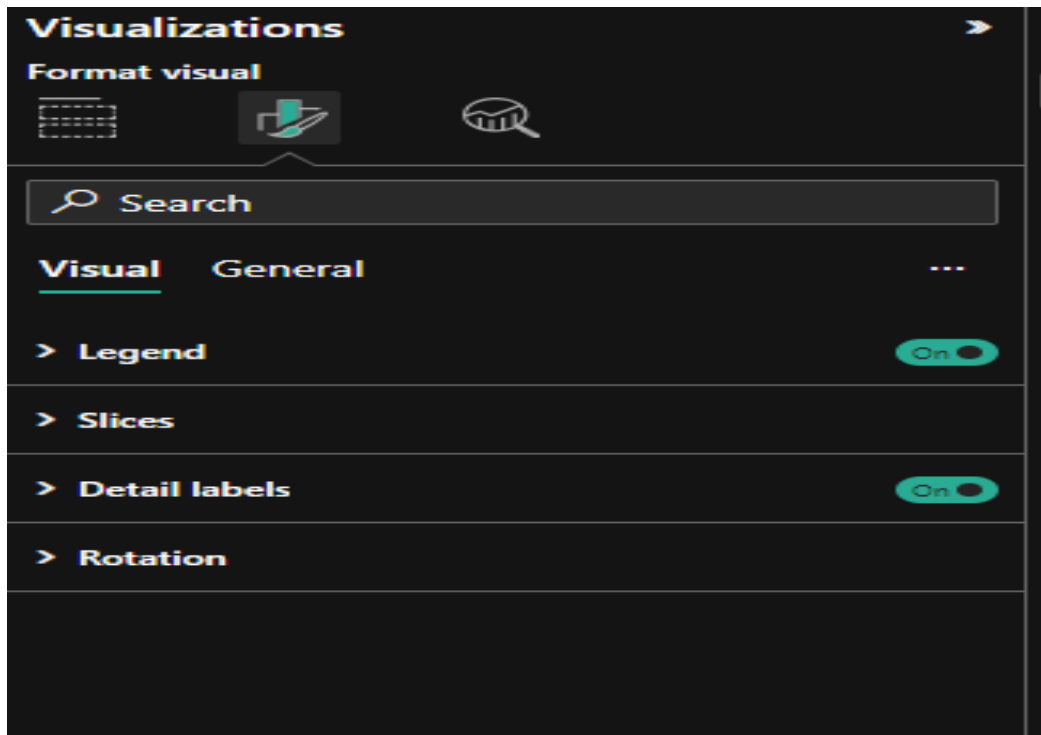
Count of Customer ID by State

Formatting PowerBI Visuals

Visualization-> Format Visual-> Visual(Chart element formatting) and General for Chart canvas formatting,=.

2. Number of customers by each city and customer segment

A **clustered column chart** is a type of bar chart where data values are represented by vertical bars (columns) grouped (clustered) together for easy comparison. Each cluster represents a category, and within each cluster, multiple columns show different series of data. This type of chart is useful for comparing multiple datasets across categories.
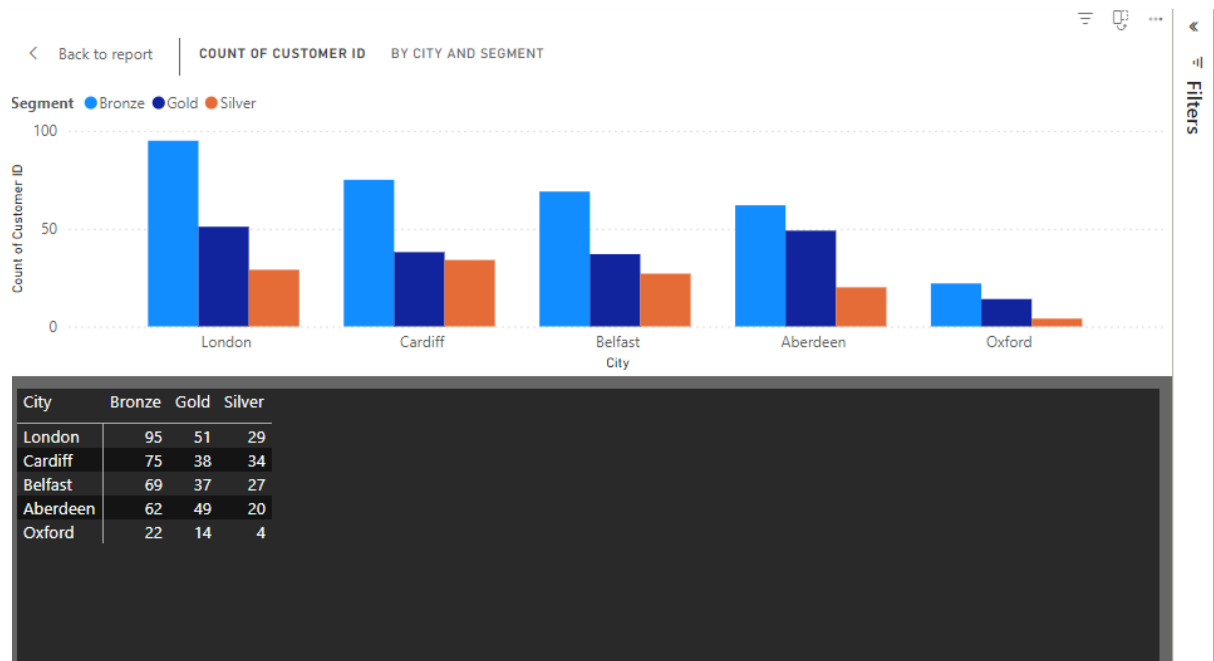
1. **Categories on the x-axis**: These are the groups you're comparing (e.g., years, departments, products).
2. **Columns within clusters**: Each category has multiple columns representing different data series (e.g., sales for different regions or products).
3. **Values on the y-axis**: The height of the columns indicates the value for each series within the category.

Y-Axis: Numerical Value: customer id.

X Axis: Data Category: Cities

Legend: Breaks X axis to further categories: Customer segment

X axis analysed further based on customer segment(This becomes the Legend)

Filters

Count of Customer ID by City and Segment

Segment ● Bronze ● Gold ● Silver

City	Bronze	Gold	Silver
London	95	51	29
Cardiff	75	38	34
Belfast	69	37	27
Aberdeen	62	49	20
Oxford	22	14	4

4. Distribution of Sales over time

Sales over Time Visualization

A **line chart** is a type of graph that displays data points connected by straight lines to show trends over time or relationships between two variables. It is commonly used to track changes or trends in data over intervals, such as time (e.g., days, months, years).

1. **Choosing the Right Chart**:

   - To show **sales over time**, a **line chart** is ideal as it displays changes over a period.
   - Select the **line chart** from the visualization pane.

   Setting the Y-Axis:

   - The y-axis represents the value to be analysed.
   - In this case, use the **Invoice ID** for the y-axis and choose **distinct count** of Invoice IDs to represent the number of transactions.

   Setting the X-Axis:

- The x-axis represents time.
- Use the **Order Date** field, which records the date of each sale, for the x-axis.

Simplifying the Chart:

- Initial charts may be hard to read due to too many data points (e.g., daily data).
- **Remove the day hierarchy** from the date field to simplify the chart, leaving only **year, quarter, and month**.

Insights from the Chart:

- The chart shows general **upward trends** in sales with **monthly variations**.
- Example: Sales transactions in **June 2013 (51)** vs. **July 2013 (17)**, and **December 2013 (61)** vs. **January 2014 (31)**.
- The supermarket manager could investigate reasons for these **declines in sales**.
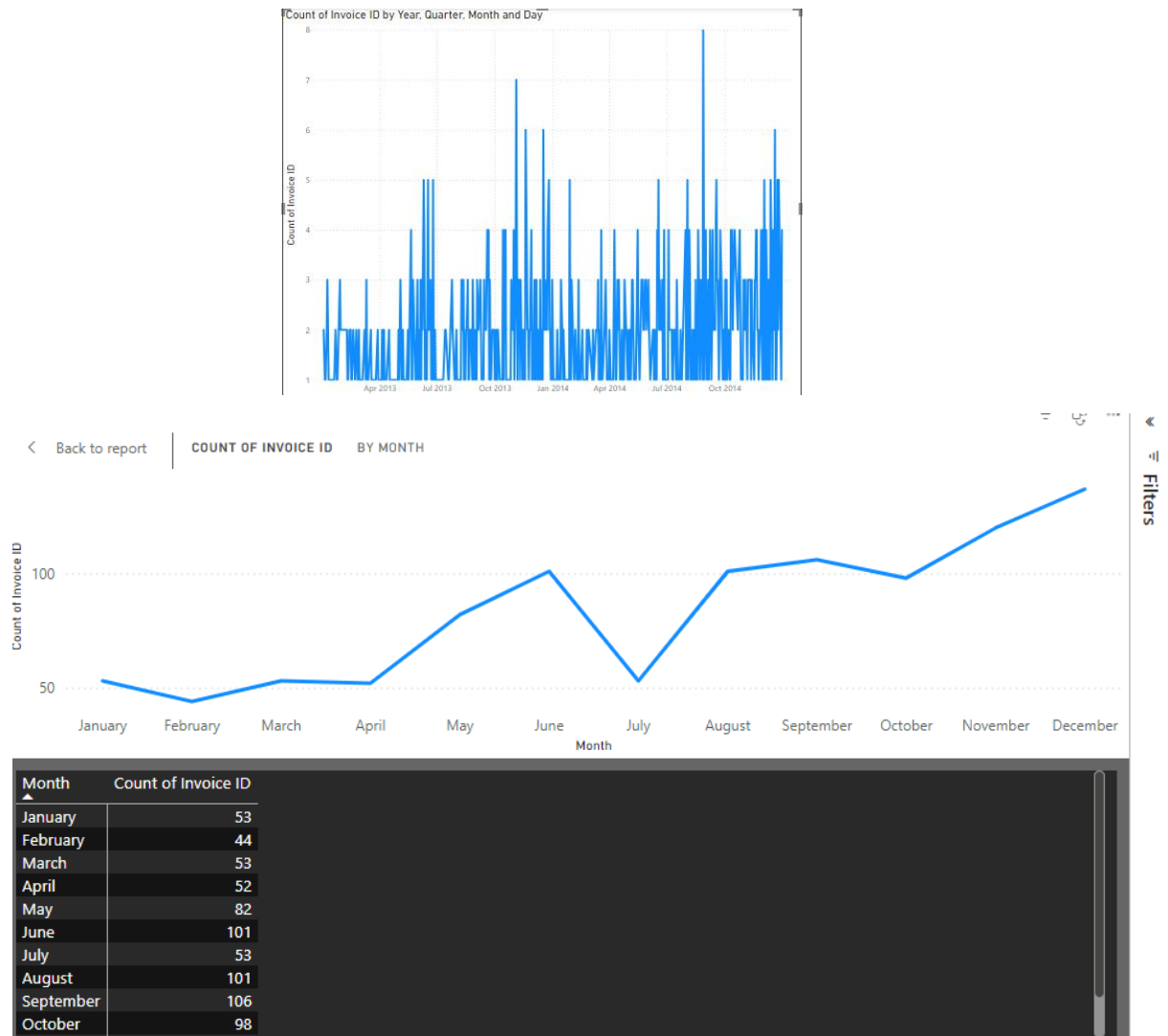
Formatting:

Use the **Format Painter** to rename and customize labels:

- Rename the chart to **"Sales Over Time"**.
- Rename the y-axis to **"Number of Sales"**.
- Apply similar formatting to other charts, such as renaming the y-axis of a column chart to **"Number of Customers"**.

<u>**Y-Axis is**</u> always the value, so we shall consider invoice-id which gives the number of sales transactions(Distict Count)

- **In Time related graphs, X-Axis is always the time.** The time related information in the data set is found in the order date field which records the date of each sale, for the x-axis.

Too many data values or variations lead to clutter.

Month	Count of Invoice ID
January	53
February	44
March	53
April	52
May	82
June	101
July	53
August	101
September	106
October	98

5. Revenue by Payment Method and City Visualization

1. **Objective**:

   Create a visualization to analyze **sales revenue** by payment method and city.

2. **Choosing the Chart Type**:

   - Opt for a **horizontal clustered bar chart** to accommodate three variables (payment method, city, and revenue).
   - In horizontal charts, the axes are interchanged:

     - **X-Axis**: Numerical value (total price).
     - **Y-Axis**: Categorical value (payment method).

3. **Setting Up the Chart**:

- Add **Total Price** to the **X-Axis**.
- Add **Payment Method** to the **Y-Axis**.
- This setup provides insights into revenue distribution by payment method.

4. **Adding City**

- Include **City** as the **Legend** to differentiate revenue contributions by city.
- Insights revealed:
- **Highest Revenue**: Achieved through **E-wallet transactions**, followed by **Cash** and **Credit Cards**.
- **Transaction Distribution**: London shows the highest number of transactions across all payment types due to having the largest customer base.
- Cardiff exhibits higher cash and credit card transaction revenue compared to other cities, unlike London, Burden, and Belfast.

5. **Using Focus Mode**:

- Click on the visual and select **Focus Mode** (top right corner) to zoom in for detailed analysis.

6. **Formatting the Chart**:

- Rename the chart title to **"Revenue Distribution by Payment Method and City."**
- Update the **Y-Axis** label to **"Payment Method."**
- Update the **X-Axis** label to **"Total Revenue."**

7. **Navigation**:

- Use the **Back** option to return to the main dashboard after analyzing the visual.

Payment	Aberdeen	Belfast	Cardiff	London	Oxford
Ewallet	1,288.00	1,298.00	1,191.00	1,570.50	286.50
Cash	1,240.00	1,002.00	1,442.50	1,512.50	303.50
Credit card	957.50	1,048.00	1,342.50	1,385.00	220.00

6. Revenue by Store Visualization

1. **Objective**:

   - Analyze which store has earned the highest revenue.

2. **Choosing the Chart Type**:

   - Use a **donut chart** to visualize the revenue distribution among stores.

3. **Setting Up the Chart**:

   - Drag the **Total Price** field to the **Values** area to represent revenue.
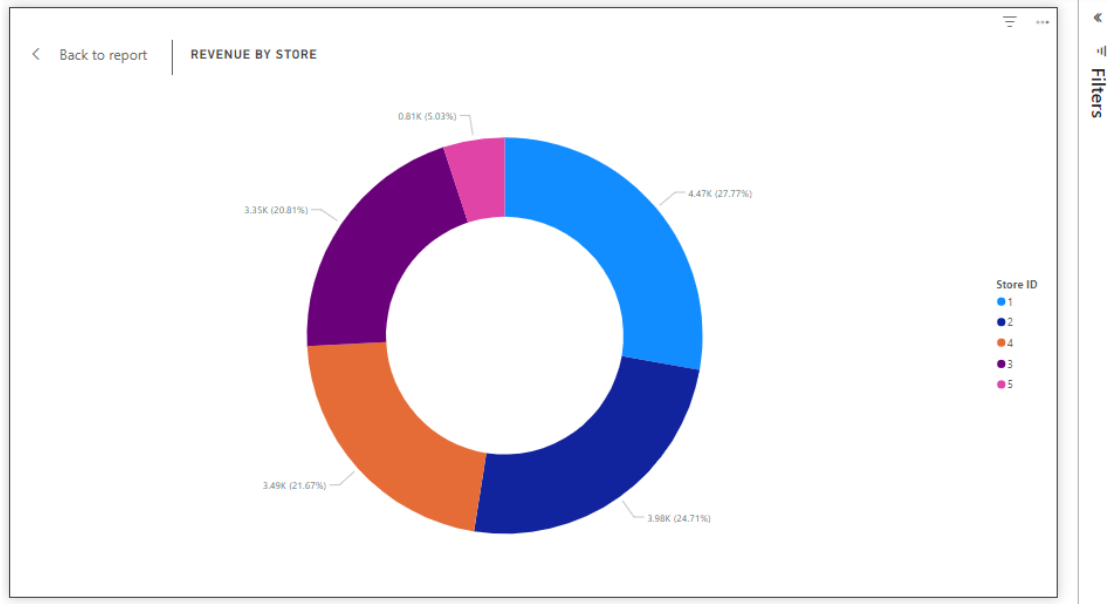   - Use **Store ID** as the **Legend** to categorize the revenue by each store.

4. **Insights Revealed**:

   The chart visually displays that:

   - **Store 1** has earned the highest revenue.
   - Followed by **Store 2**, **Store 4**, **Store 3**, and **Store 5**.

5. **Final Touch**:

   Rename the chart to **"Revenue by Store"** for clarity.

7. <u>Map Visual for Revenue Distribution</u>

1. **Objective**:

   - Visualize revenue distribution by city using a map.

2. **Creating a New Page**:

   - Open a new page in the Power BI report to accommodate the new visual.
   - Duplicate the previous page to retain formatting but delete all existing visuals.

3. **Adding the Map Visual**:

   - Select the **Map Visual (Basic Map)** from the visualization pane and place it on the canvas.
   - Add the **Location** field (City) to the map to represent data points.

4. **Handling Latitude and Longitude**:

   - No need for separate latitude and longitude values; the map visual can function with just city names.

5. **Customizing Bubble Size**:

- The map displays bubbles representing each city. Adjust the bubble size based on the **Sales Figure** to indicate revenue:
  - **Smaller Bubbles**: Indicate lower sales revenue (e.g., Oxford).
  - **Larger Bubbles**: Indicate higher sales revenue (e.g., London).

6. **Formatting the Visual**:

   - Use the **Format Painter** from the first page:

     Select a visual, click on **Format Painter**, then apply it to the map on the second page.

     - Rename the visual to **"Revenue Distribution Map"** for clarity.