**Guidelines for Data Visualization and Analysis Project**

**About the Project:**

In this project, you will be working with a dataset from the Superstore, aiming to answer 30 scenario-based questions through data visualisation and analysis. Your objective is to select the best chart for each question, explain your choice. This project will showcase your proficiency in data visualisation, critical thinking, and effective communication.

**Skills Required:**

* Proficiency in data visualisation concepts and techniques.
* Familiarity with Tableau or a similar data visualisation tool.
* Strong analytical and problem-solving skills.
* Ability to choose appropriate charts based on data characteristics and question requirements.
* Clear and concise communication skills.

**Deliverables:**

* A Google document containing solutions to the scenario based questions including the screenshot of relevant chart picked for each scenario, presented in a concise and well-structured format. Make sure to provide explanations that highlight your problem-solving skills.

**Rubrics for Assessment:**

Question Responses:

* Accuracy and completeness of answers for all 30 questions.
* Clear and concise explanations that address the question's context.

Chart Selection and Explanation:

* Thoughtful rationale for choosing specific chart types.
* Justification based on data characteristics, context, and communication goals.

Creative Enhancements:

* Effective use of creative elements to enhance visualisation quality.
* Enhancements that contribute to better understanding or engagement.

**Note**:

* Duplicate this document and proceed to write your solutions.
* For each scenario and question, provide a justification for the choice of chart type. Explain why it is the best option to visualise the data effectively.
* Attach screenshots of the charts you have created in Tableau for each scenario and question using the Superstore dataset. Label them clearly to match the corresponding questions in the Google Document.
* Submit the duplicated google doc file after completion.

Use these guidelines to structure your data visualisation and analysis project. Remember to maintain consistency in your responses, explanations, and visualisation styles. This project will not only demonstrate your skills but also your ability to effectively communicate complex information through visualisations. Good luck!

**Problem Statement: Choose the Best chart for any 30 scenario based questions from Superstore Dataset.**

Imagine you are a data enthusiast aiming to excel in data visualisation and analysis. In this task, you have been given any 30 scenario-based questions derived from the Superstore dataset, and your objective is to provide insightful answers using appropriate charts. For each question, you need to select a chart that best represents the data, explain why you chose that specific chart, and then proceed to build the chosen chart using Tableau.

Your responses should be succinct, organised, and illustrative of your problem-solving capabilities.

**Dataset Link:**

<https://community.tableau.com/s/question/0D54T00000CWeX8SAL/sample-superstore-sales-excelxls>

**Please keep in mind:**

1. **Answer Completion**: Ensure that you furnish answers for all any 30 questions and build charts for them.
2. **Encouraged Creativity**: Don't hesitate to employ visuals, creative elements, or any other innovative approaches to enhance the quality of your responses.

By completing this task effectively, you'll not only demonstrate your proficiency in data visualisation and analysis but also showcase your ability to effectively communicate complex concepts through both text and charts.

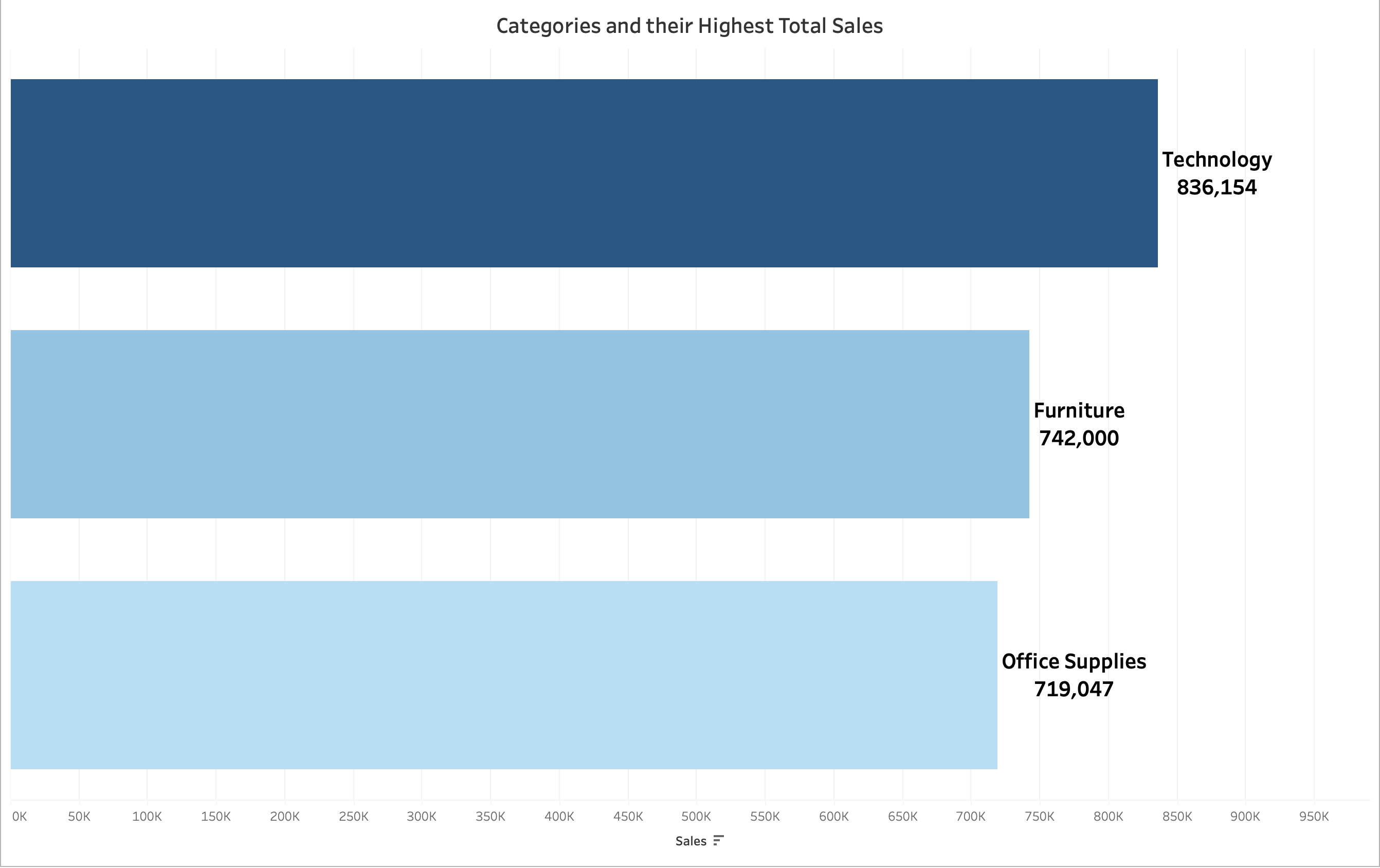
**Good luck!**

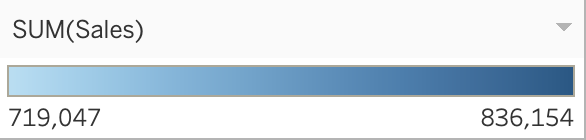
**Questions:**

1. Which product categories have the highest total sales in the "Superstore" dataset?

Ans.

**Chart**: Bar Chart



****

**Why Bar Chart?**

It is ideal for comparing total sales across categories. It allows for easy comparison of values and identifies the categories with the highest sales.

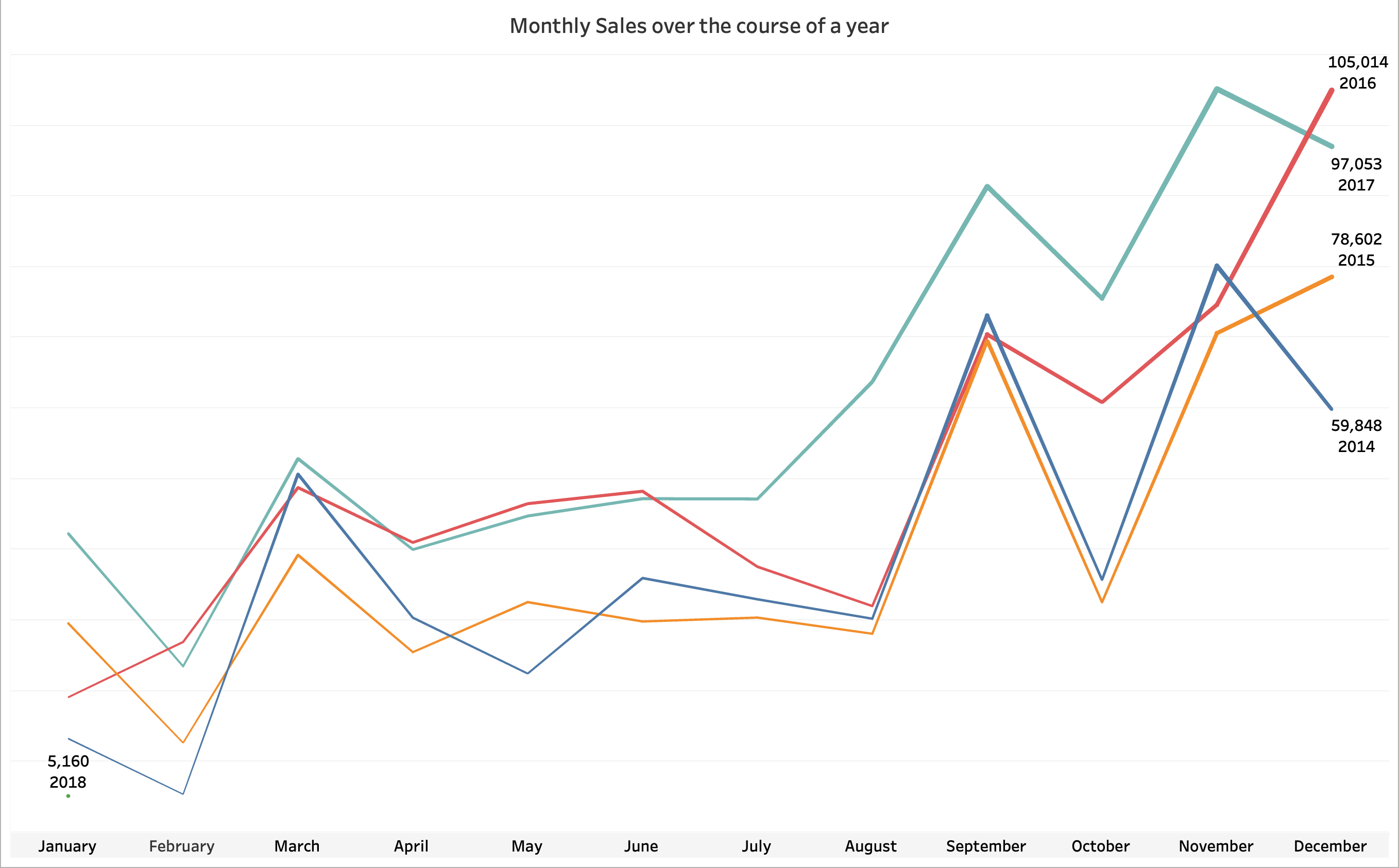
**Insight:**

From the Graph it is clear that **Technology** is the **highest total sales** category with the value of **8,36,154/-** among other categories.

1. How do the monthly sales amounts change over the course of a year?

Ans.

**Chart:** Line Chart



**Why line chart?**

It Effectively shows changes in **monthly sales over time**. It’s easy to observe trends, seasonal trends and variations throughout the year.

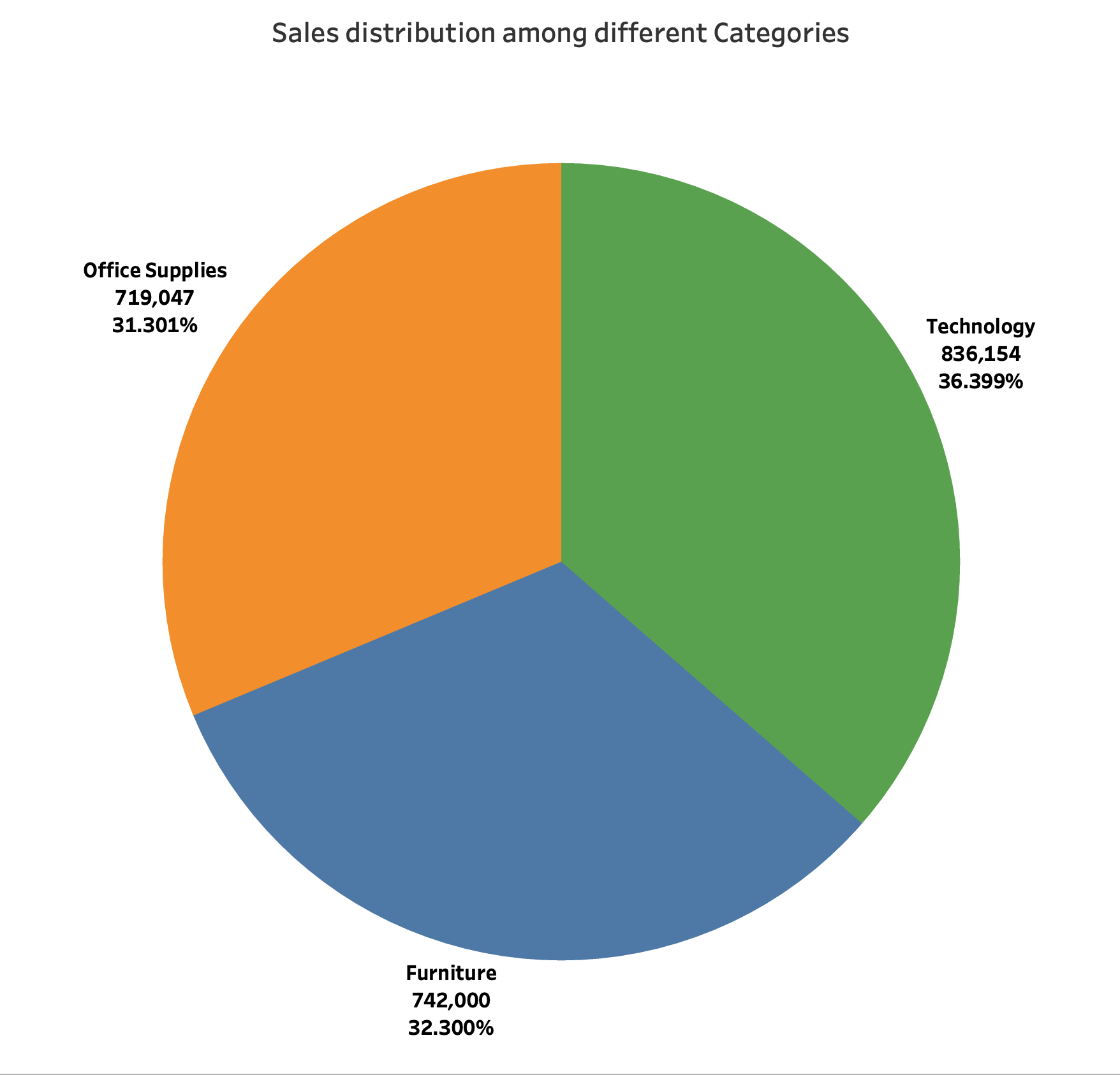
**Insight:**

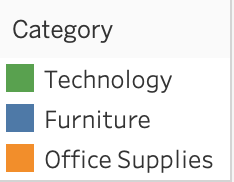
From the multiple line chart, we gain valuable insights into the monthly sales trends throughout the year. We observe a steady increase in sales from **2014 (69,546) to 2016 (96,999)**. In **2017**, sales saw a significant rise up to **November(118,448)** but experienced a sudden decline by the end of **December(83,829).**

1. How is the total sales amount distributed among different product categories?

Ans.

**Chart:** Pie Chart



****

**Why pie chart?**

It visually represents the distribution of Sales among different categories. It’s useful for showing proportions and the contribution of each category to total Sales.

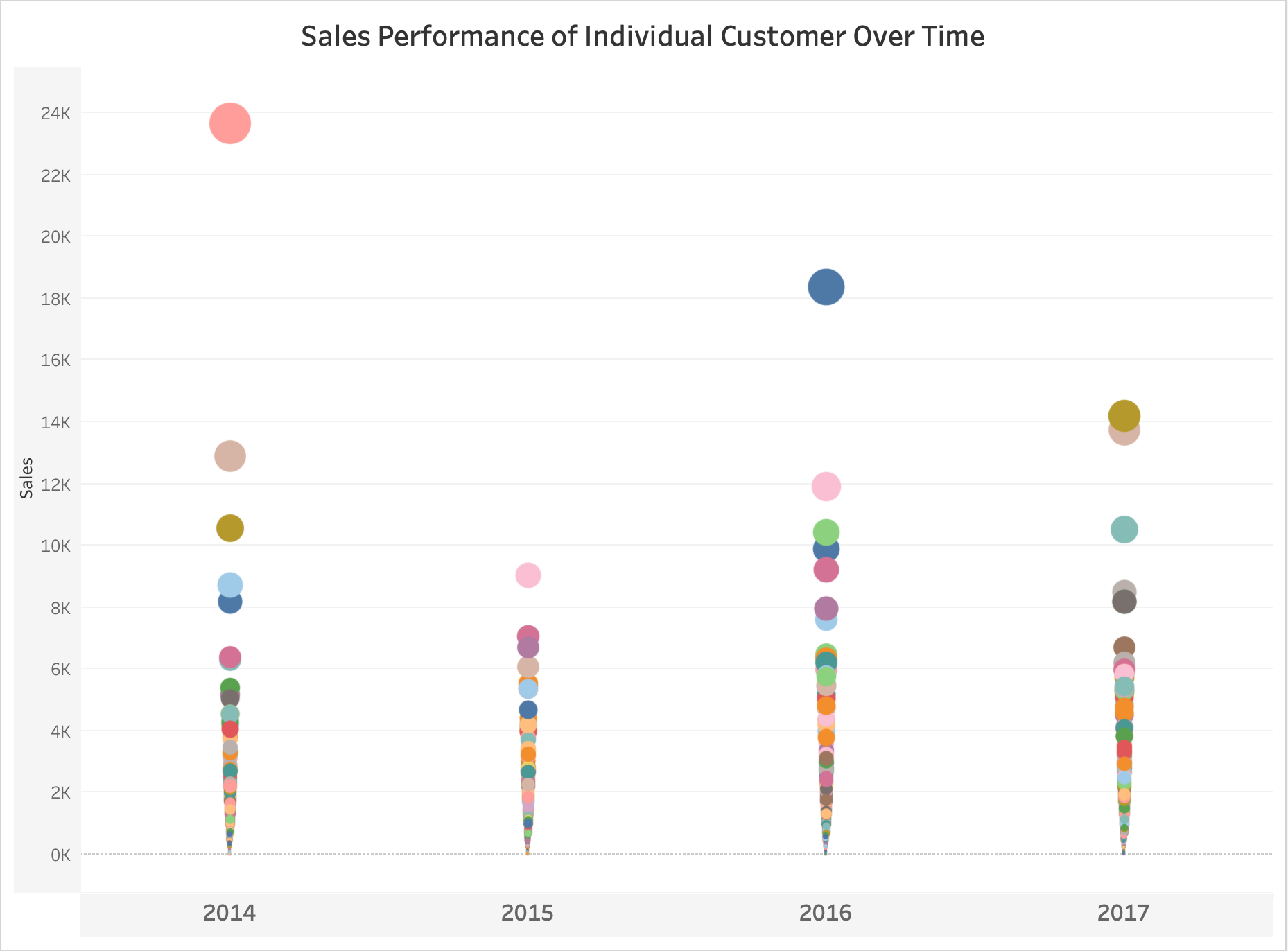
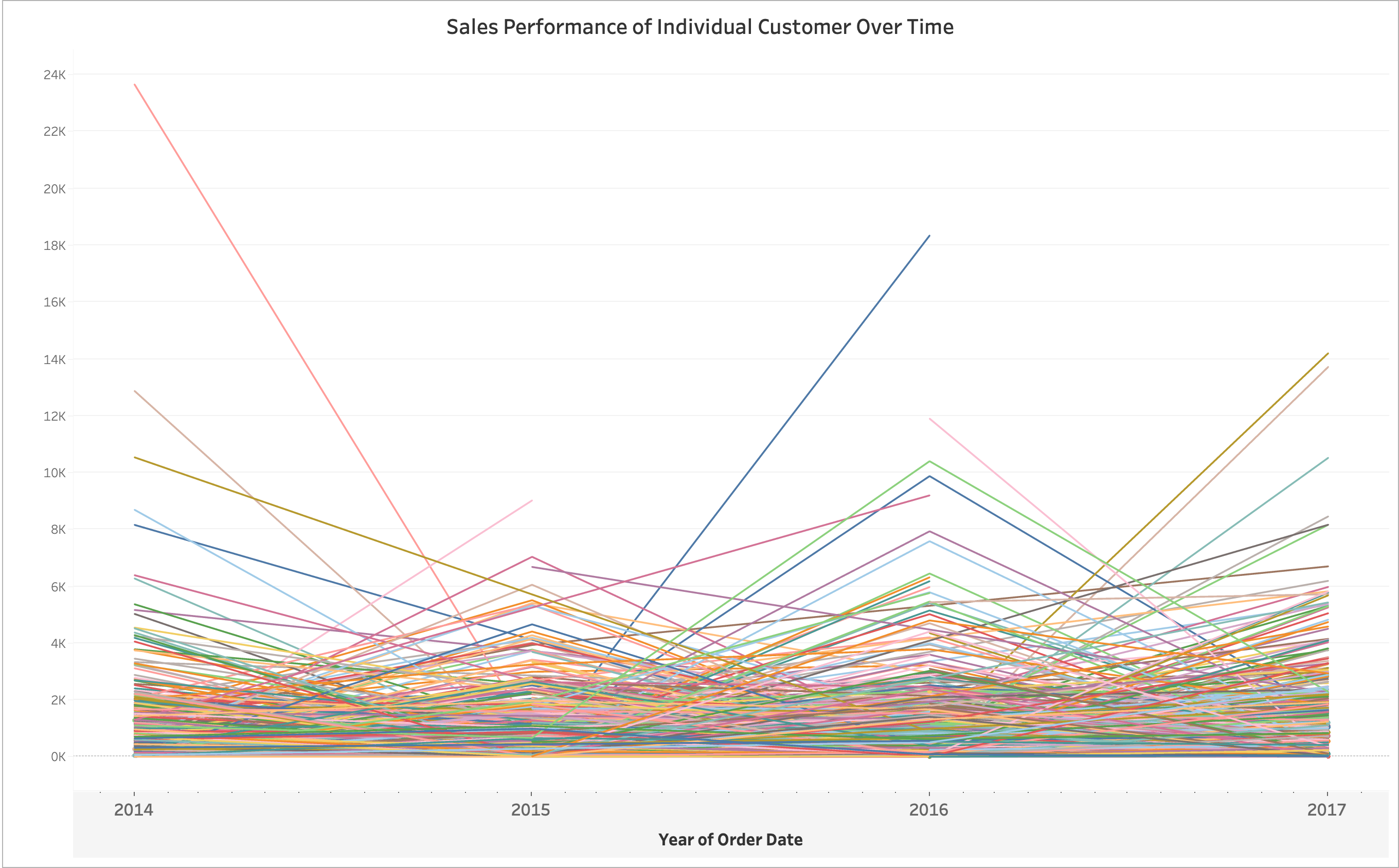
**Insight:**

From the graph, All the **three categories show almost equal contribution** to total sales. Starting from lowest contributor, **Office Supplies** - contributed **31.30%**, **Furniture** - contributed **32.30%** and **Technology** - contributed **36.40%** which is higher than the other two but still all three lie around the near percentage value.

1. Can we analyze the sales performance of individual customers over time?

Ans.

**Chart:** Line Chart or Scatter Plot

****

**Why Line or Scatter chart?**

Line chart can show sales trends for individual customers over time, while a scatter plot can show sales volume over time with each point representing a customer’s purchase.

**Insight:**

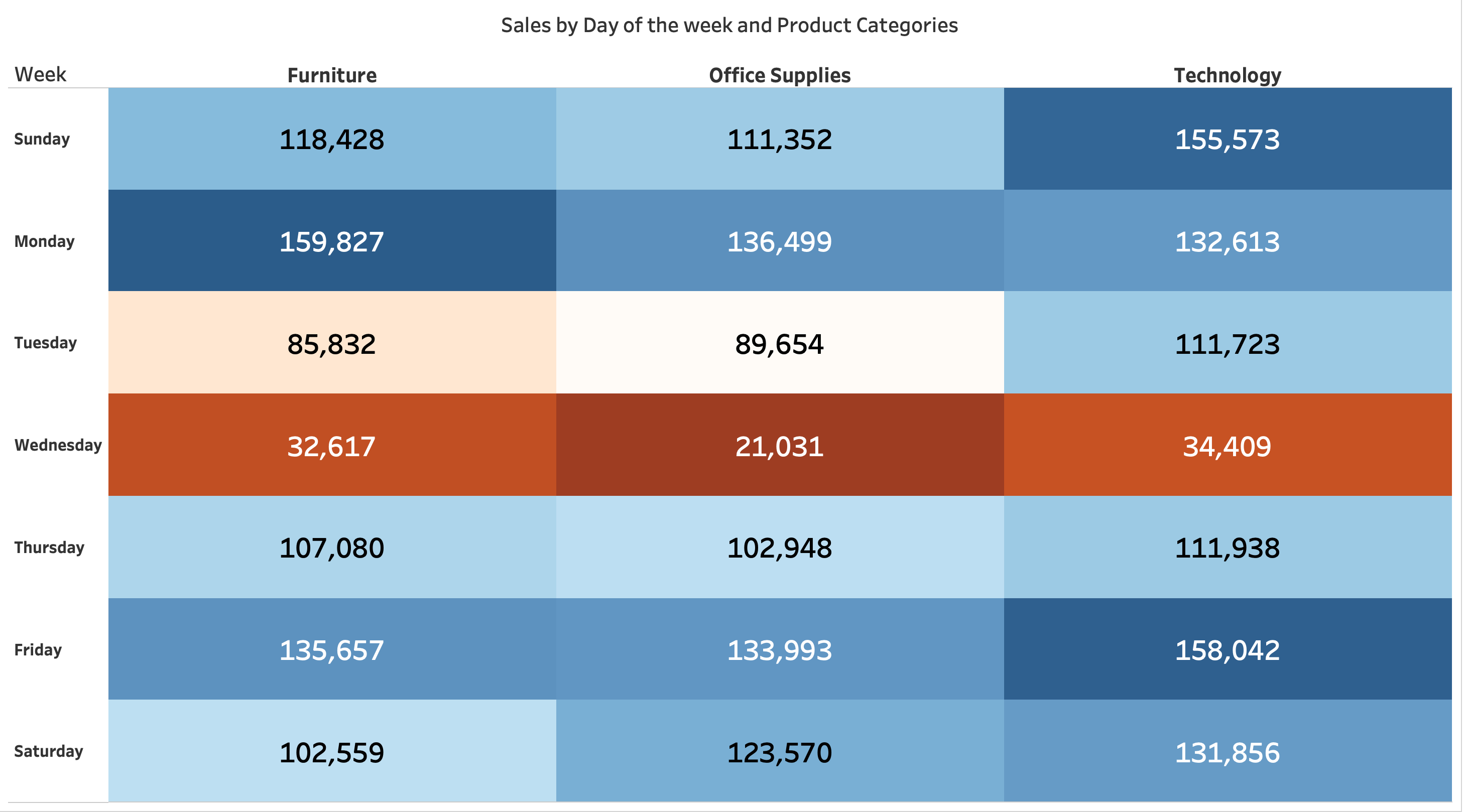
Most of the orders have seasonality factors.

Customers base is diverse, With some customer being more active than other others.

1. How do sales vary based on different days of the week and product categories?

Ans.

**Chart:** Heatmap

****

**Why Heatmap chart?**

Heatmap is effective for showing variations in sales across days and categories. The color intensity makes it easy to identify high and low sales periods.

**Insight:**

**Stronger Sales on Certain Days:** Some product categories might show significantly higher sales on specific days of the week. For example, the "Office Supplies" category has dark blue colors for days (like Monday or Friday), it suggests higher sales for this category during the mid-week period.

**Weaker Sales on Certain Days:** Conversely, if some product categories show consistently lower sales on certain days (e.g., mid-week (Tuesday and Wednesday)), it indicates potentially lower demand for these products during those days.

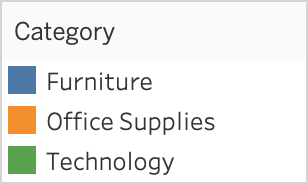
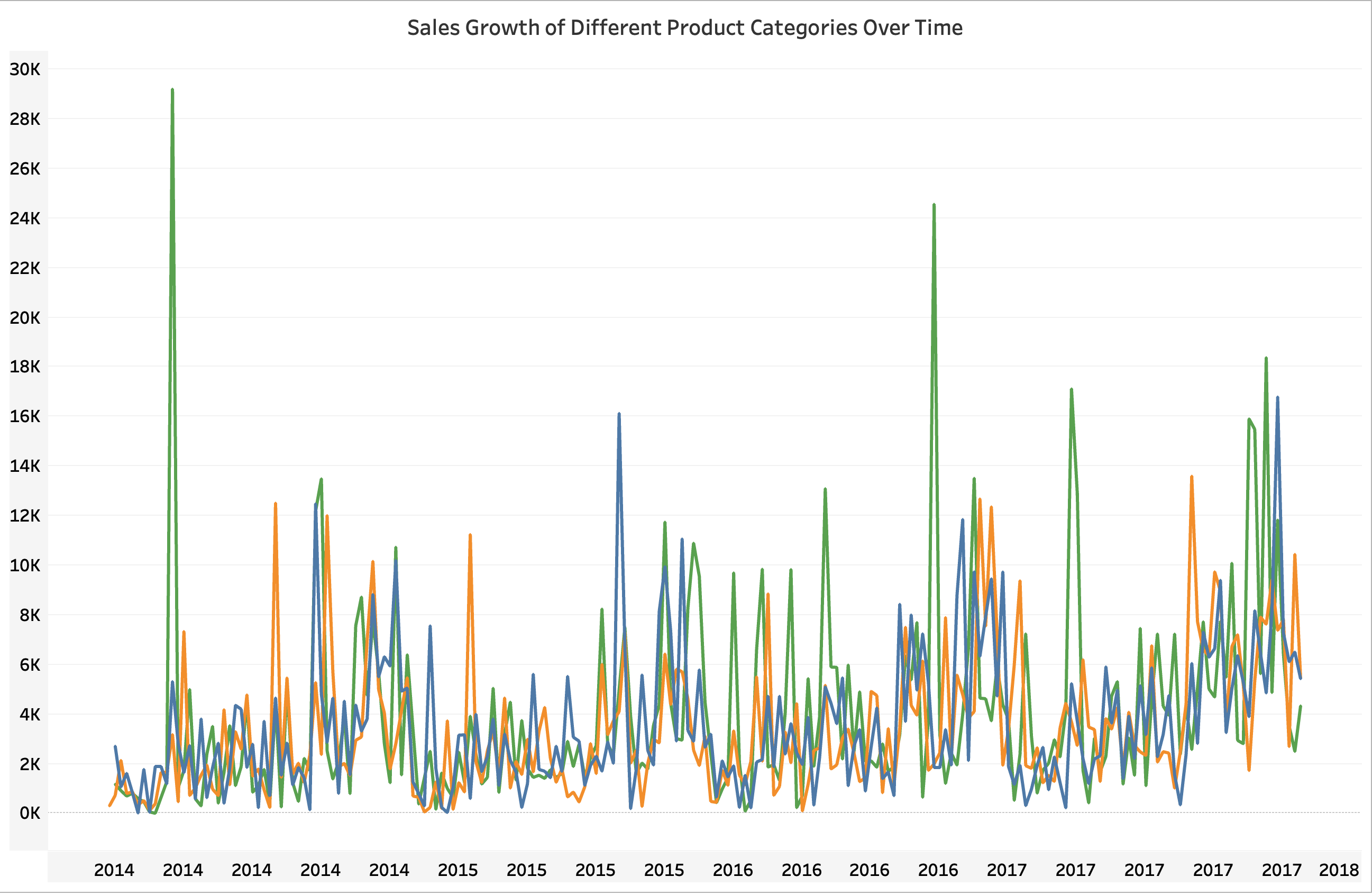
**Identifying Patterns:** Observing the general color intensity across rows and columns can reveal overall patterns. If, for example, the whole row for "Wednesday" is relatively red color, it implies lower overall sales across all categories during that day of the week.

**Category-Specific Trends:** It helps to understand whether there are specific categories that perform better or worse on particular days. For instance, the "Technology" category might have higher sales on Mondays as businesses might be more likely to place technology-related orders at the beginning of the week.

1. Can we visualise the sales growth of different product categories over time?

Ans.

**Chart:** Line Chart

****

**Why line chart?**

It can **display the growth trajectory** of each **category over time**, making it easy to compare trends and identify categories with significant growth.

**Insight:**

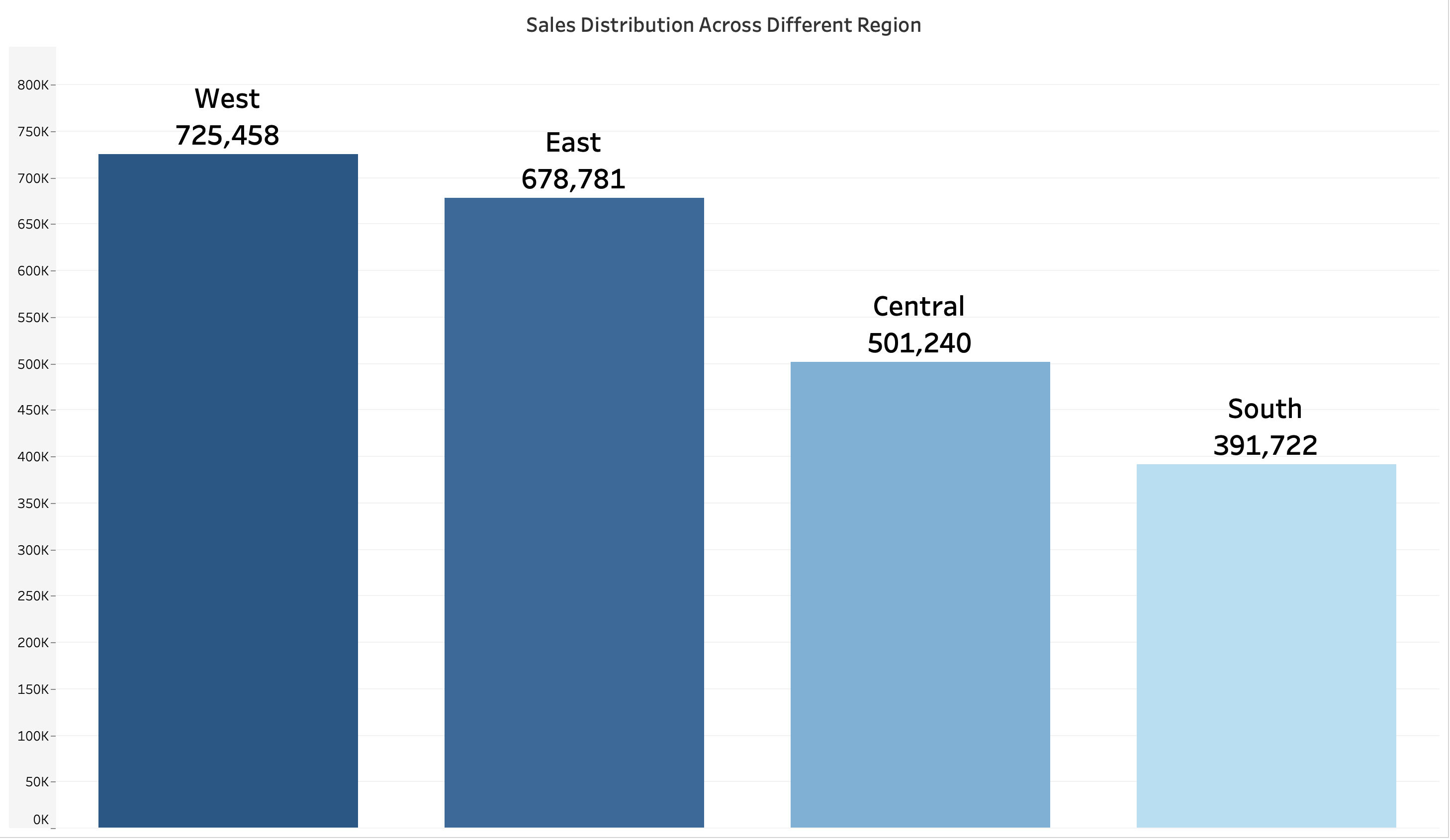
The chart shows that the "Technology" category consistently has higher sales and a steeper growth trend compared to "Furniture," it suggests that the Technology product line is more successful and potentially has a greater potential for future growth. Also, Thw sales of "Office Supplies" show a seasonal spike during back-to-school periods, it could inform future marketing and inventory management strategies for that category.

By analyzing this chart effectively, we can gain a comprehensive understanding of the sales performance of different product categories, identify growth opportunities, optimize inventory levels, and implement targeted marketing campaigns.

1. How does the sales distribution vary across different regions in the "Superstore" dataset?

Ans.

**Chart:** Bar Chart

****

**Why Bar chart?**

It effectively highlights the differences in **total sales across regions**, making it easy to compare values visually. It allows for quick identification of which regions are performing better or worse in terms of sales.

**Insight:**

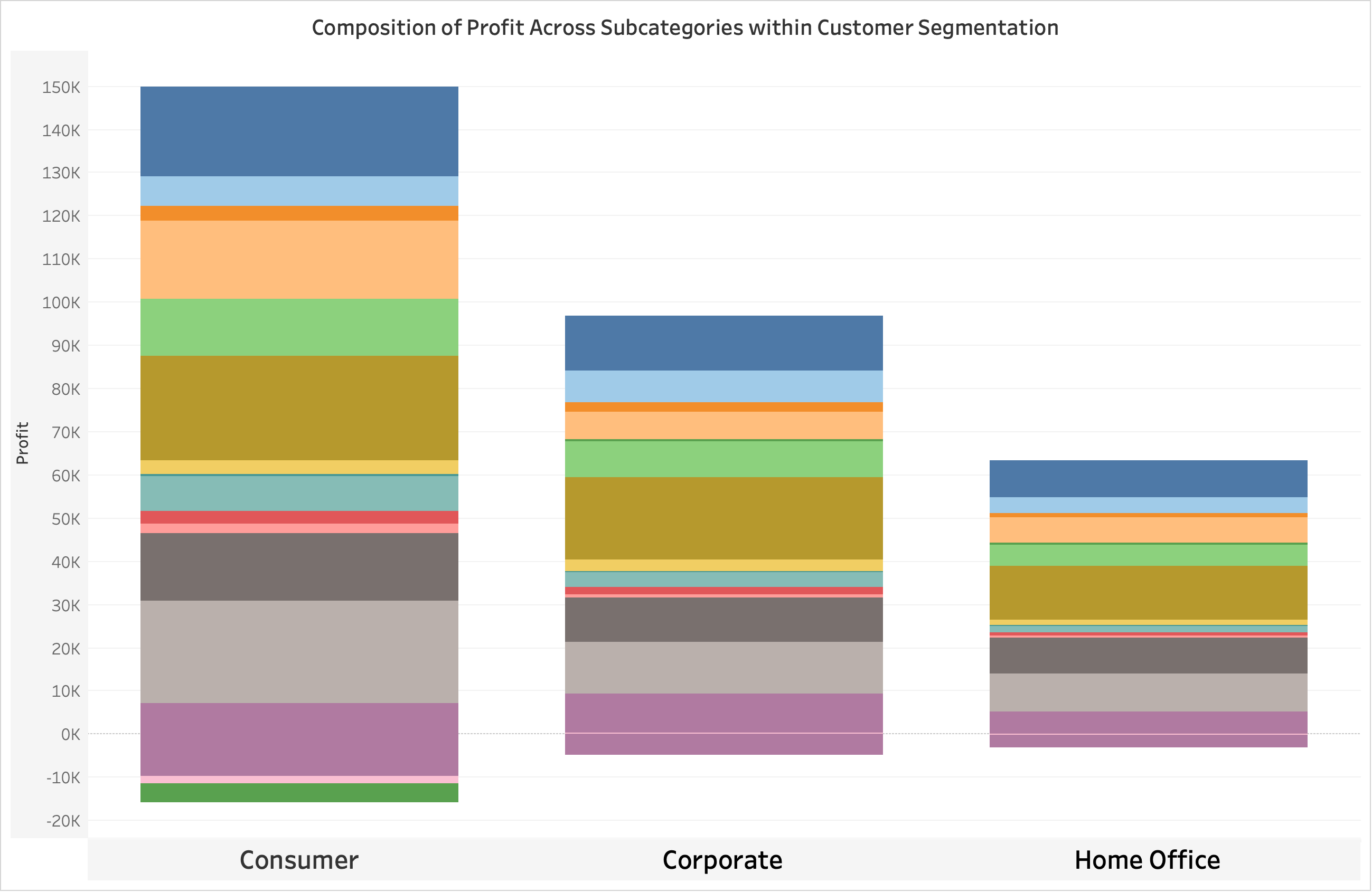
**West** is doing much better than other region. Might be because of higher population or central state is located in West region.

|  |  |
| --- | --- |
| **Region** | **Sales** |
| Central | 501239.89 |
| East | 678781.24 |
| South | 391721.90 |
| West | 725457.82 |

1. Can we visualise the composition of profits across various subcategories within different customer segments?

Ans.

**Chart:** Stacked Bar Chart

****

**Why Stacked Bar chart?**

A stacked bar chart shows the composition of profits across subcategories and segments, allowing for easy comparison and assessment of segment profitability.

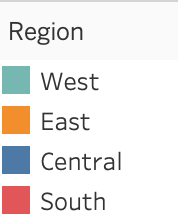
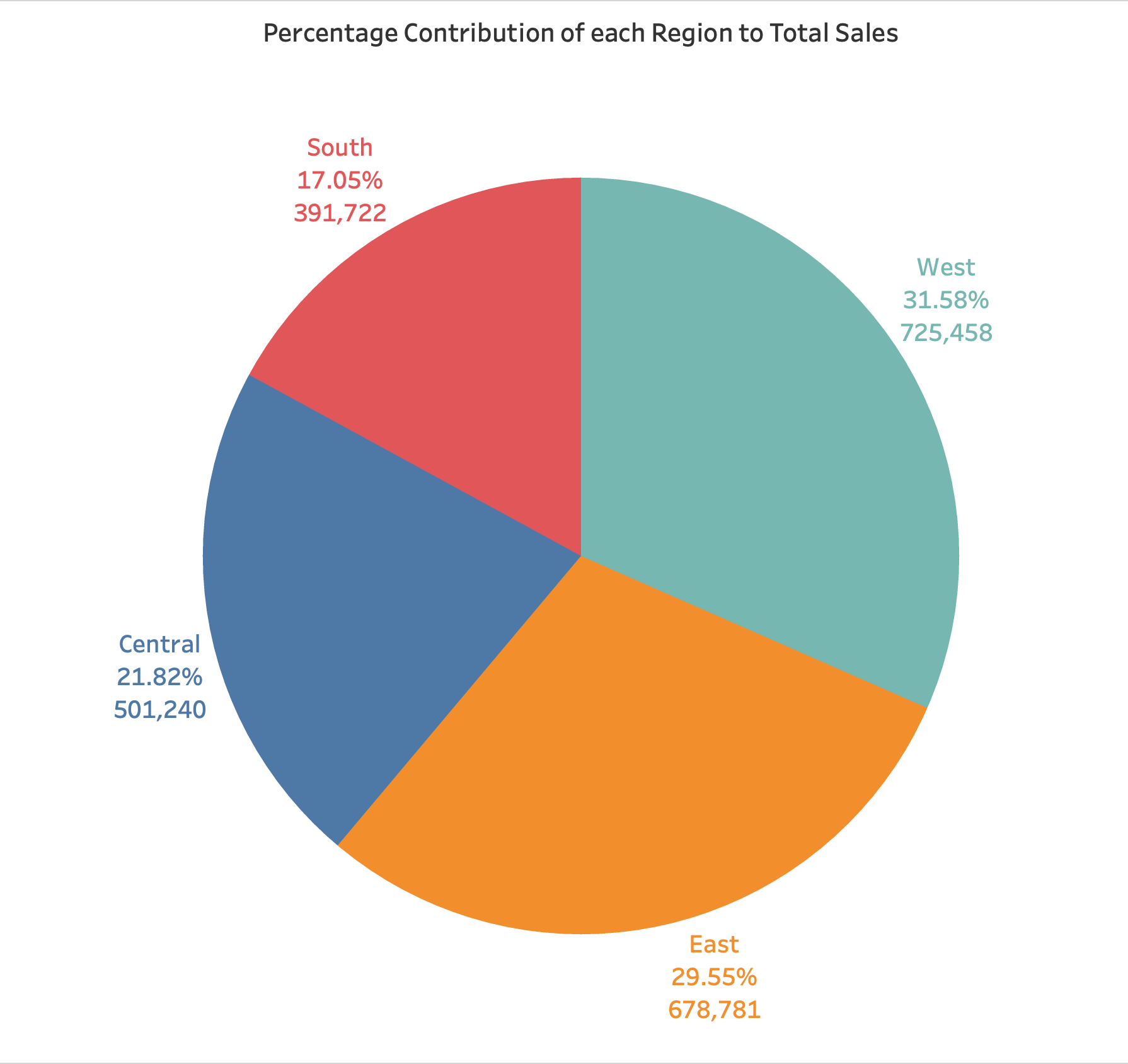
**Insight:**

The chart clearly highlights which subcategories are the most profitable for each customer segment. You can see which subcategories contribute the largest portion of the total profit for each segment. For instance, you might find that "Copiers" are the most significant profit driver for all segment, while "Phones" are more critical for the "Home Office" segment.

1. What is the percentage contribution of each region to the overall sales?

Ans.

**Chart:** Pie Chart



**Why pie chart?**

These charts effectively depict the contribution of each region to total sales, making it easy to identify the most and least contributing regions

**Insight:**

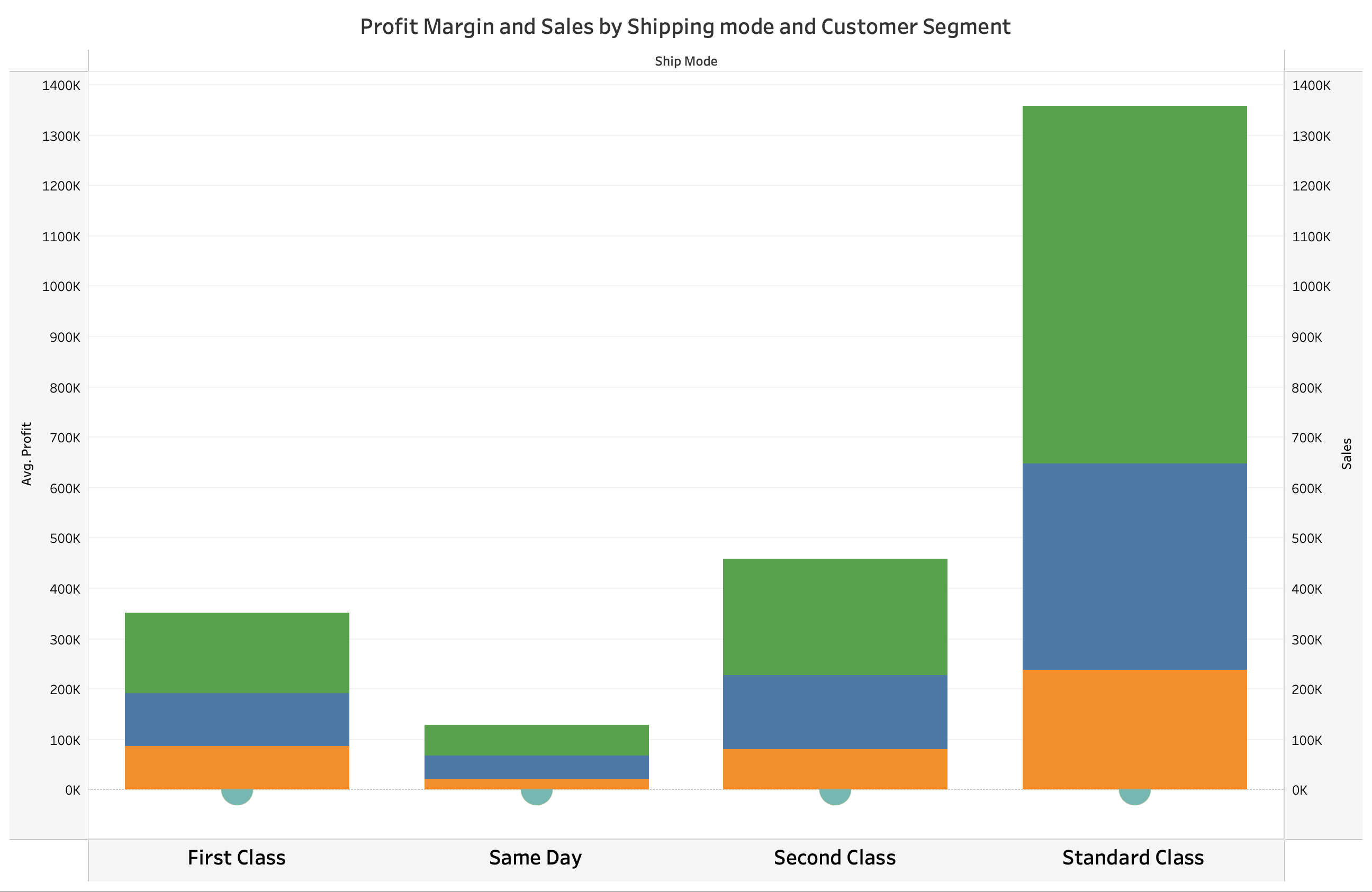
From the graph, It is clear that west has contributed 31.58% of sales. East has contributed 29.55% of sales. Central has contributed 21.82% of sales and South has contributed 17.05% of sales.

|  |  |
| --- | --- |
| **Region** | **Percent of total Sales** |
| West | 31.58% |
| East | 29.55% |
| Central | 21.82% |
| South | 17.05% |

1. Can we visualise the profit margins associated with different shipping modes and customer segments?

Ans.

**Chart:** Bar Chart with Duel Axis

****

**Why bar chart with a dual chart?**

It can show profit margins alongside sales figures, allowing for comparison across shipping modes and customer segments.

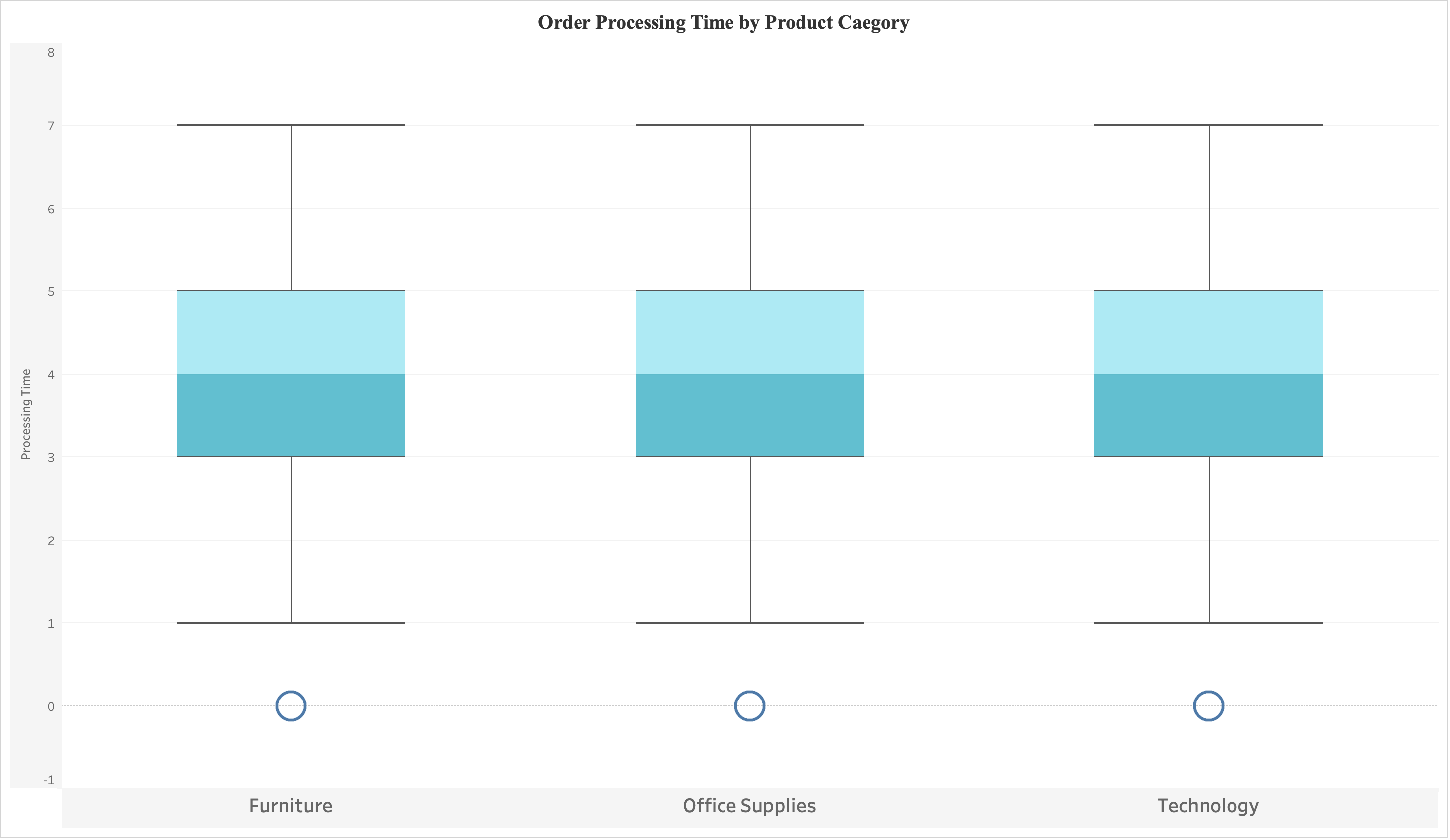
**Insight:**

The Profit margin of Consumer is more across every shipping mode and in particular the segment of consumer in Standard Class shipping mode is the highest. Meaning that standard shipping mode is more preferred than the others.

1. How long does it take to process orders for different product categories?

Ans.

**Chart:** Box Plot

****

**Why Box Plot chart?**

A box plot can display the processing time distribution for each category, highlighting the median, quartiles, and any outliers. access the data from previous cell.

**Insight:**

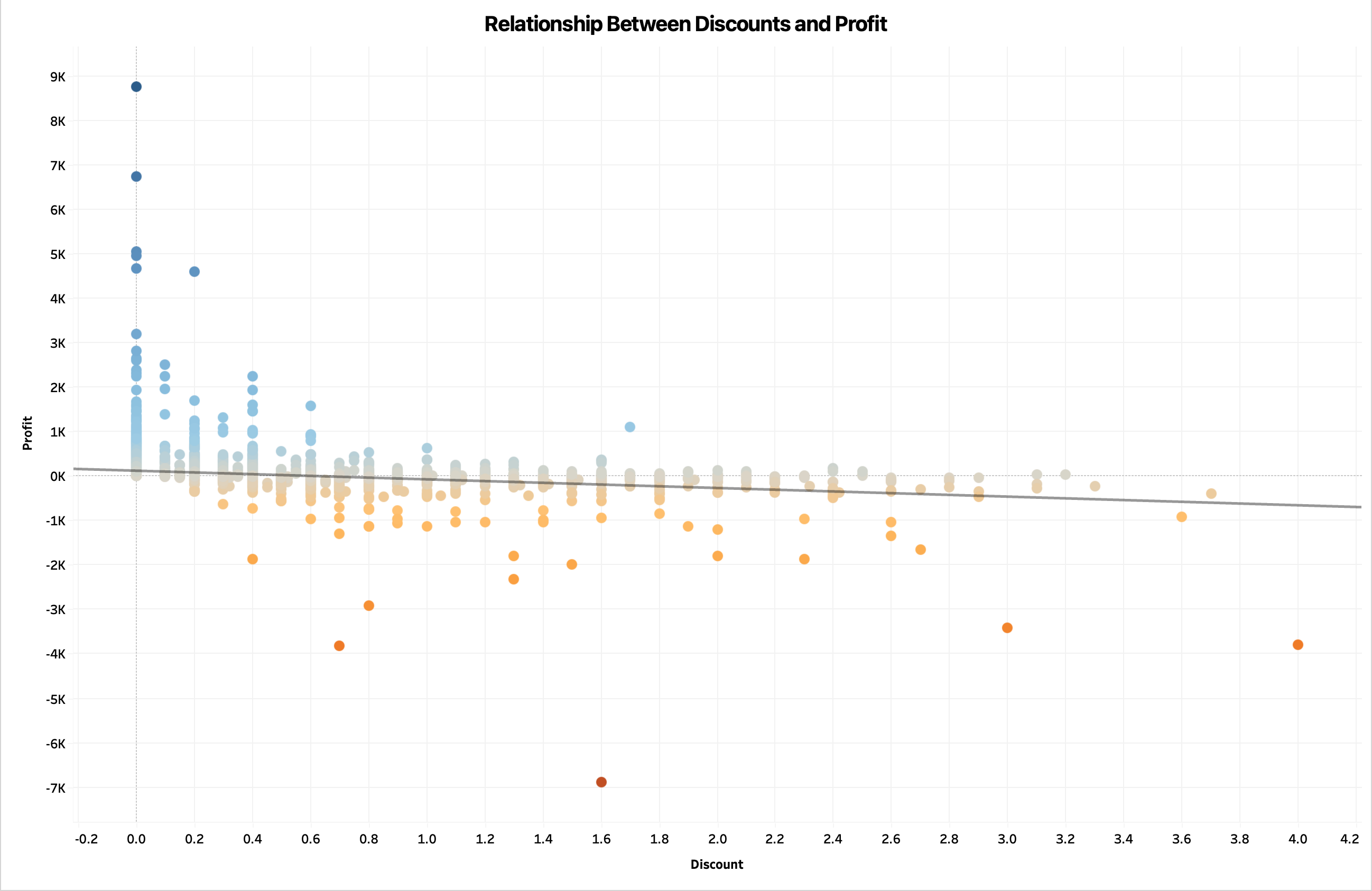
From the graph I can see that each categories generally have same

processing time for order. 0 to 7 represents the week days and processing time = Ship date – order date. Each box plot whisker has same 0 to 7 range and each categories have Q1, Q2(middle) and Q3 representing the higher whisker density inside box.

1. How do discounts affect overall profit?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. Can we visualise the relationship between product sales and profitability for different product categories?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the distribution of order quantities for products in the dataset?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. How do the profit distributions vary across different product categories?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. Can we compare the shipping time distributions for different shipping modes?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the monthly trend in the number of orders shipped?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. How do different customer segments perform in terms of sales and discount rates?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What are the sales and profit trends across different product subcategories and regions in the Superstore dataset?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the average delivery duration for different regions and ship modes?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. How has the average order quantity changed over the years for various product categories?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. Can we visualise the correlation between discount rates and order quantities for different customer segments?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the proportion of orders returned in each region within the Superstore dataset?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. Can you compare the profit of different products for different subcategories?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. Which shipping mode is the most commonly used in the Sample Superstore dataset?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. How does the sales performance of different regions evolve throughout the quarters of a year?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the distribution of order priorities across different product categories?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the relationship between discounts and sales?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. How does the average order value differ between repeat customers and new customers?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**

1. What is the geographical distribution of returns and its impact on overall profitability?

Ans.

**Chart:**

**Why \_ chart?**

**Insight:**