Bike Sales Data Analysis using Power BI

Shweta Ajay Shinde
Masters in Data Analytics, San Jose State University
Data 230: Data Visualization
Instructor: Venkata Duvvuri
October 9, 2024

Power BI

Report on Advanced Analytics for Bike Sales Data using Power BI

1. Introduction:

As a data analyst, you're tasked with building reports and dashboards to support key business decisions using Bike sales data. Different teams in the organization have unique requirements:

- **Product Team:** They want to know which bikes are not selling as well as others.
- **Sales Team:** They are focused on forecasting sales for the upcoming year.
- **Warehouse Team:** They need an overview of how the warehousing and shipping locations are performing globally.

2. Data Set:

From the dataset, we have below listed column we can use for Data Visualization.

- Product Category
- Product Subcategory
- Product Name
- Product Description
- Product Price
- Order Total
- Shipping Method

3. Power BI Advances Features:

Power BI offers advanced analytics tools that will help create insights and visuals for each of these teams. The platform enables you to quickly generate reports and share them through dashboards, helping teams make decisions based on:

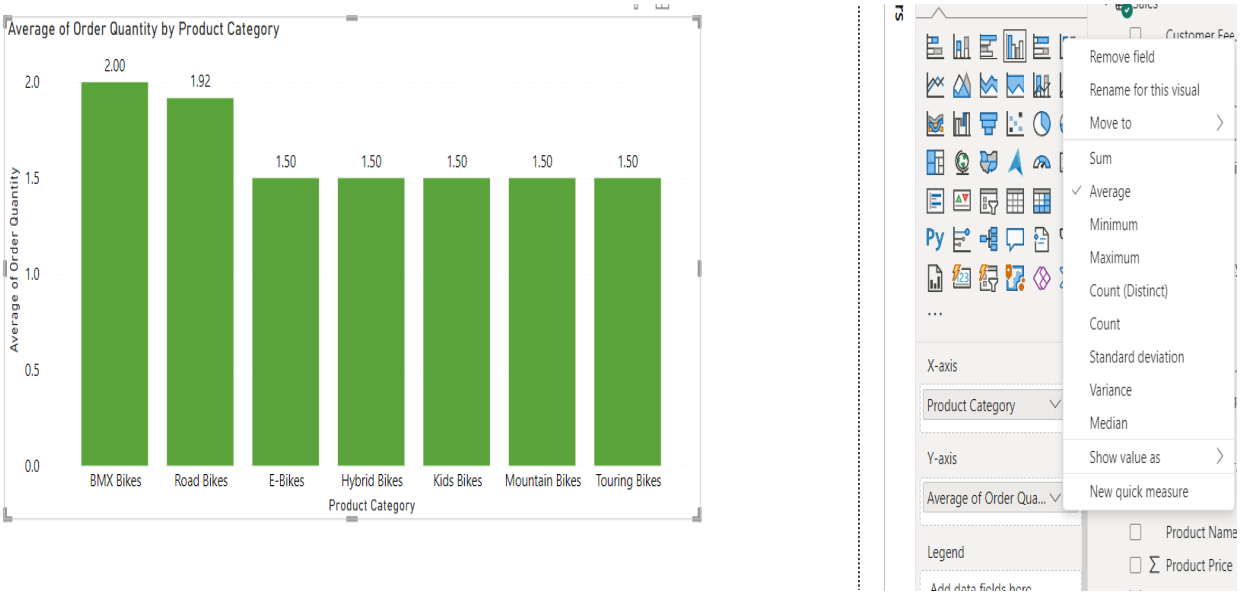
a) Exploring Statistical Summary:

Statistics help you understand your data by showing its distribution and key patterns. A statistical summary provides a quick overview, helping you spot trends, averages, and any outliers in your data.

By exploring a statistical summary, you can easily see patterns in customer behavior and product performance. For example, you can create a report for the Supply Chain team to show how often certain products are ordered. You can access these quick functions by right clicking the Y-axis and selecting **Average** field in the Visualizations pane, as illustrated in the Figure 1.

Power BI

Figure 1:
Average Order Quantity by Product Category



b) Histogram:

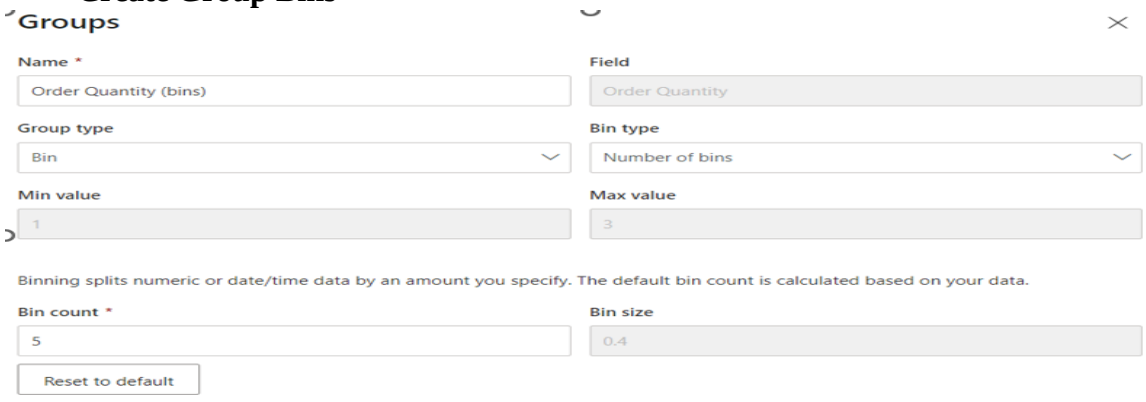
Histograms and bell curves are the most common way to display statistics about your semantic models. Figure 3 can be a major help to Warehouse Team to understand the quantity of order they get.

Binning is like grouping but applies to continuous data, like numbers or dates. It divides data into equal-sized groups (bins).

To create a histogram in Power BI using a clustered column chart, follow these steps as shown in Figure 2:

1. Select the clustered column chart icon from the Visualization pane.
2. In the Fields pane, right-click the field you want to analyze (like *OrderQty*) and choose New Group.
3. In the Groups window, rename the group to *Order Bins (Buckets)*.
4. Set Group type to Bin and Bin Type to Number of bins.

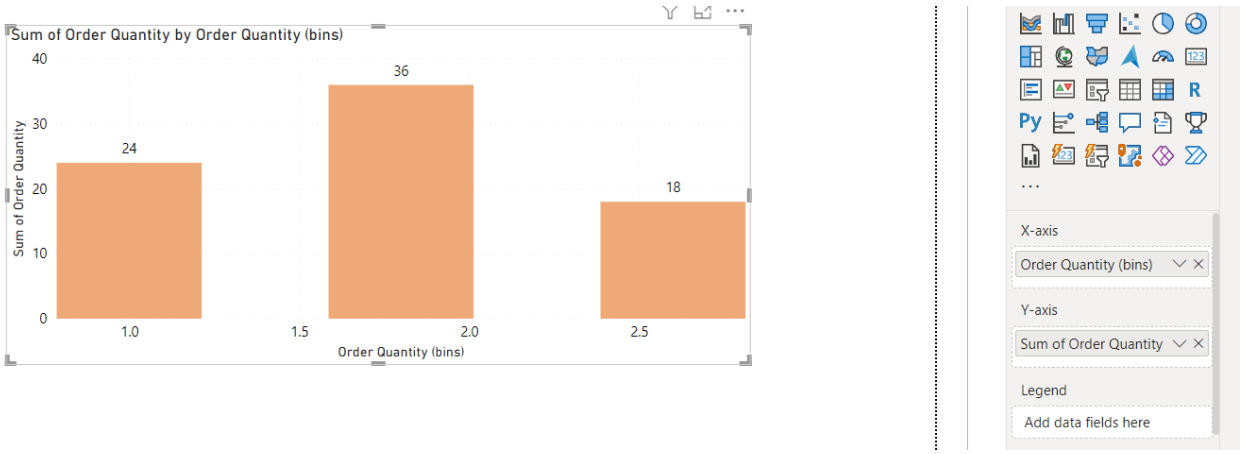
Figure 2:
Create Group Bins



To finish the visual as Figure 3:

1. Drag **OrderQty** from the Fields pane to the **Y-axis** field.
2. Drag **Order Bins (Buckets)** to the Y-**Axis** field.

Figure 3:
Histogram



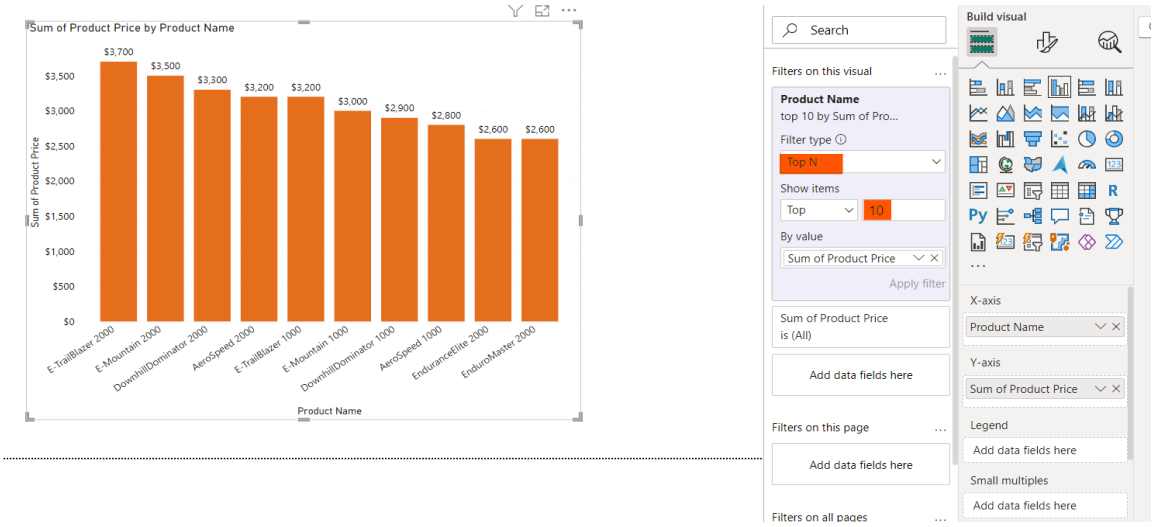
c) Top N

Top N analysis helps show the top (or bottom) items in a list, like the top 10 selling products. Supply Chain teams can use Top N analysis to focus on the most critical products

To apply the Top N filter as shown in Figure 4:

1. Select the **Product Name** field on your report.
2. In the **Filters** pane, choose **Top N** from the Filter type list.
3. Set **Top** to 10.
4. Choose **Product Price** as the value to filter by.

Figure 4:
TOP N



Power BI

Alternate way to calculate TOPN using Figure 5 formula.

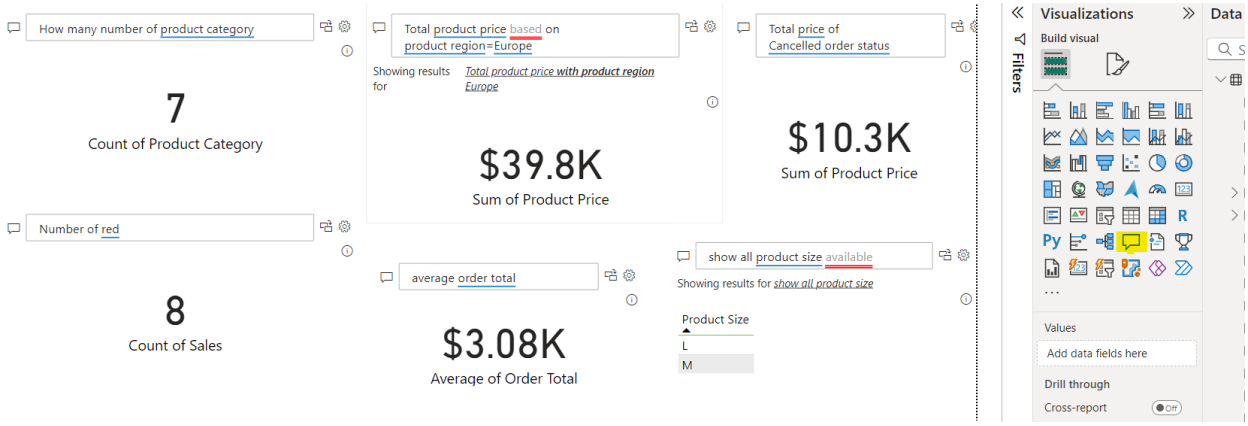
Figure 5:
TOPN Function

```
Top 10 Products = SUMX ( TOPN ( 10, Sales, Sales[Product Price]), [Product Price] )
```

d) Q & A

You've created a report for the Finance team, and now the team members have questions about various other views or insights that they are interested in. Power BI has a built-in Q&A visual that allows users to ask their own questions and get answers, so you don't have to address each individual question. The Q&A visual is an effective tool because it allows users to quickly get answers about the data independently, which saves time for everyone involved. The Q&A visual is unique in that it does not require knowledge of Power BI to use the visual; users can ask their question and they, too, can create insightful visuals as shown in Figure 6.

Figure 6:
Q and A



e) Identify Outliers

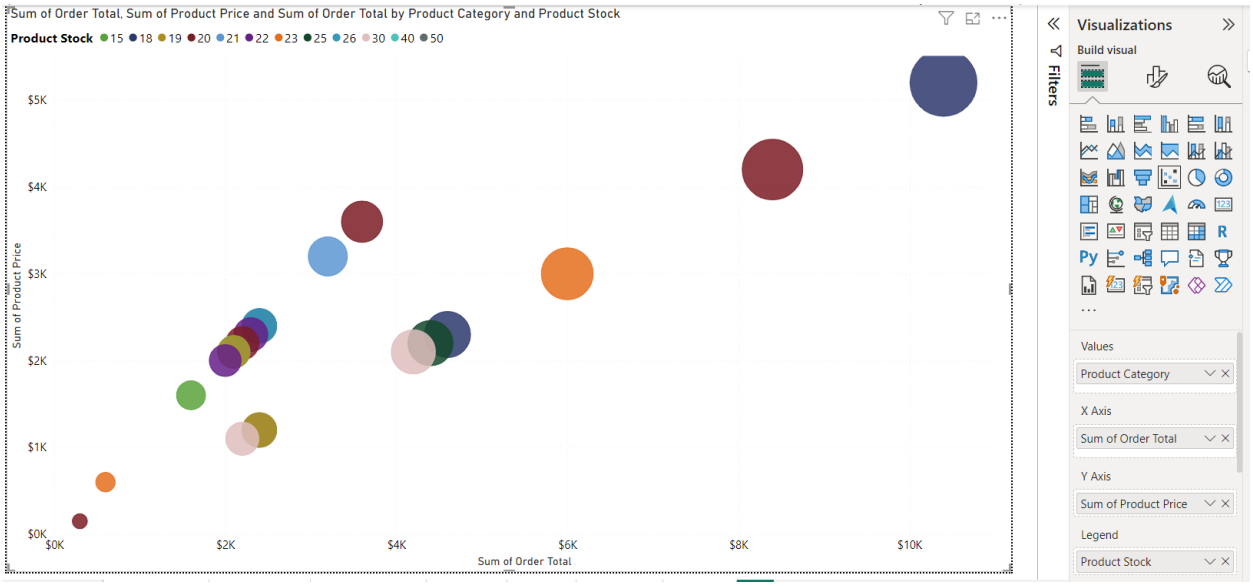
An outlier is a data point that stands out because it doesn't match the expected pattern based on historical data. Identifying outliers is important, as they can reveal valuable insights for business decisions.

For example, while analyzing a shipping warehouse, you see a sudden spike in orders for a certain product category. To get outliers as Figure 7:

To add a scatter chart to your Power BI report, follow these steps:

1. **Select the Scatter Chart**: In the Visualizations pane, click on the scatter chart icon.
2. **Add Data to the X Axis**: Drag the **Orders Shipped** field from the Fields pane and drop it into the **X Axis**.
3. **Add Data to the Y Axis**: Drag the **Qty Orders** field from the Fields pane and drop it into the **Y Axis**.

Figure 7:
Outliers



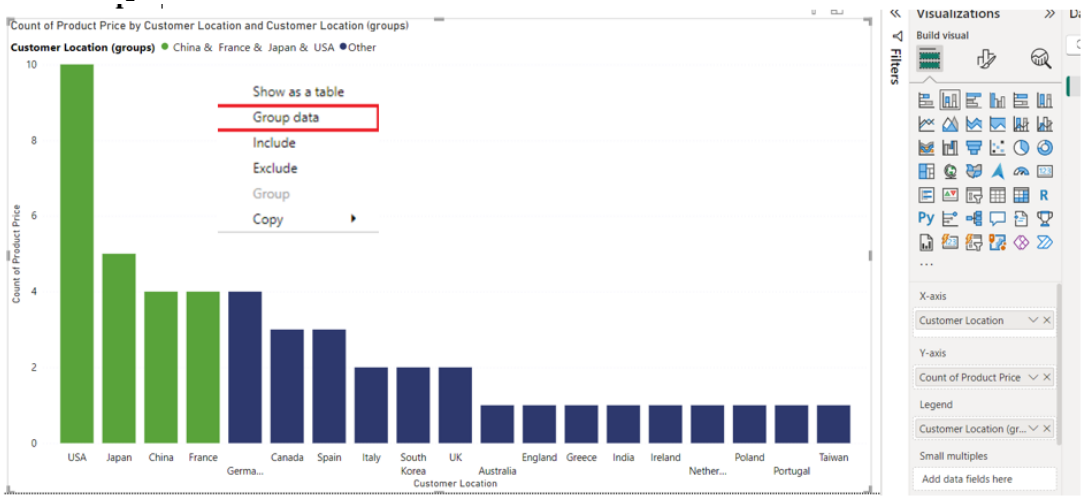
f) Group

When you create visuals in Power BI Desktop, the data is automatically grouped based on the underlying values. You can refine how these groups are displayed and create new groups or bins as needed.

- Grouping: Used for categorical data. You can combine two or more data points into a single group. Please refer to Figure 8

To create the group, use **Ctrl + click** to select the data points on the visual that you want to group. Right-click one of those selected data points and then select the **Group data** option.

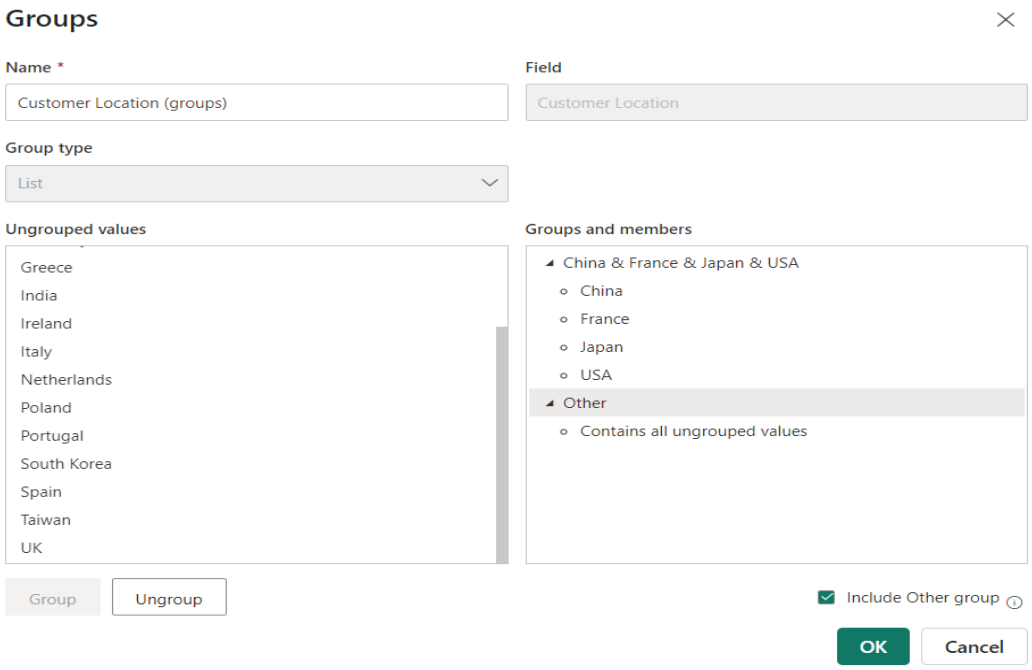
Figure 8:
Groups



To edit a group in Power BI as showed in Figure 9:

1. **Find the Group Field**: In the **Legend** bucket or **Fields** pane.
2. **Right-Click**: Select **Edit Groups**.
3. **Review Groups**: In the Groups window, see the existing groups and their members.
4. **Refresh Data**: New items will go into the *other* group if ungrouped.

Figure 9:
Edit groups



g) Clustering Techniques

Clustering helps you find similar data points that are different from the rest. In Power BI, clustering groups similar data automatically. For example, you can use it to find patterns in sales data, like grouping customers by age or location. clustering helps the Sales and Marketing teams identify patterns, trends, and outliers in customer behavior and product performance.

To start, add a scatter chart to your report. Then, put **Order Qty** on the x-axis, **Sales** on the y-axis, and **Product Price** in the Values field. The chart will show lots of data points, as in Figure 10. The new cluster field is added to the Legend section of your scatter chart. You can now use it for cross-highlighting, like any other legend. You'll also find the cluster field in your field list and can use it in other visuals.

To edit the cluster, right-click the cluster field and select Edit clusters.

Figure 10:
Clustering

4. Conclusion:

Power BI offers a variety of advanced features that make data analysis more insightful and efficient for different teams. Tools like statistical summaries, histograms, and clustering help the Sales and Marketing teams identify patterns, trends, and outliers in customer behavior and product performance. Supply Chain teams can use Top N analysis to focus on the most critical products and suppliers, while the Finance team can benefit from the Q&A feature, quickly getting answers about budgets and forecasts. Additionally, Operations teams can use grouping and clustering techniques to segment data for better resource management. Overall, these features empower all teams to make informed, data-driven decisions.

References

Microsoft. (n.d.). *Introduction to performing analytics in Power BI*. Microsoft Learn. https://learn.microsoft.com/en-us/training/modules/perform-analytics-power-bi/1-introduction

Shinde, Shweta (2024). *Bike sales analysis [Power BI project file]*. GitHub. https://github.com/ShindeShwetaK/All-PowerBi-Projects/blob/main/Power-BI%20Project/Bike_Sales_analysis.pbix