

# **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 visualization and analysis. Your objective is to select the best chart for each question, explain your choice. This project will showcase your proficiency in data visualization, critical thinking, and effective communication.

## **Skills Required:**

- Proficiency in data visualization concepts and techniques.
- Familiarity with Tableau or a similar data visualization 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 charts 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 visualization quality.
- Enhancements that contribute to better understanding or engagement.

## **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 visualization 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, organized, 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 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.

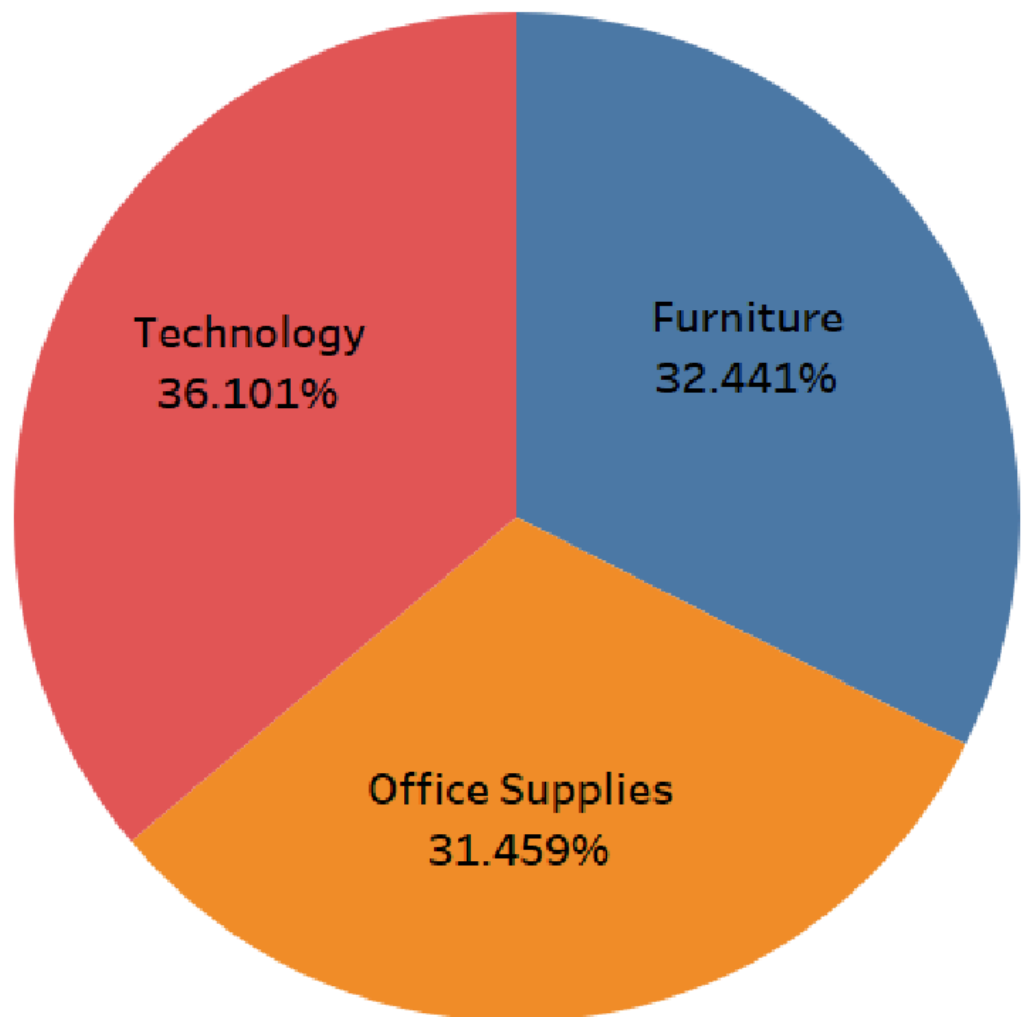
By completing this task effectively, you'll not only demonstrate your proficiency in data visualization and analysis but also showcase your ability to effectively communicate complex concepts through both text and charts.

**Good luck!**

**Questions:**

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

**<Product Category with highest total sales>**

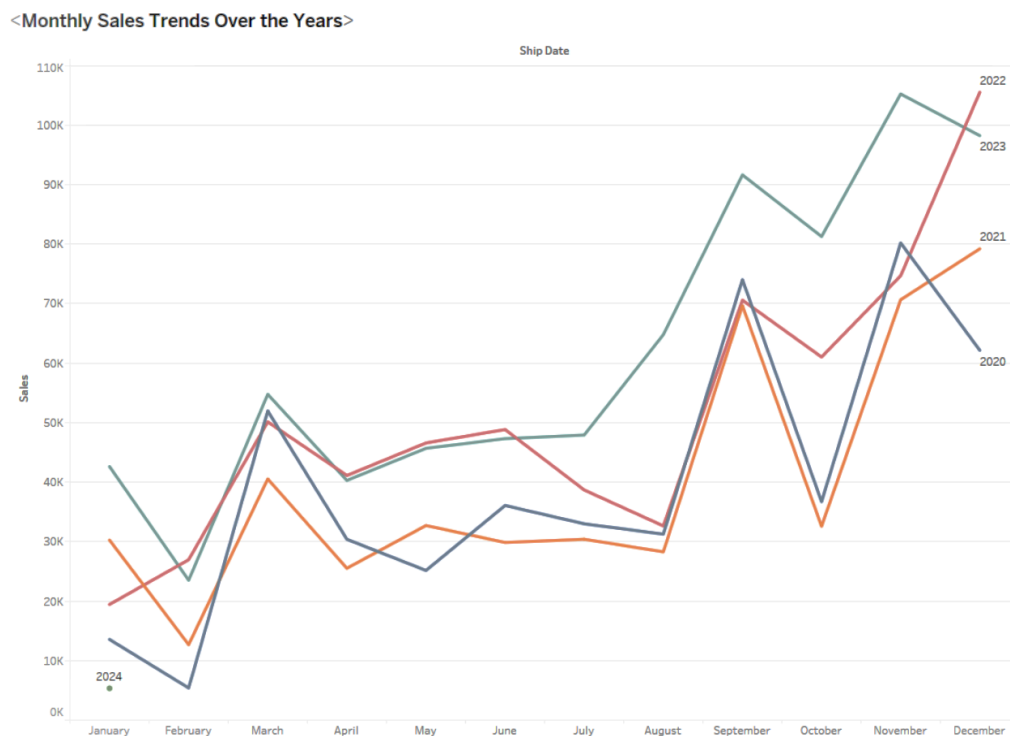


According to the pie chart , it can be seen that the 'Technology' category has the highest total sales of 36.101%.

This chart has been used to identify the product categories with the highest total sales in the "Superstore" dataset, the pie chart is used to visually compare the total sales across different categories. A pie chart is an effective choice as it allows for a quick comparison of sales

values between categories.

## 2. How do the monthly sales amounts change over the course of year?

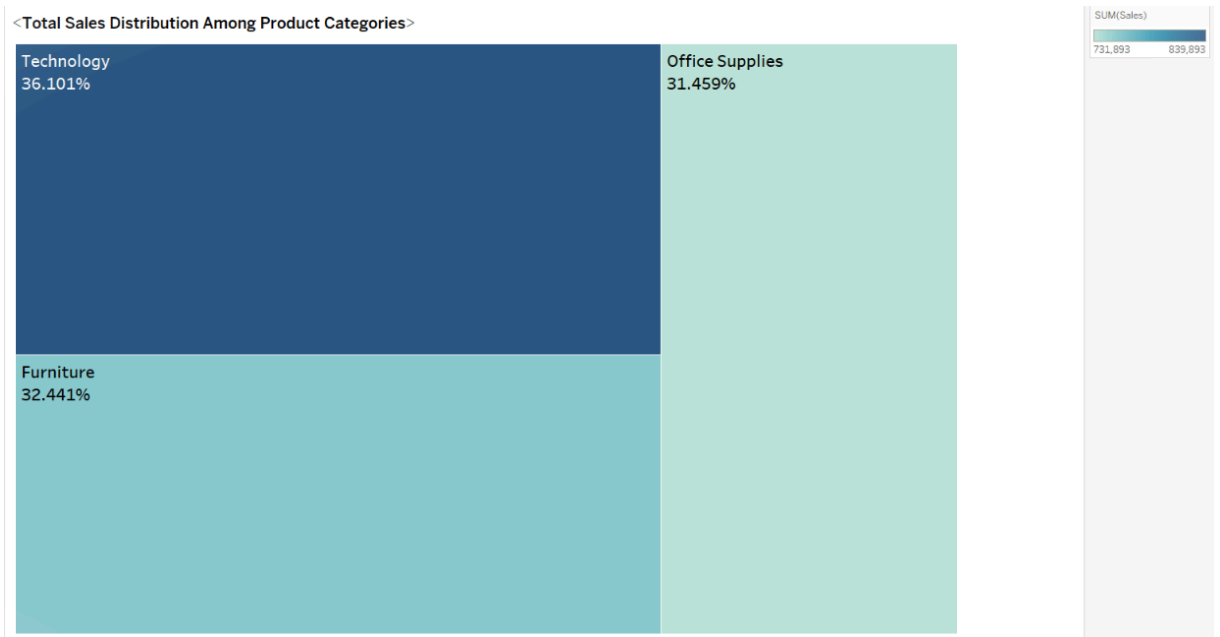


- **Seasonal Variations** : It can be seen that over the past 4 years the sales tend to peak from February to March, then it goes stagnant from April to August and then with continuous peak in trend from August to the end of the year.
- This trend tends to follow the same pattern for 4 years but with an increase every year.

This chart has been used to analyze the change in monthly sales amounts over the years, A line chart in Tableau has been chosen for its ability to display trends and variations in sales over time.

## 3. How is the total sales amount distributed among different product

## categories?

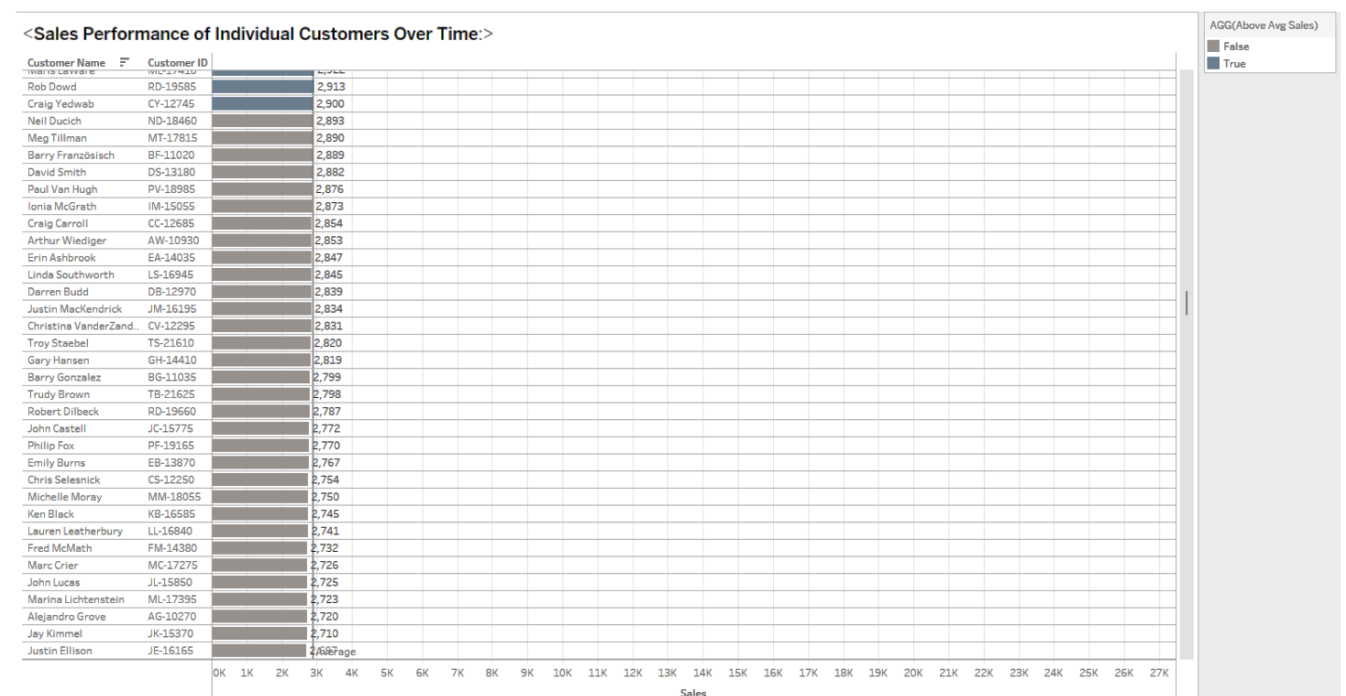
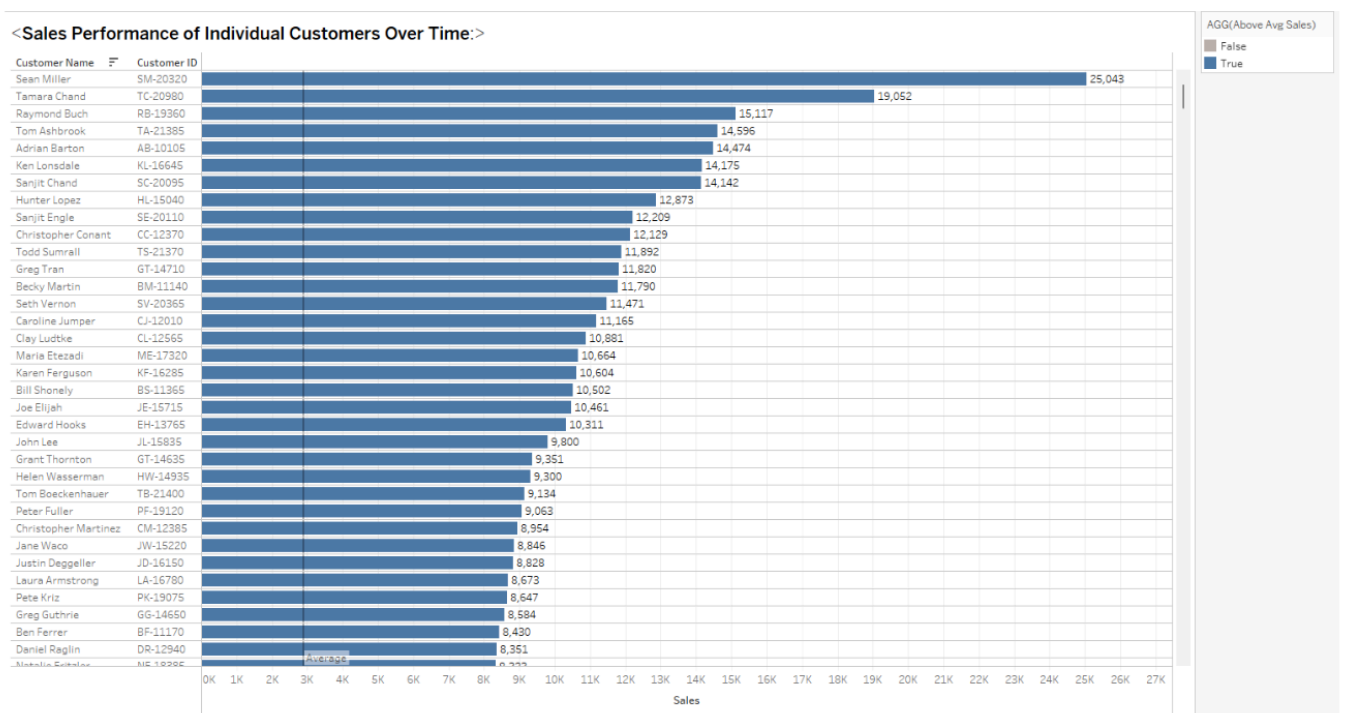


### Insights from the Total Sales Amount Distribution :-

- **Technology (36.1%)** : This category appears to contribute the most to total sales.
- **Office Supplies (31.4%)** : It contributes a significant share but slightly less compared to Technology.
- **Furniture (32.4%)** : While substantial, it shows a slightly lower percentage compared to Office Supplies but higher than Technology.

By visualizing the total sales amount distribution among different product categories using the heatmap, it's easier to grasp the proportional contributions of each category to the overall sales. The insights derived can assist in identifying focus areas for marketing strategies or investment decisions within these product categories.

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



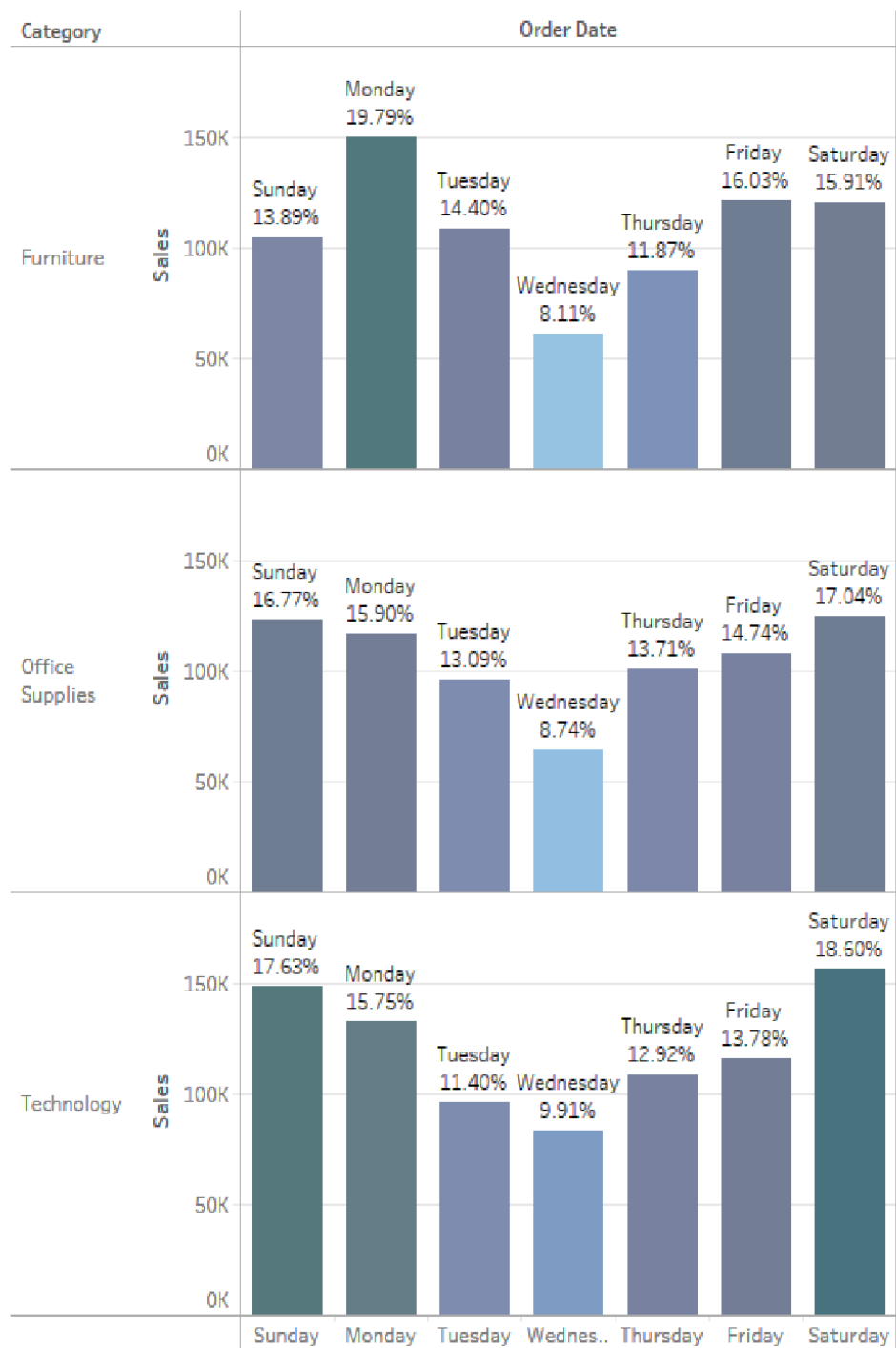
Insights from the Horizontal Bar Chart :-

- **Customer Sales Performance** : Assess individual customer sales by observing the lengths of horizontal bars. Longer bars represent higher sales.
- **Identifying Top Customers** : Identify top-performing customers by their bar lengths, placed at the top of the chart.
- **Sales Trends Over Time** : If segmented, analyze changes in sales performance trends across various time intervals (yearly, monthly) for each customer.

This horizontal bar chart allows for a clear visualization of individual customer sales performance over time, helping to identify top customers and understand their purchasing patterns. Insights derived from this visualization aid in customer segmentation, targeted marketing strategies, and identifying areas for customer engagement and retention.

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

### <Sales Variation Based on Days of the Week and Product Categories>



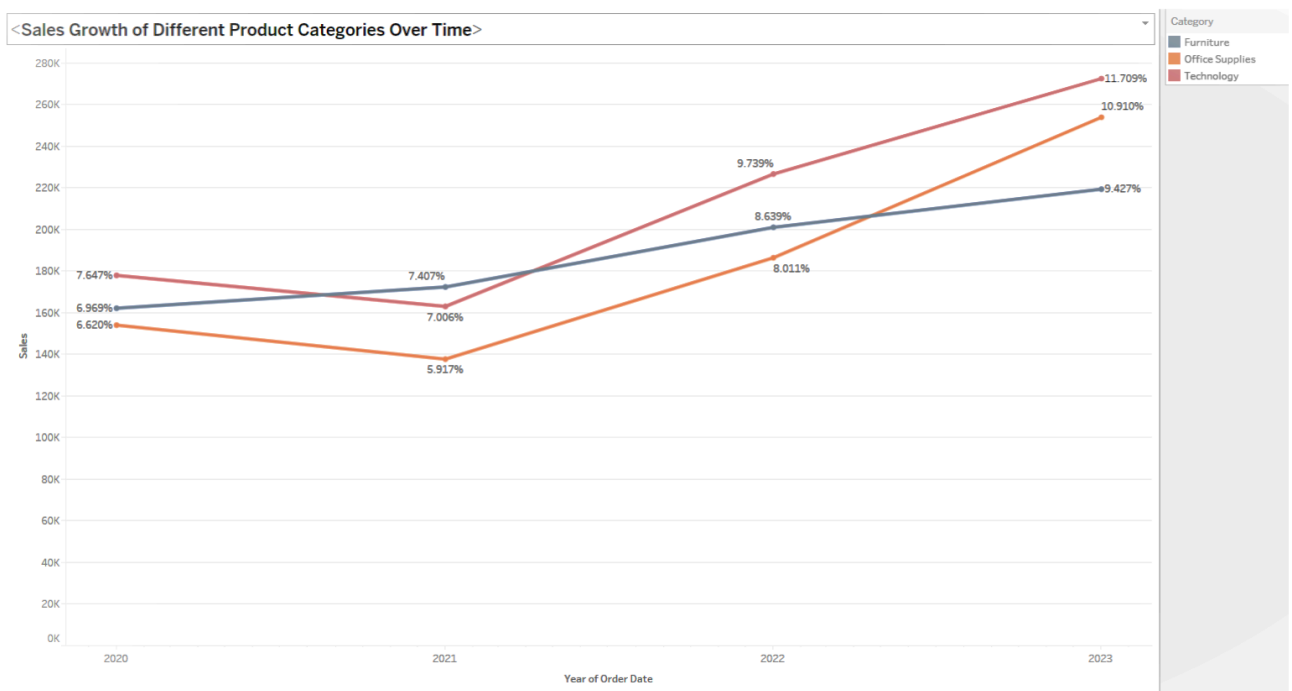
- **Sales Peaking at Week's Start and End** : Notice higher sales volumes at the beginning (Monday) and end (Friday and Saturday) of the week across different product categories.
- **Weekday Influence on Sales** : Identify whether certain categories consistently perform better on particular days or exhibit patterns based on weekdays.



- **Impact on Inventory and Marketing** : Insights help in optimizing inventory management and devising targeted marketing strategies aligned with specific days and product categories.

This visualization approach allows for a comprehensive understanding of how sales vary across weekdays for each product category. Insights gleaned can assist in devising targeted promotions, adjusting inventory levels, and understanding customer preferences based on specific days of the week and product categories.

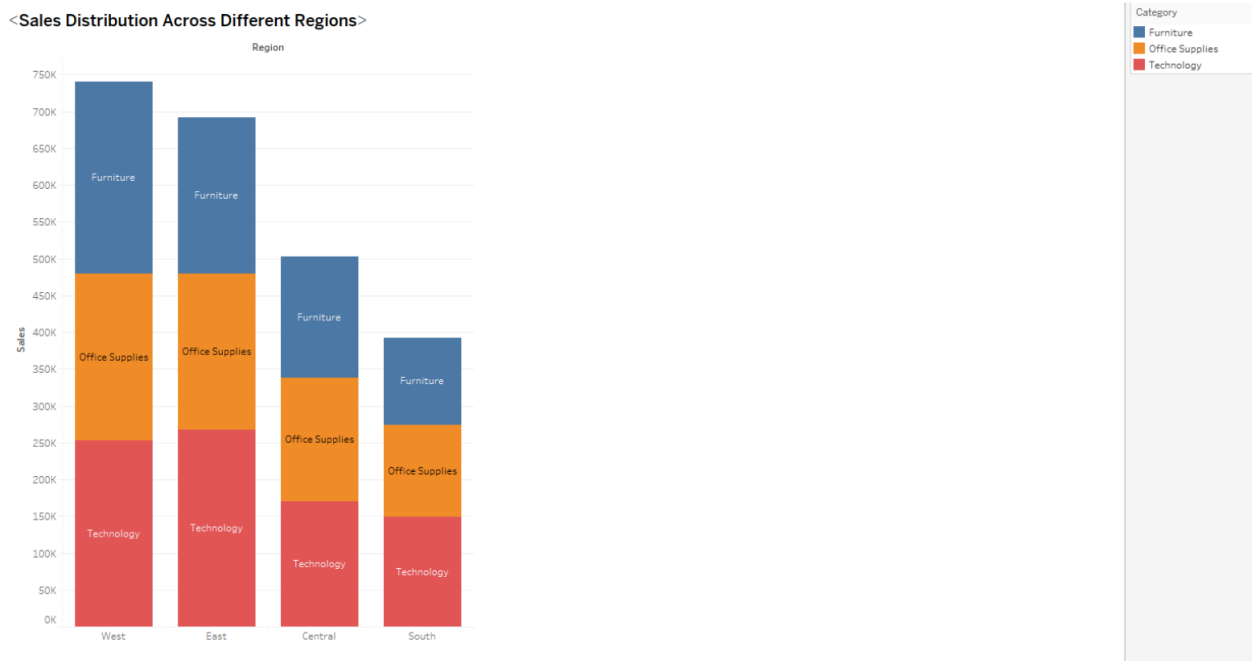
## 6. Can we visualize the sales growth of different product categories over time?



- **Sales Growth Comparison** : We can see the growth trajectories of different product categories represented by distinct lines on the chart.
- **Identifying Trends** : The technology and office supplies have tended to show more growth over the years, whereas the growth of furniture is not very high.
- **Highlighting Patterns** : There is a relative performance of each category, particularly focusing on technology and office supplies, which have demonstrated higher sales growth compared to furniture.
- **Strategic Decision-Making** : Insights derived can inform strategic decisions, such as allocating resources, focusing on specific product categories, or adjusting marketing strategies based on sales growth patterns.

The line chart effectively illustrates the sales growth trends of various product categories over time. Insights obtained help in identifying the performance of different categories and making informed business decisions to leverage or improve sales performance.

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

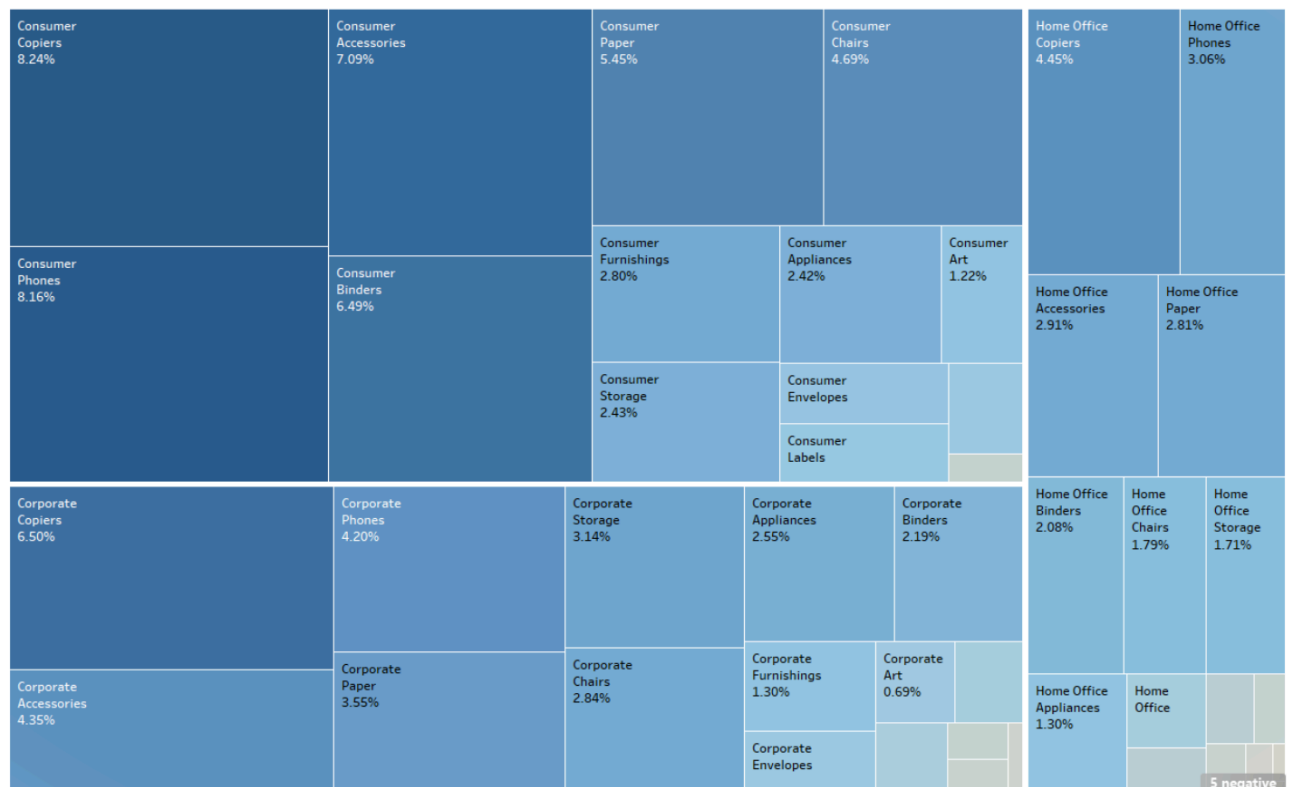


- **Regional Sales Distribution** : The chart visually showcases how sales are distributed across product categories within each region.
- **Regional Comparisons** : By comparing the heights of stacked bars for different regions, insights into sales dominance or underperformance in specific regions (such as West and East) can be gained.
- **Identifying Top Categories** : Technology is the category that contributes the majority of the sales in all the region which is followed by office supplies. However furniture tends to perform very well in the west region

The stacked bar chart effectively represents sales distribution across different regions, showcasing the dominance of specific product categories in each region. Insights obtained can guide decisions on regional targeting, marketing strategies, or product assortment to maximize sales potential in specific areas.

## 8. Can we visualize the composition of profits across various subcategories within different customer segments?

<Composition of Profits Across Customer Segments and Subcategories>



- **Profit Composition** : The tree map illustrates the breakdown of profits across subcategories within the main categories (technology, office supplies, furniture) and further categorizes profits based on different customer segments.
- **Relative Profit Contribution** : Larger rectangles indicate higher profits within each main category, offering an immediate visual understanding of which subcategories contribute most significantly to profits.
- **Segment-wise Comparison** : The color-coded segments help compare profit distribution across different customer segments within each subcategory, revealing which segments are more profitable for specific subcategories.

The tree map effectively presents the profit composition across various subcategories within different customer segments. Insights derived from this visualization can assist in identifying profitable product segments, understanding customer behavior, and optimizing marketing strategies.

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

<Percentage Contribution of Each Region to Overall Sales>



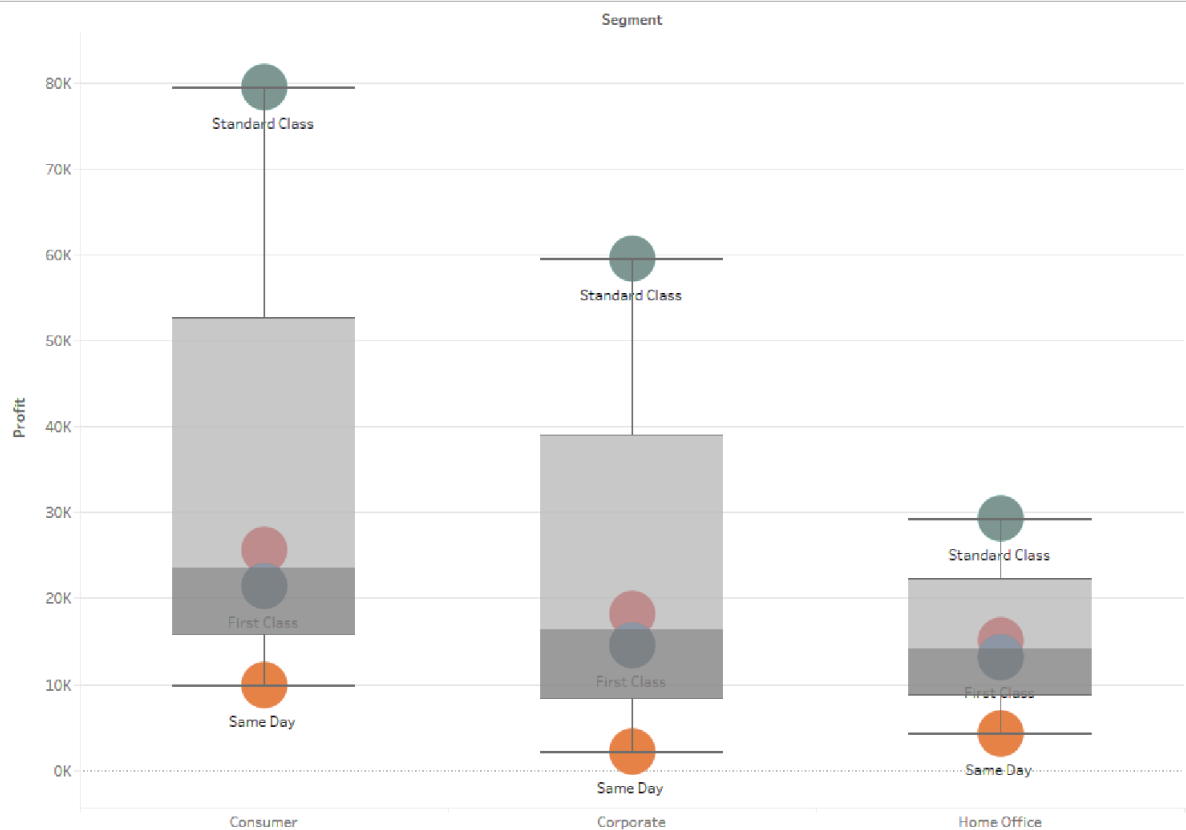
- **Sales Contribution by Region** : The map graph visually represents the percentage contribution of each region to the overall sales. Darker regions indicate higher sales contributions.
- **Regional Disparities** : Clear distinctions are observed between regions based on their sales shares, highlighting the significance of each region in the total sales figure.
- **Major Contributors** : Insights reveal the dominant sales-contributing regions. For instance, if the South region of the US exhibits the highest sales percentage, followed by Canada, it signifies the importance of these regions in generating revenue for the Superstore.

Overall, the map graph effectively illustrates the percentage contribution of each region to the overall sales, emphasizing the sales distribution across different geographic areas. The insights derived from this visualization can guide targeted marketing efforts, resource

allocation, and regional business strategies.

## 10. Can we visualize the profit margins associated with different shipping modes and customer segments?

<Profit Margins Across Shipping Modes and Customer Segments>

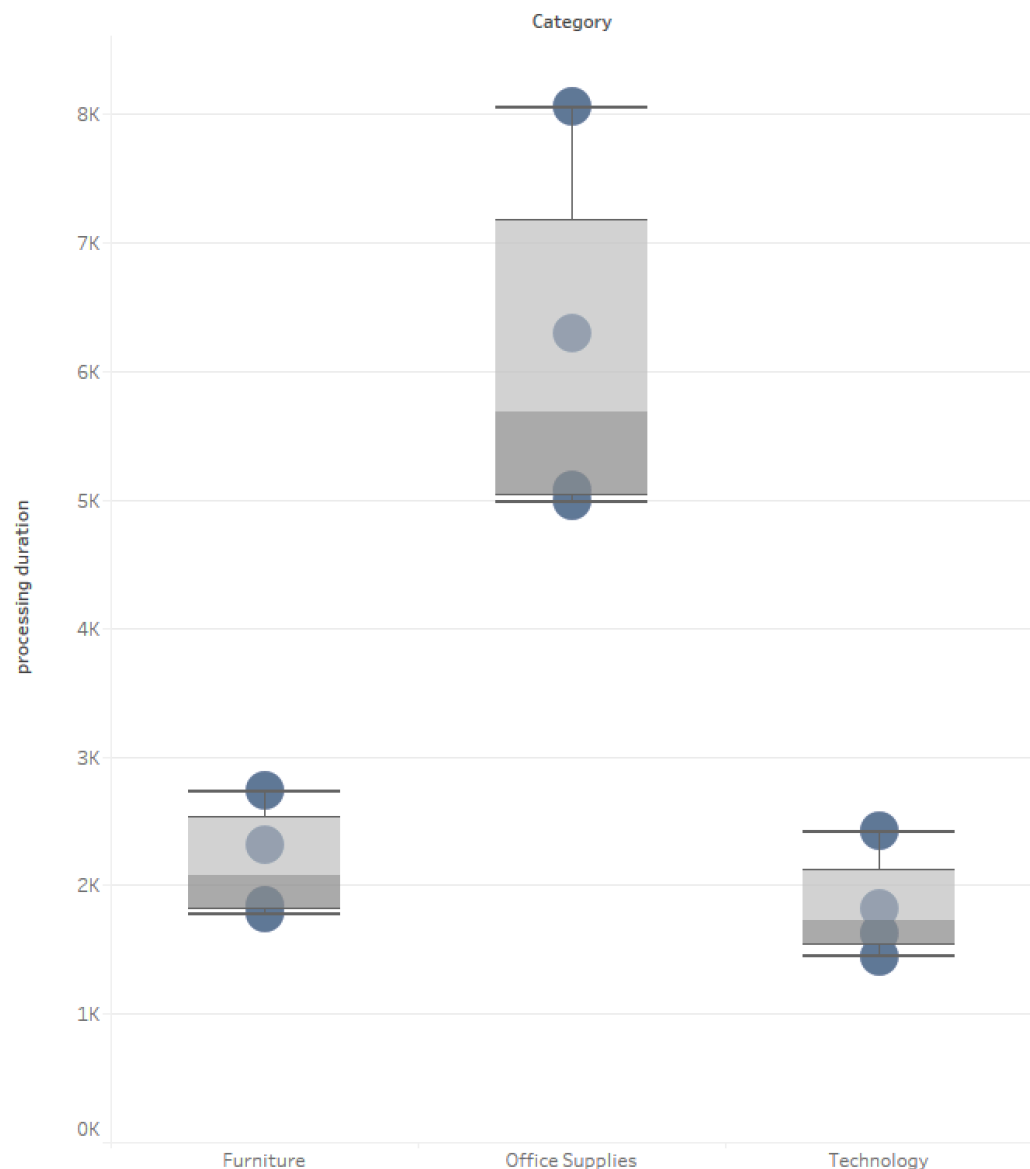


- **Profit Margin Comparison** : The box and whisker plot effectively showcase the distribution of profit margins across shipping modes and customer segments.
- **Variability in Profits** : the box plot enables the identification of differences in profit margin variability between shipping modes and customer segments .
- **Segment-Specific Profitability** : The plot reveals segment-specific profitability trends. The consumer segment demonstrates larger profits across various shipping modes, it indicates the consumer segment's significance in driving higher profits for the Superstore and all the segments prefer standard class as the mode of shipping.

Overall, the box and whisker plot visually represent the distribution and variability of profit margins associated with different shipping modes and customer segments. The insights derived from this visualization aid in understanding profit trends, identifying profitable segments, and making informed decisions regarding shipping strategies and customer targeting.

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

<Processing Time for Different Product Categories>



- **Processing Time Comparison** : The box plot effectively showcases the

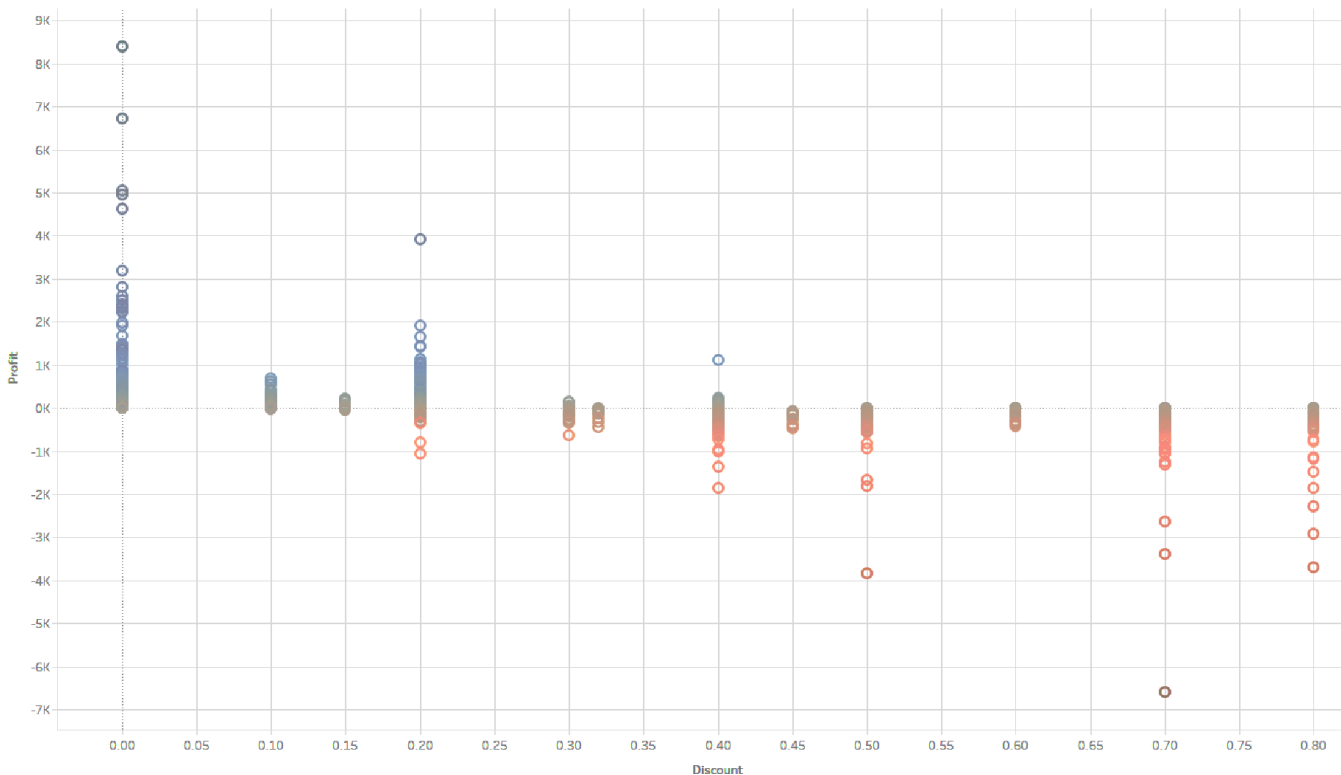
distribution of order processing times across different product categories.

- **Variability in Processing Times** : Analysis of the box plot enables the identification of differences in processing time variability between product categories.
- **Fastest Processing Category** : By observing the box plot, we can determine that the office supply category demonstrates the fastest order processing time compared to other categories. This insight is valuable for understanding efficiency in processing orders and identifying areas for improvement.

The box plot visually represents the distribution and variability of order processing times across various product categories. The insights derived from this visualization aid in understanding processing time trends, pinpointing categories with faster or slower processing times, and making informed decisions to optimize operational efficiency.

## 12. How do discounts affect overall profit?

<Effect of Discounts on Overall Profit>

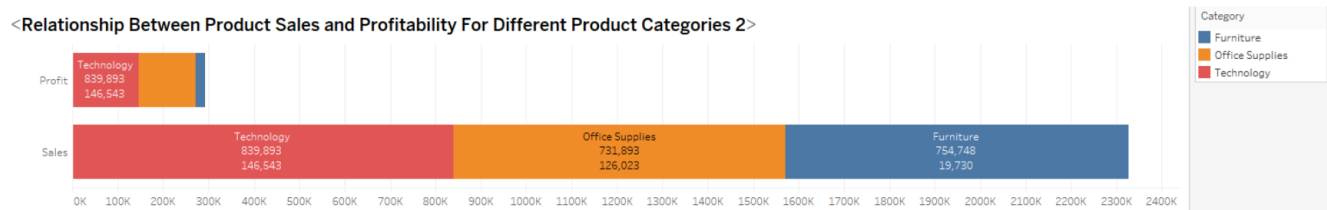
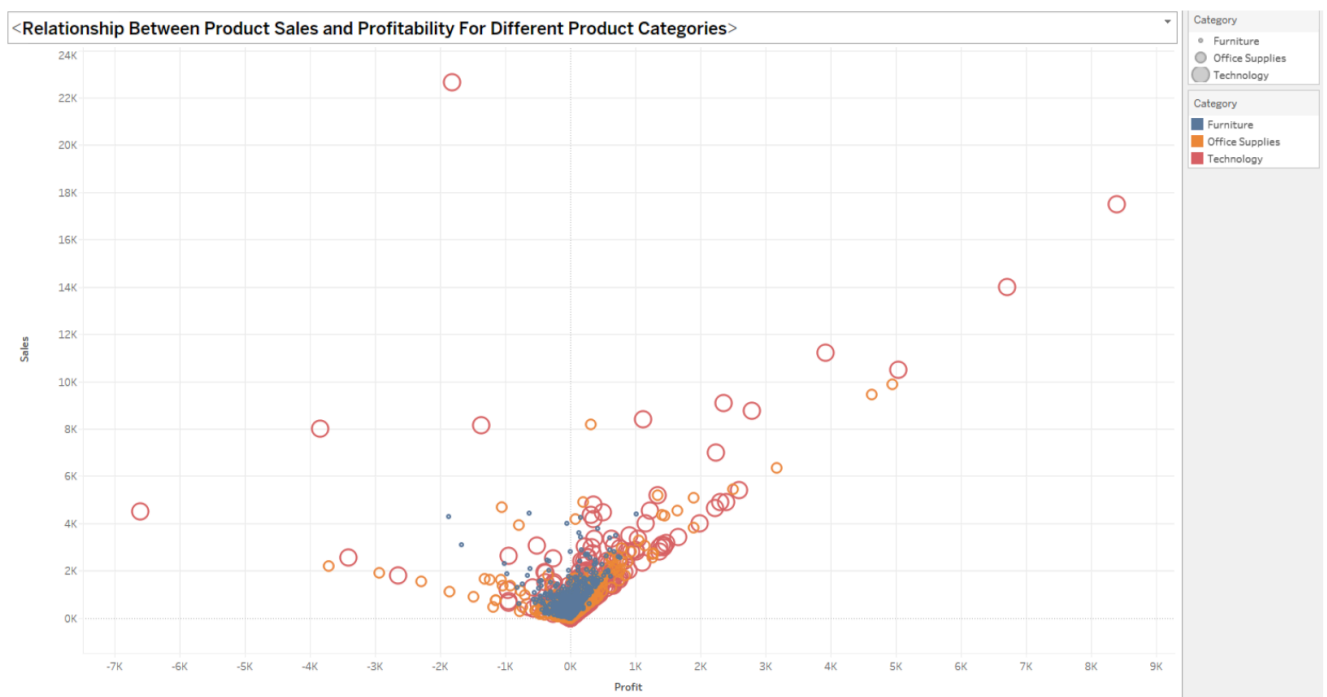


- **Impact of Discounts on Profit** : Analysis of the scatter plot provides insights into how discounts affect overall profit. Observing the distribution of data points can reveal trends or patterns indicating how different discount levels influence profit margins.

- **Negative Profit with Higher Discounts :** From the data in the scatter plot, it can be observed that transactions with higher discount rates often result in negative profits. This insight suggests that excessive discounts might lead to reduced profitability or even losses for certain transactions.

The scatter plot effectively visualizes the relationship between discounts and profits, highlighting the impact of varying discount rates on overall profitability. The insights gained from this visualization can aid in making informed decisions regarding discount strategies to optimize profitability while managing margins effectively.

### 13. Can we visualize the relationship between product sales and profitability for different product categories?



- **Correlation between Sales and Profitability :** Observing the scatter plot, if a significant number of points show low sales resulting in negative profitability, it suggests a potential negative correlation between sales and profitability for those specific

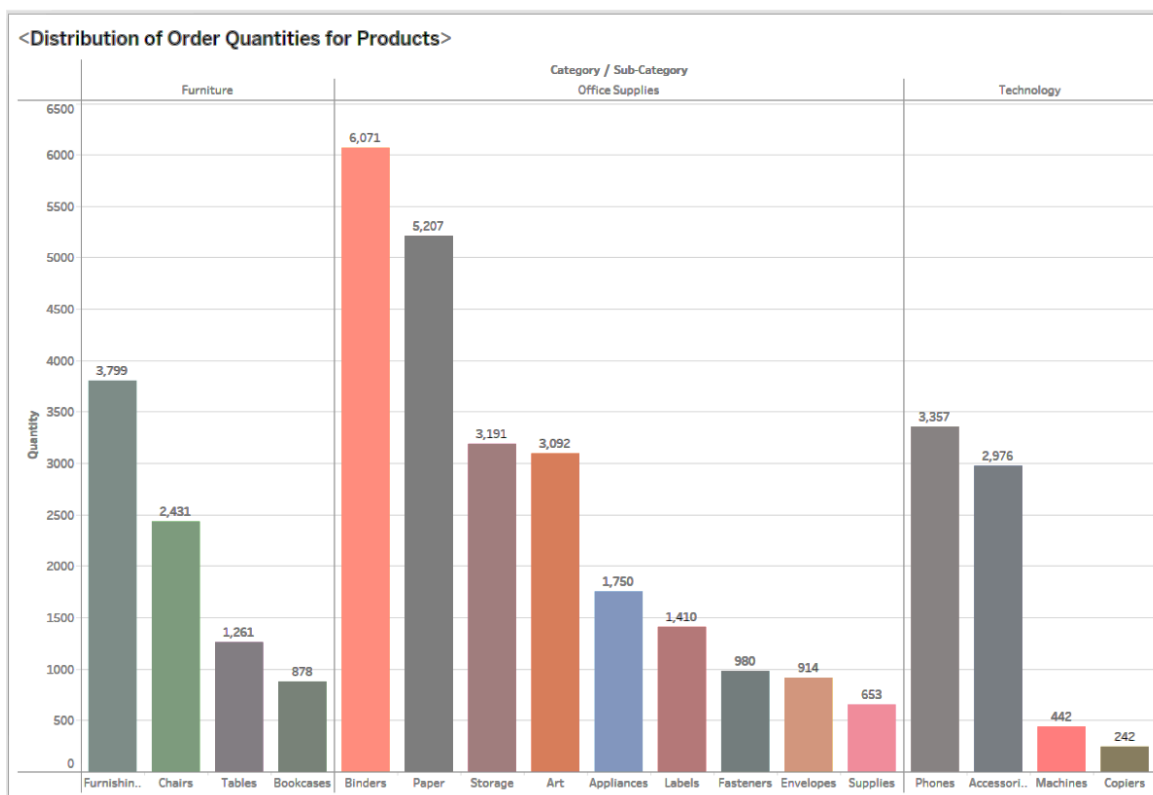


categories.

- **Profitability Composition** : From the stacked bar chart, insights into the proportion of positive and negative profits within each product category can be gained. Categories with more negative segments imply poor profitability despite considerable sales.

In summary, this dual visualization strategy effectively captures the relationship between sales and profitability for different product categories, revealing trends and insights about their performance. It allows stakeholders to identify categories that might need further analysis or strategy adjustments for enhanced profitability.

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

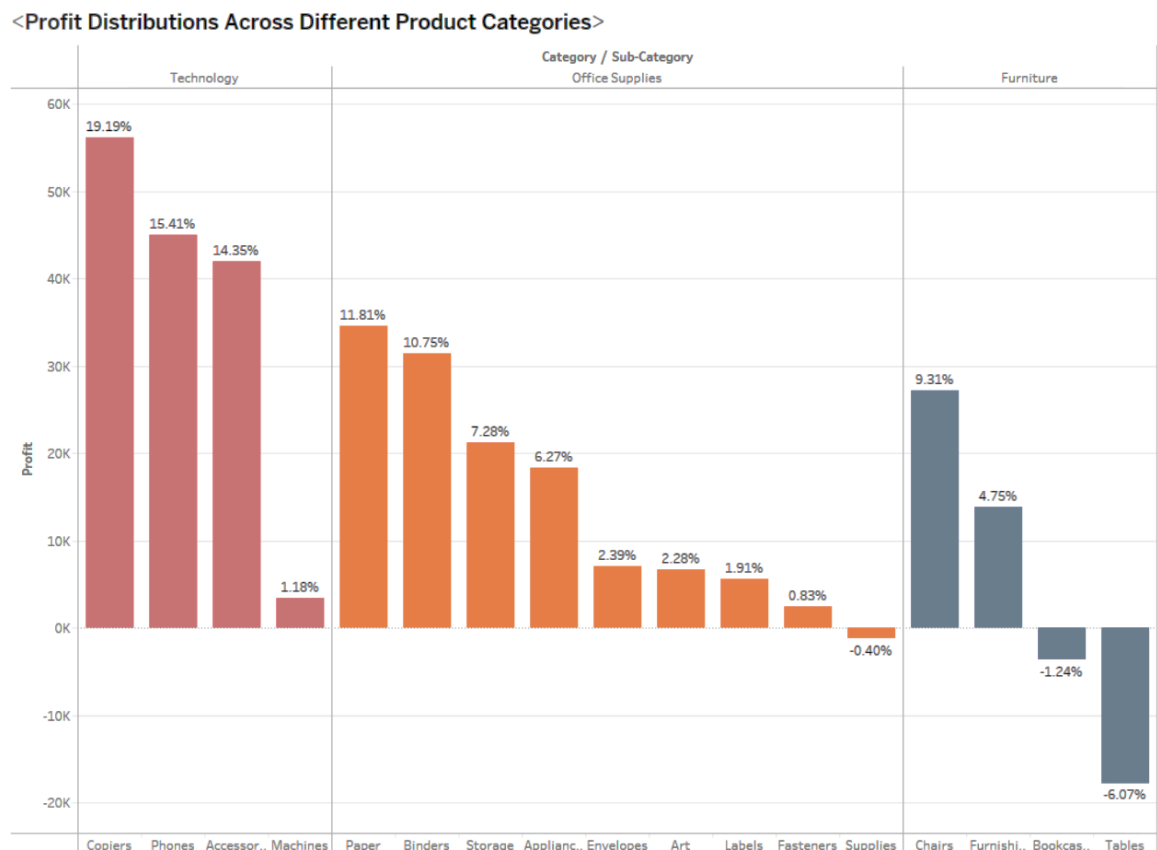


- **Category-Wise Order Quantities** : The chart would show each category's total order quantities, and the comparison between categories would indicate which category has higher overall orders.
- **Sub-Category Variation** : Within each category, observing the sub-category bars provides insights into the order quantities' variation. Office supplies and furniture might show higher variability or diverse ranges of products compared to other categories.

In summary, the grouped bar chart effectively showcases the distribution of order quantities across categories and their sub-categories in the Superstore dataset. It provides

insights into the varying scales of orders within different product groups, highlighting categories and subcategories with notable quantities and variations.

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

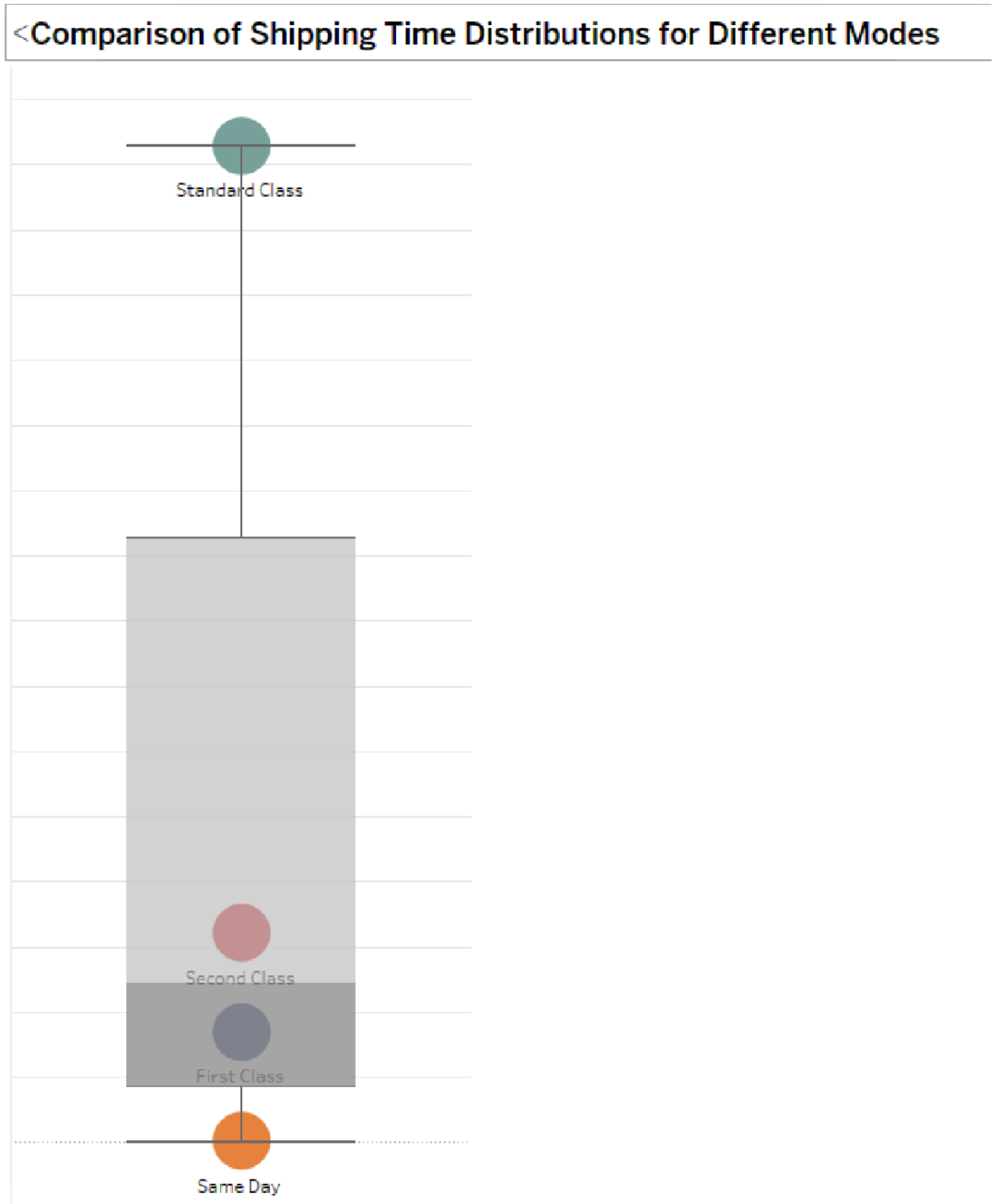


- Category-Wise Profit Analysis :** The chart would showcase the profit distribution across different product categories. It would highlight categories with the highest overall profits, such as technology and office supplies, and categories with potentially negative profits, such as certain products in the furniture category (like bookcase and tables).
- Product-Level Profits :** Within each category, the comparison between products' bars indicates which products are contributing positively or negatively to the overall profit. For instance, specific products within the furniture category might show negative profits, impacting the category's overall profitability.

In summary, the grouped bar chart effectively represents the profit distributions across different product categories, highlighting categories with high profits and specific products

within categories that might be affecting profitability negatively.

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



- **Comparison of Shipping Time:** The box and whisker plot will demonstrate the differences in shipping times across different shipping modes. It would clearly indicate if the standard class indeed has higher shipping times compared to other classes.

In summary, the box and whisker plot effectively visualize the distribution of shipping times for different shipping modes, facilitating a clear comparison and providing insights into the variability and central tendency of delivery durations across various shipping classes.

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

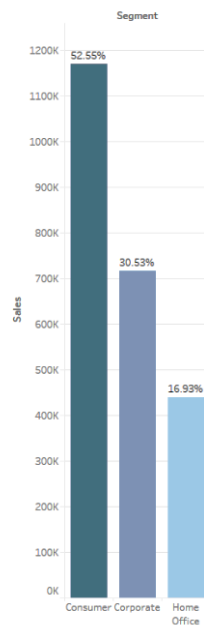


- **Identifying Seasonal Trends :** There's a repetitive pattern seen during specific months (such as the observed similarity from March to November), it might indicate a seasonal trend or cyclic behavior in order shipments during those months.

The line chart would effectively present the monthly trend in the number of orders shipped, allowing for the identification of any regular patterns or trends, and potentially highlighting any specific months with consistent behavior or significant changes in order shipments.

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

<Sales and Discount Rates Across Customer Segments>

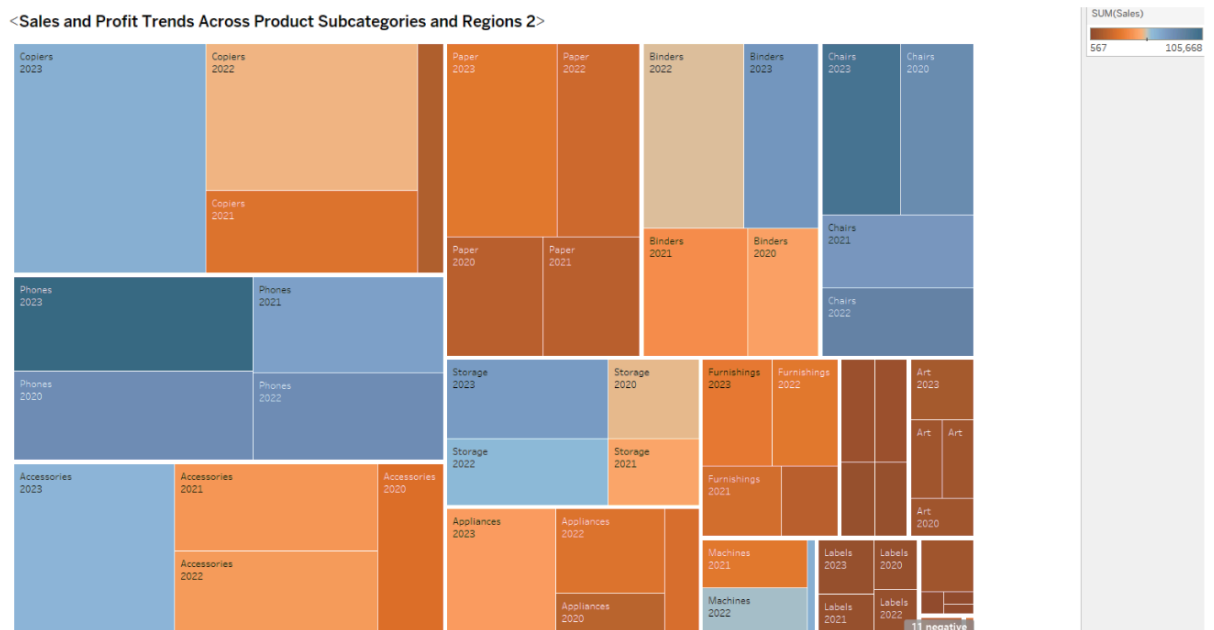
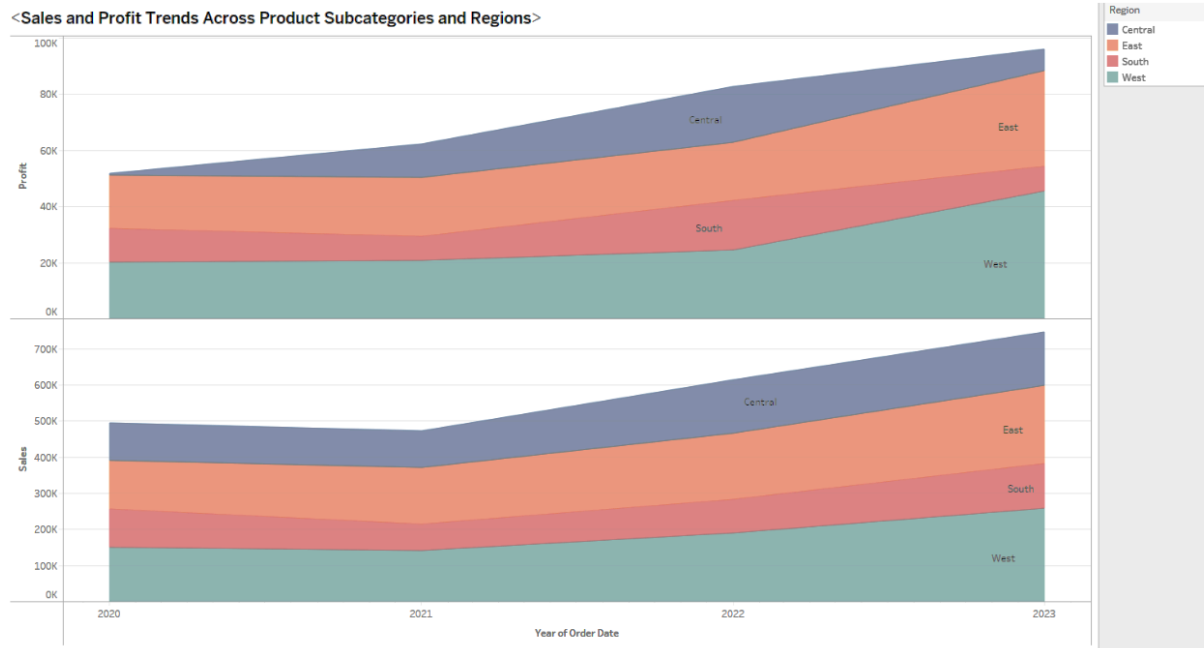


SUM(Discount)	
268.1	832.3

- **Sales Performance** : The chart illustrates the relative performance of different customer segments in terms of sales. For instance, if the consumer category appears with higher bars, it indicates higher sales compared to other segments.
- **Discount Rates** : The darker color in the chart corresponds to a higher discount rate. Thus, darker segments indicate higher discounts applied within a particular customer segment.

The chosen bar chart effectively presents a comparison between different customer segments concerning their sales performance and discount rates. It allows for a visual comparison to identify segments with higher sales and significant discount rates, providing insights into each segment's performance.

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

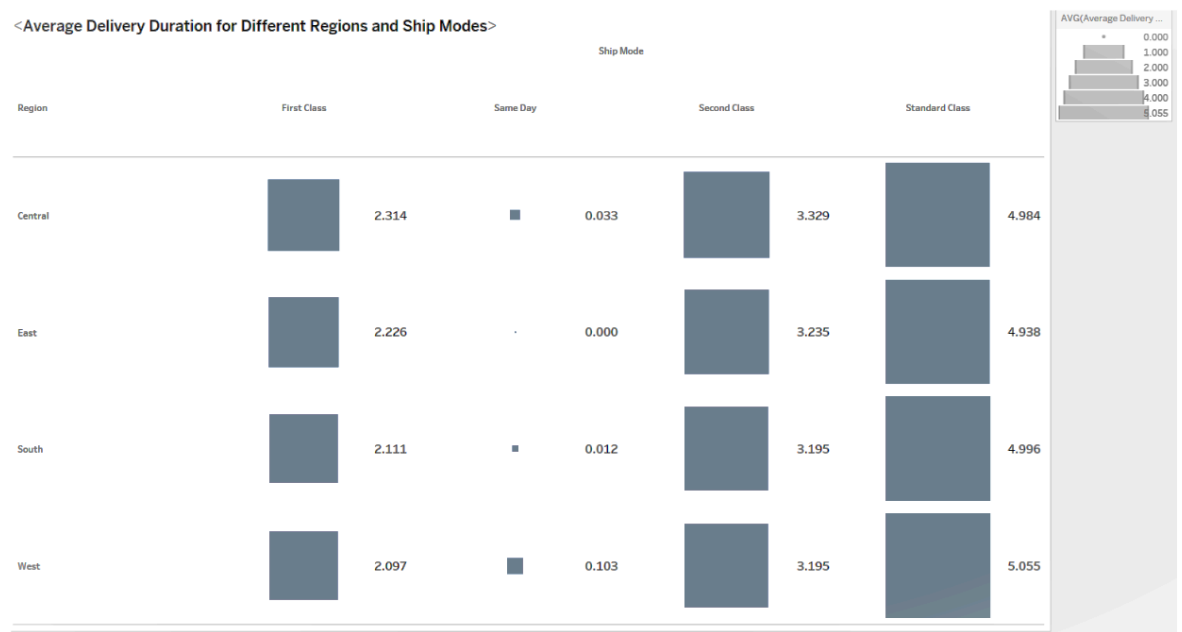


- Area Chart Insights :** The area chart will display trends over the years for both sales and profit across different regions. Rising areas in the chart represent growing sales and profit trends, while declining areas denote decreasing trends.
- Treemap Insights :** The treemap will illustrate the contribution of each product subcategory to sales and profit. Rectangles with blue or bluer colors indicate

subcategories with a healthy ratio of profit to sales, signifying good performance. Conversely, red or warmer-colored rectangles denote subcategories where profit is not proportional to sales, highlighting areas needing improvement.

In summary, utilizing an area chart and a treemap in Tableau effectively presents insights into the trends of sales and profit across regions over time, as well as the performance of different product subcategories in terms of their contribution to sales and profit.

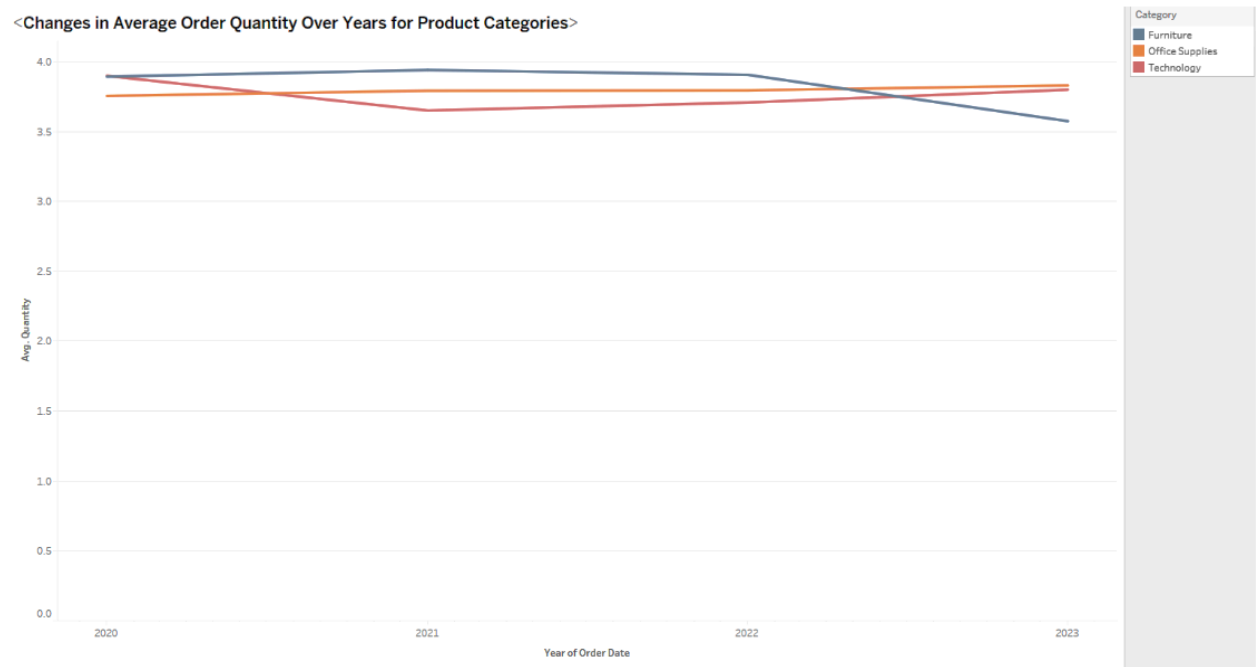
## 20. What is the average delivery duration for different regions and ship modes?



- **Heatmap Insights** : The heatmap will provide an instant overview of the average delivery duration across different regions and shipping modes. larger boxes signify longer average delivery times, while smaller boxes indicate shorter average delivery times.
- Insights can be drawn by observing the differences in box sizes. For instance, if certain regions or shipping modes consistently appear as larger boxes, it suggests longer average delivery durations in those areas or for those modes.

Utilizing a heatmap in Tableau effectively presents insights into the average delivery durations across various regions and shipping modes, allowing easy identification of areas or modes with longer or shorter delivery times.

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

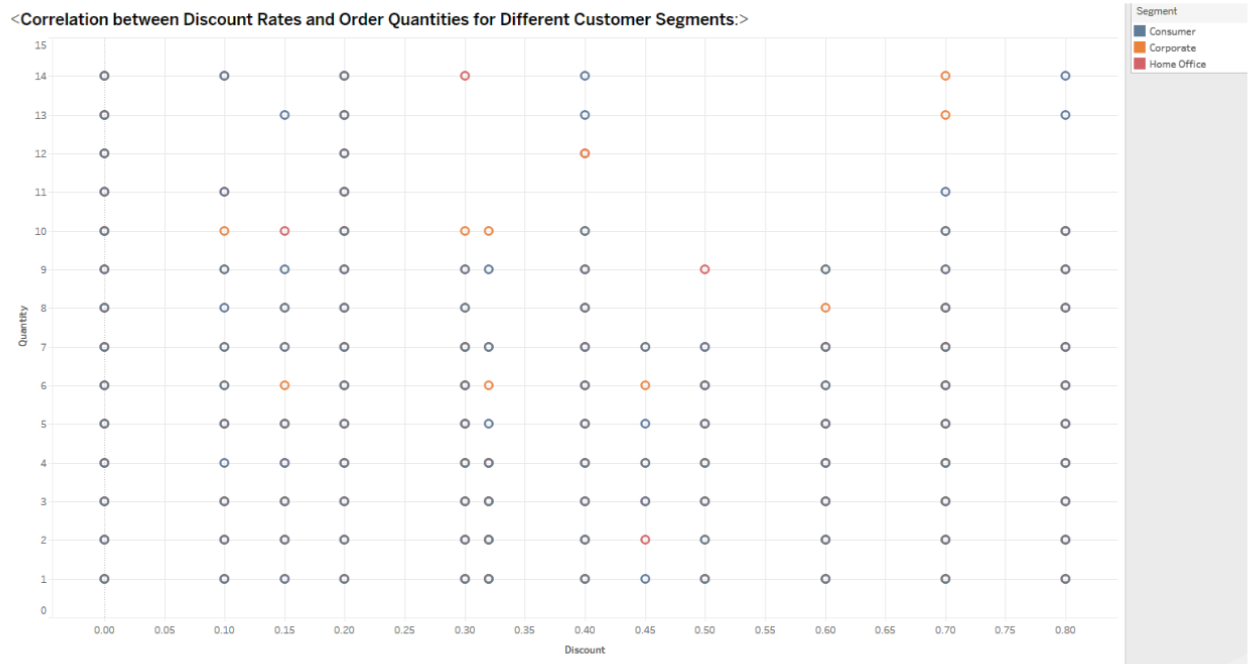


- **Line Chart Insights** : By analyzing the lines, one can easily discern the trend of average order quantity changes for different product categories over the years.
- The declining trend in the line representing furniture indicates a decreasing average order quantity for furniture products over the years.

Using a line chart in Tableau effectively presents the change in average order quantity over the years for various product categories, allowing easy comparison and identification of trends, such as the decline in average order quantity for furniture over time in the Superstore dataset.

## 22. Can we visualize the correlation between discount rates and order quantities for different customer segments?



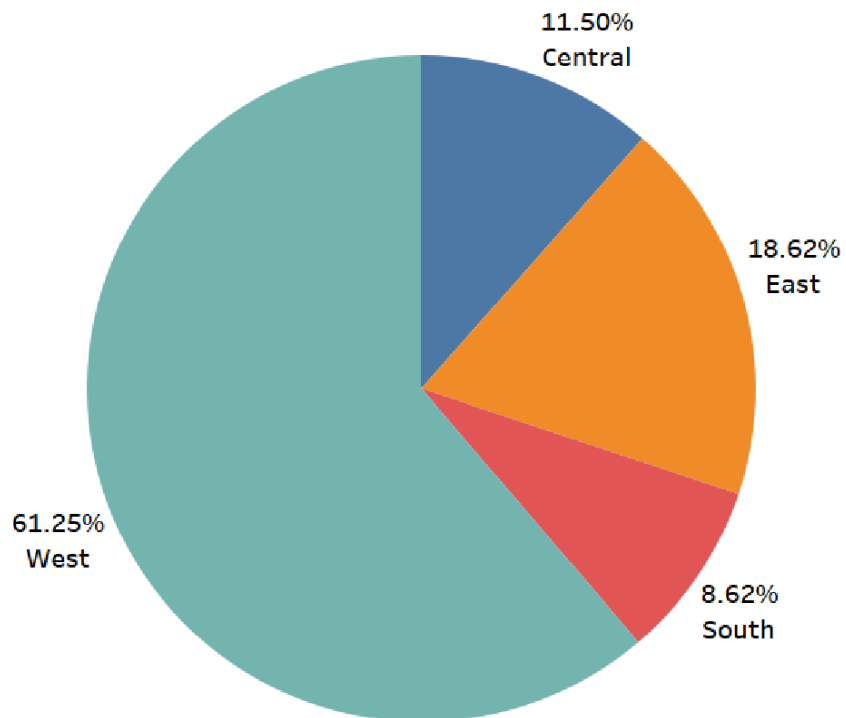


- **Correlation Observation :** Analyzing the scatter plot allows for an understanding of how discount rates impact order quantities across different customer segments.
- The points clustering in a positive slope indicates a positive correlation, implying that higher discounts correlate with higher order quantities.

For instance, in the Superstore dataset, if the scatter plot indicates that the consumer segment receives higher discounts resulting in higher order quantities compared to other segments, this information can be vital for strategic decision-making in sales and marketing efforts.

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

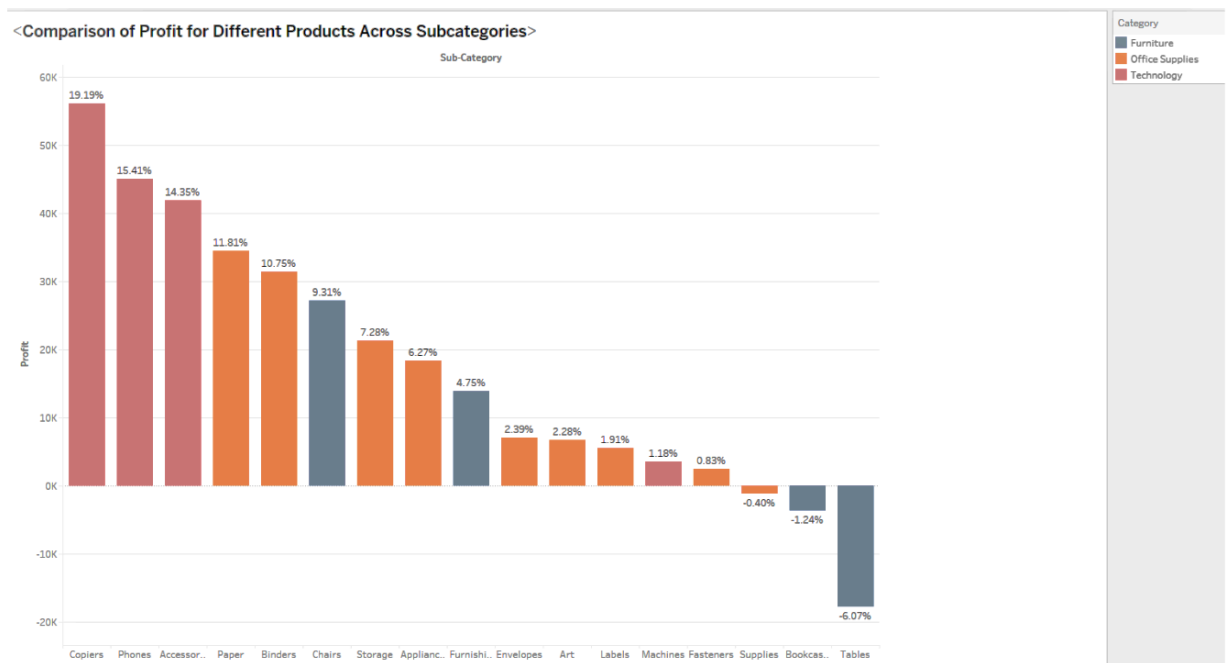
<Proportion of Orders Returned in Each Region>



- **Return Percentage Observations :** Pie chart visualization helps in quickly understanding the contribution of each region to the total returned orders.
- Analyzing the pie chart might reveal that the West region has the highest proportion of returns, followed by East, Central, and South. This insight can be crucial for identifying regions that need specific attention or improvement in terms of product quality, delivery, or customer service to reduce returns and enhance customer satisfaction.

Understanding these percentages can aid in strategizing better customer service initiatives, optimizing logistics, and refining product quality or descriptions to minimize returns in regions with higher return percentages.

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

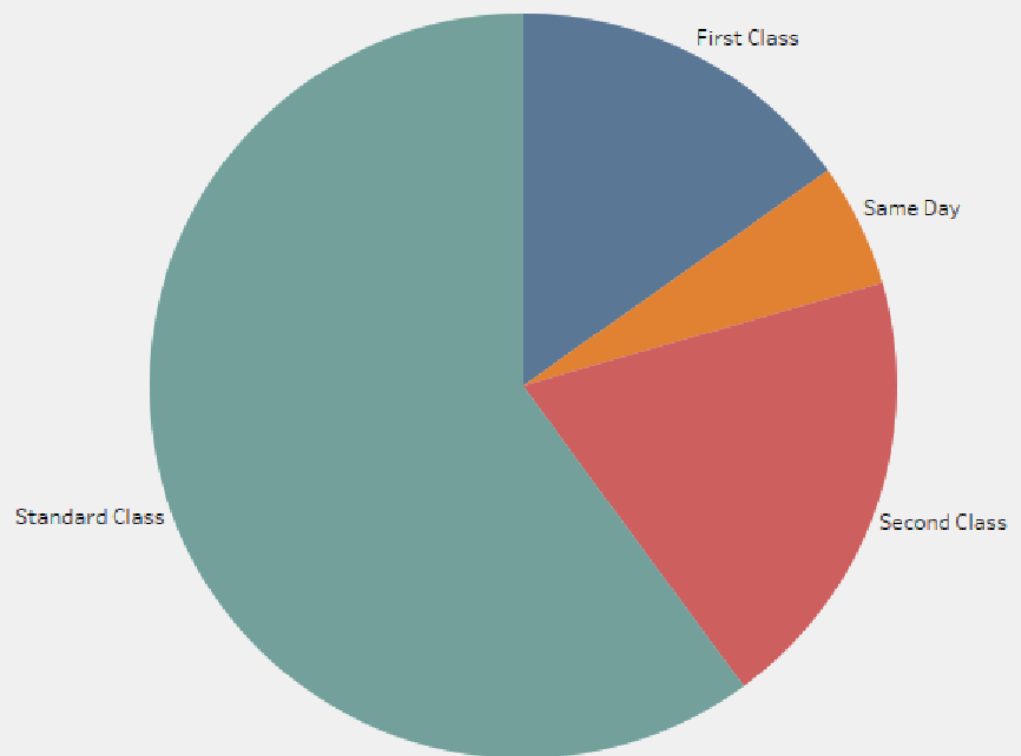


- **Profit Comparison** : By examining the bars, one can easily discern which products within each subcategory are generating higher profits and which ones are incurring losses.

From the visualization, it might be observed that products like bookcases and tables are incurring negative profits, represented by downward bars. This insight can be crucial for further analysis and decision-making, such as considering pricing adjustments, marketing strategies, or product improvements for these specific items to enhance profitability.

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

### <Most Commonly Used Shipping Mode>

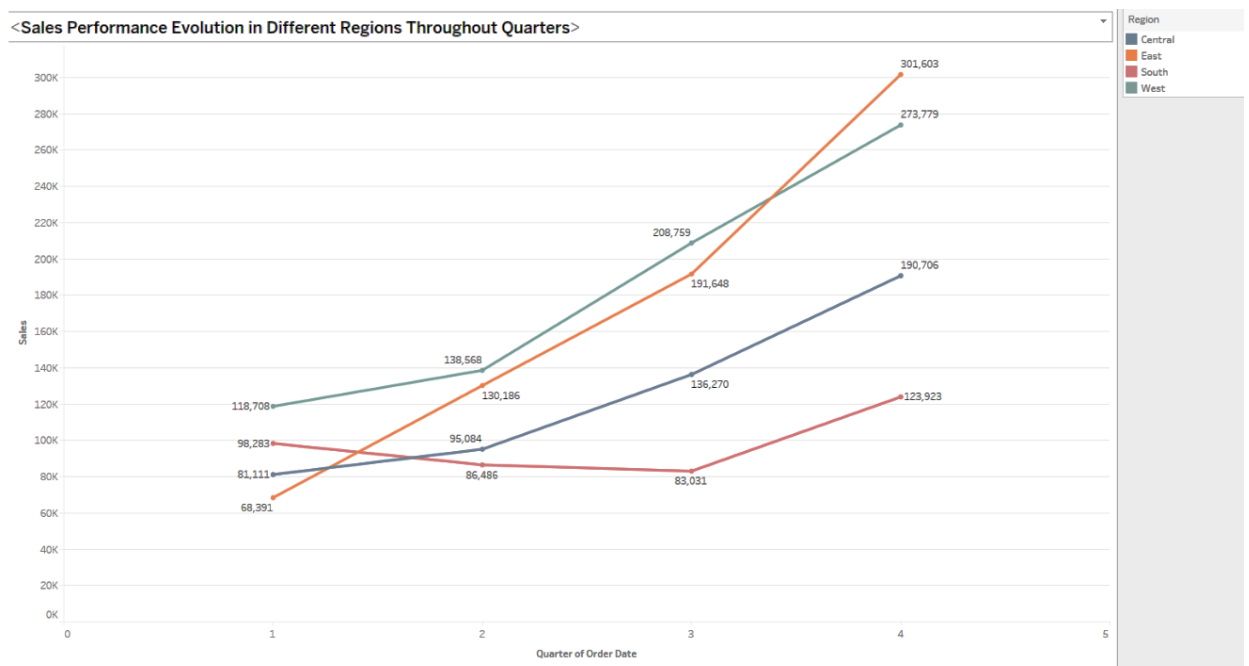


- **Usage Comparison** : The pie chart provides a clear visualization of the percentage or count of each shipping mode. In this scenario, it's apparent that the 'Standard Class' shipping mode is most commonly used, as it occupies the largest section of the pie.
- The 'Second Class' shipping mode follows as the second most popular choice.

This insight is crucial for analyzing shipping preferences, aiding decisions related to

inventory management, logistics, or optimizing shipping strategies for different modes to enhance customer satisfaction.

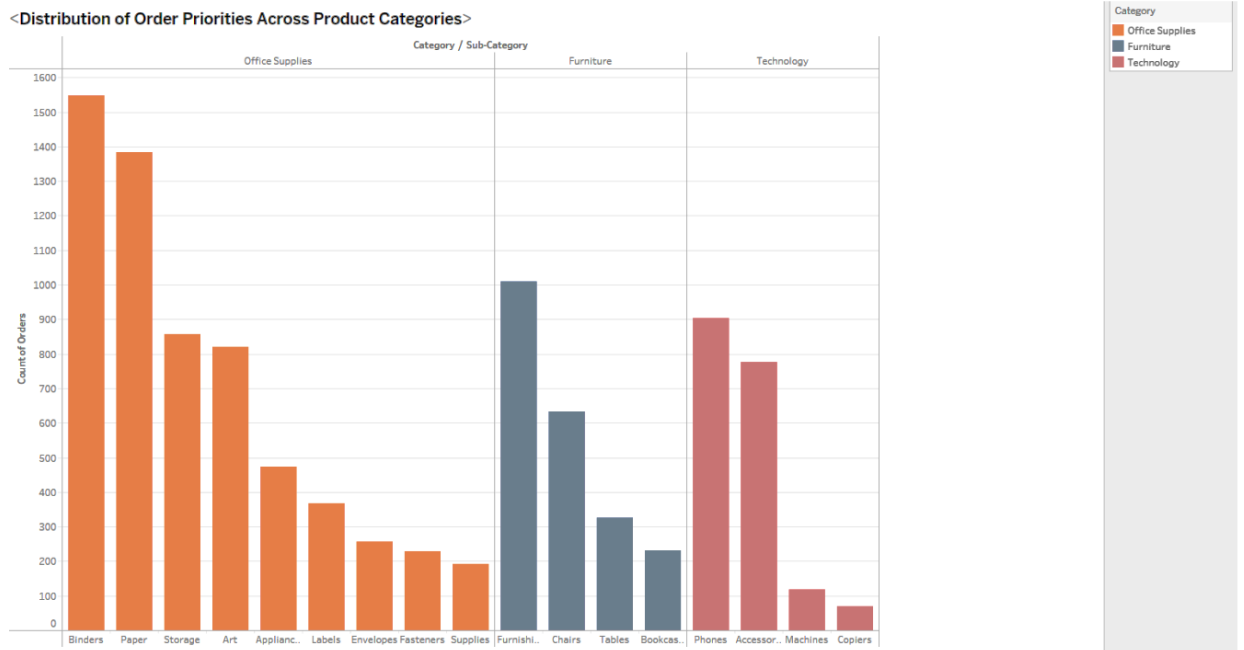
## 26. How does the sales performance of different regions evolve throughout the quarters of a year?



- The line chart will show the sales performance of different regions across the quarters of a year. By analyzing this chart, it's evident that all regions have experienced growth over the quarters. Additionally, the East region exhibits the highest growth compared to other regions, as indicated by the steeper rise or higher trajectory in its sales line.

This visualization aids in understanding the sales trends across quarters for each region, allowing for strategic planning and resource allocation to capitalize on regions with higher growth potential or addressing issues in regions with lower growth rates.

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

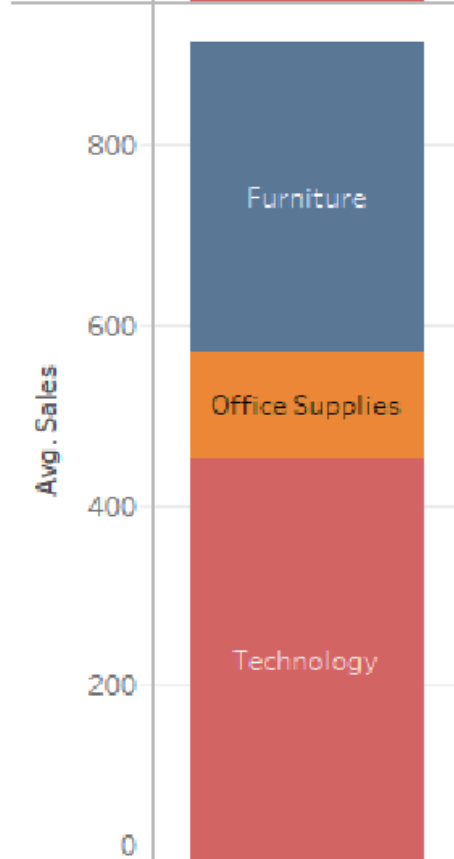
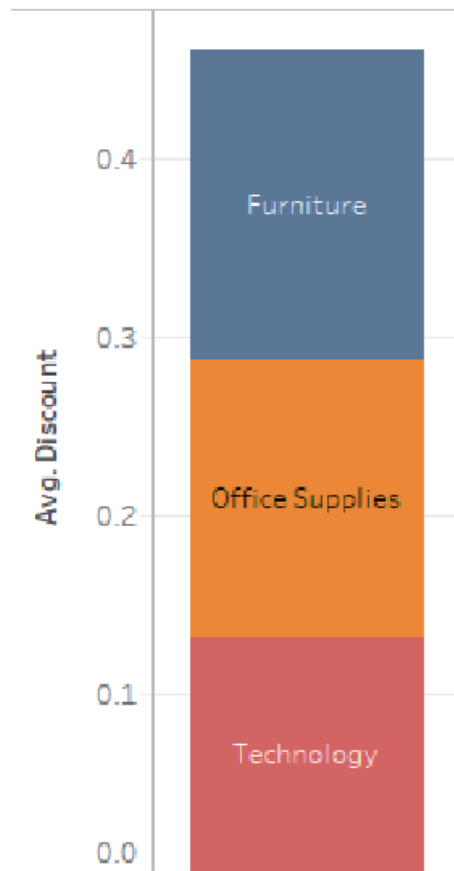


- This visualization will display the distribution of order priorities across various product categories. By examining this chart, we observe the volume of orders based on different priorities within each product category.

If the office supply category shows a higher variety of products with higher priorities, it suggests that certain products within this category are in high demand or are considered more critical by customers, leading to a larger volume of orders with higher priorities. This insight can guide inventory management or stocking strategies for high-priority items within the office supply category.

## 28. What is the relationship between discounts and sales?

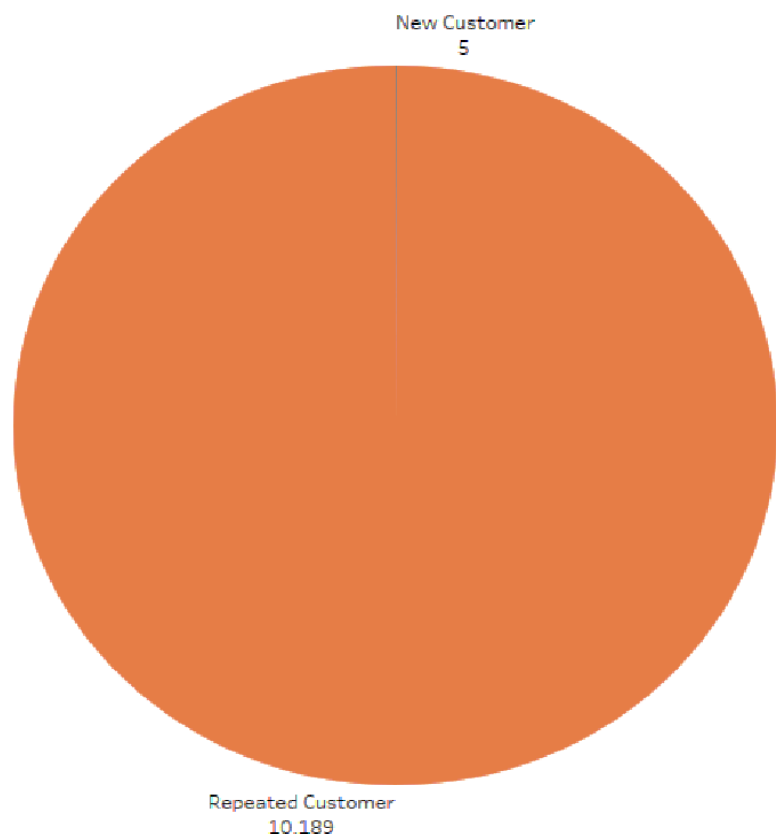
## <Relationship Between Discounts and Sales>



- This visualization enables a direct comparison between discounts and sales for each category.
- If the average discounts are similar across categories but sales in office supplies are considerably lower compared to technology and furniture, it suggests that even though discounts might be uniform, the demand or purchase behavior for office supplies is comparatively weaker. This insight can lead to further investigation into strategies to improve sales for office supply items or optimize discounting strategies to increase sales within this category.

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

### <Average Order Value Difference Between Repeat and New Customers>

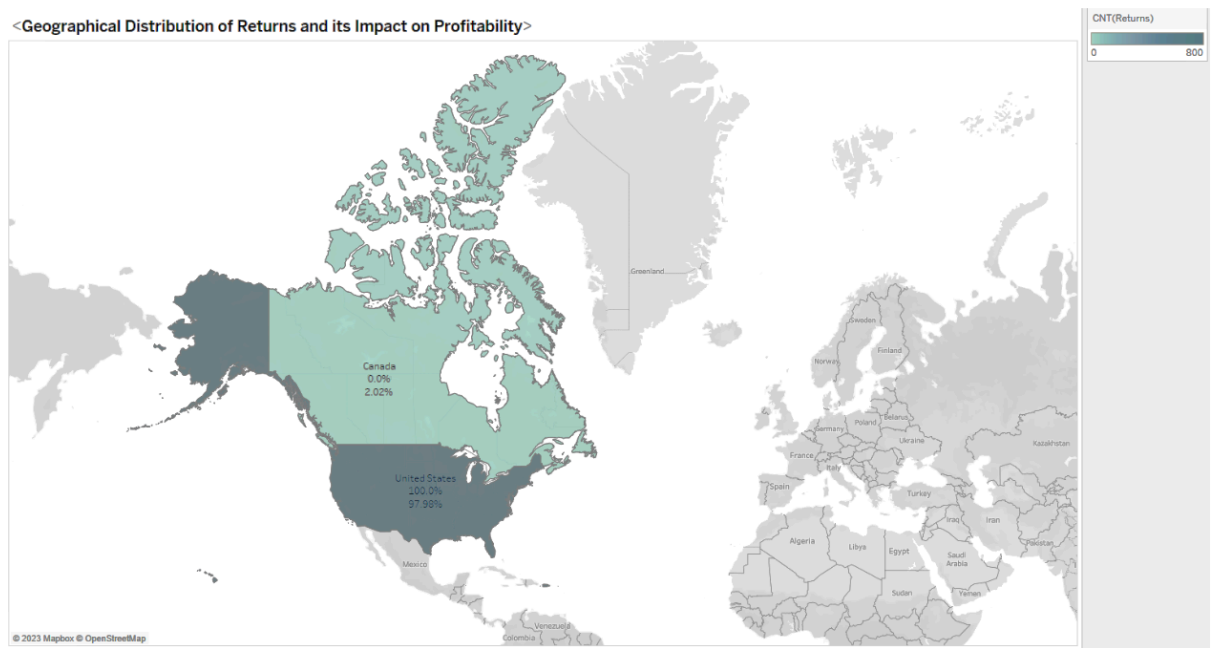


- The pie chart allows for a direct comparison of average order values between repeat and new customers.



- If the majority of customers are repeat customers and they exhibit higher average order values compared to new customers, it implies that fostering customer loyalty might positively impact the order values or overall sales revenue. Conversely, lower order values from new customers might prompt strategies to improve their buying behavior or encourage larger purchases.

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



- The geographic map for returns shows variations in return percentages across different regions, highlighting areas with higher return rates.
- The profitability visualization illustrates the contribution of each region or country to the overall profit, allowing insights into regions contributing more significantly to the company's profits despite potential returns.

By analyzing these separate visuals, one can derive insights into regions where returns are prominent and how they impact overall profitability, offering actionable insights to mitigate returns' negative effects on profits.

## **Dataset Links -**

### **1. Sample - Superstore.xls :**

<https://docs.google.com/spreadsheets/d/1shl6vJnnlkW2Qm-saMxSusbWGcoXWu6-/edit?usp=sharing&ouid=111949844480706214274&rtpof=true&sd=true>

### **2. Sample - Superstore.xls - Orders.csv :**

<https://drive.google.com/file/d/1VIBVjn9yuCibQN0CsfcgqNyz5Ne-kJa3/view?usp=sharing>

### **3. Sample - Superstore.xls - People.csv :**

<https://drive.google.com/file/d/16meo4CAzxnOBQ0GVXniM8UQ1wzE-43HO/view?usp=sharing>

### **4. Sample - Superstore.xls - Returns.csv :**

<https://drive.google.com/file/d/1o4Scv19mLkvwRNLFgGXPUKbxbpT9qr6a9/view?usp=sharing>