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?
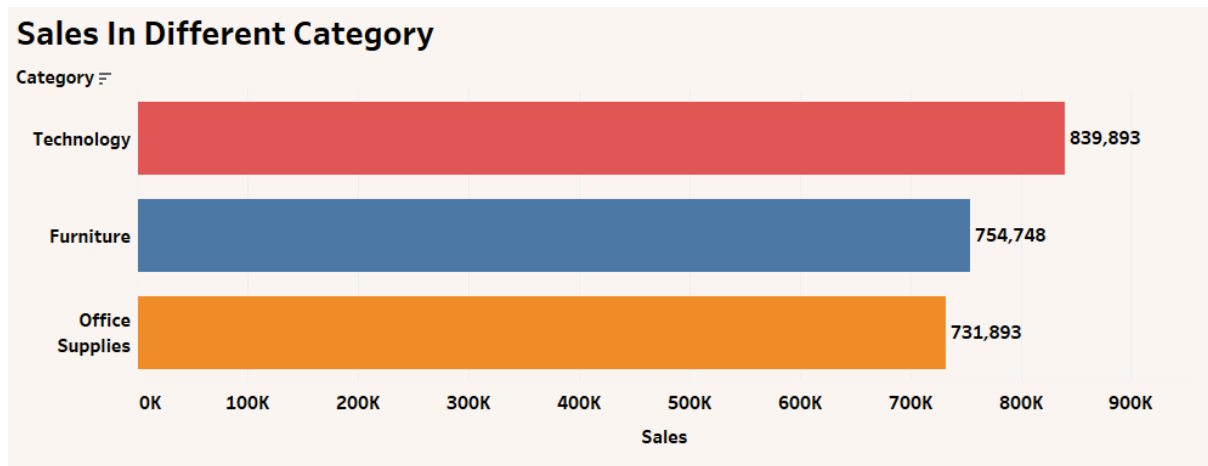

Chart Type: Horizontal Bar Chart.

Explanation:

The Horizontal bar chart clearly displays sales for each product category, allowing for a quick comparison of total sales. The bars' lengths represent the amount of sales, making it easy to determine which category performs best.

Insights:

From the chart, we can see that the **Technology** category leads in total sales, followed by **Furniture** and **Office Supplies**. This suggests that Technology items are most demanded in the store.

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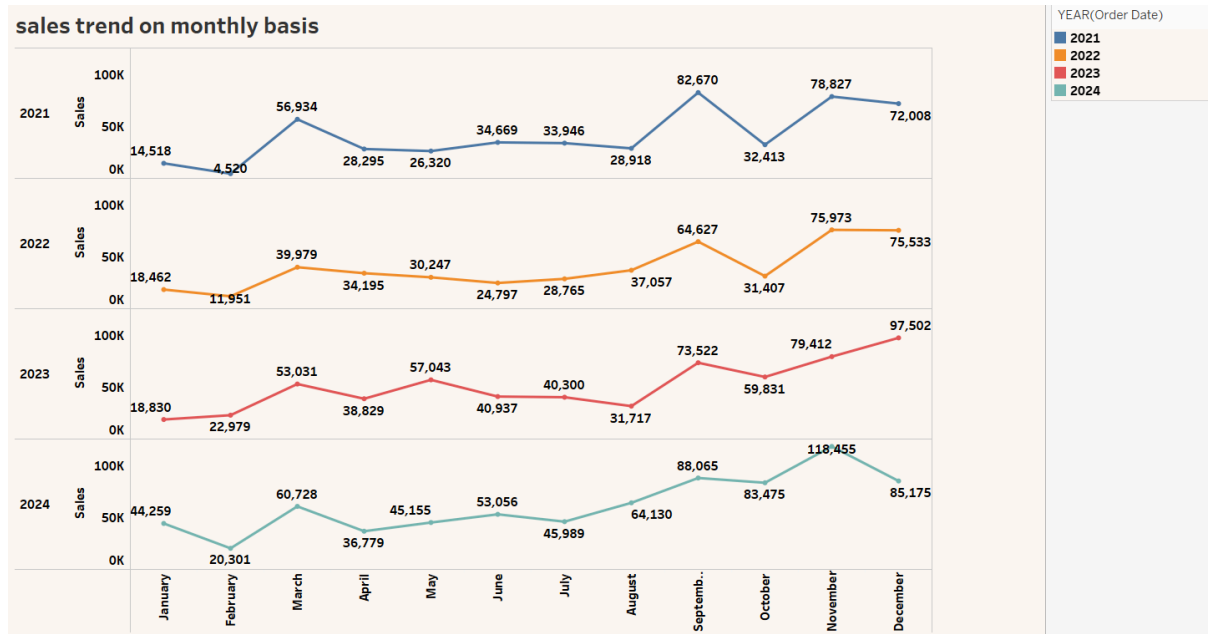
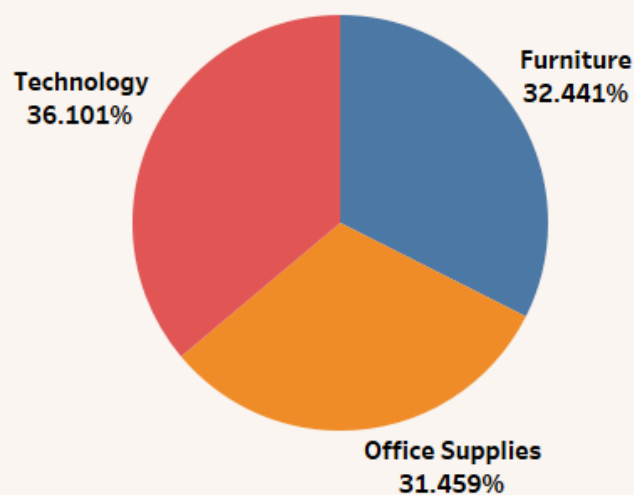

Chart Type: Line Chart

Explanation:

A line chart is ideal for showing changes over time, with months on the x-axis and year and sales on the y-axis.The trend line makes it easy to observe fluctuations throughout the year.

Insights:

The line chart shows that sales peak in **September**,**November** and **December**, likely due to holiday shopping trends. The lowest sales occur in **February**, suggesting a possible off-season period.

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


Chart Type: Pie Chart

Explanation:

A pie chart shows how total sales are divided among different categories. Each slice represents the proportion of sales contributed by a product category.

Insights:

The pie chart reveals that **Technology** contributes the largest share to total sales, followed by **Office Supplies** and **Furniture**. This indicates that Technology is the dominant category, representing a significant portion of overall revenue.

4. Can we analyze the sales performance of individual customers over time?
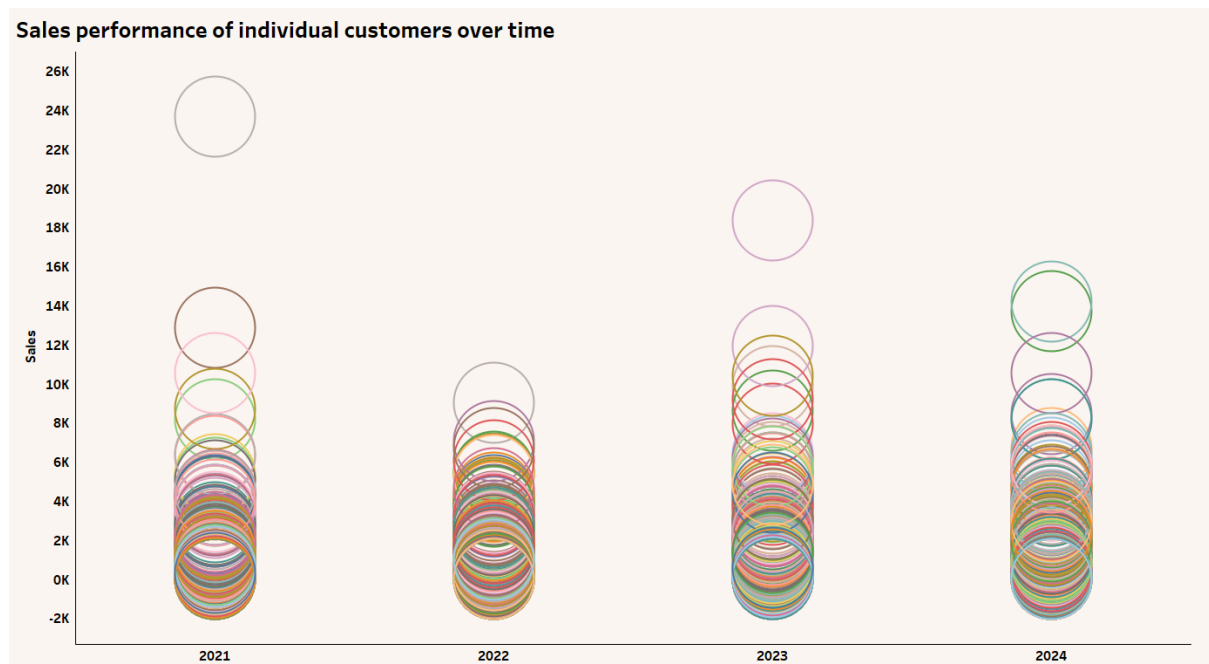
Chart Type : Circle View Chart

Explanation:

Using the Circle View Chart , we can easily visualise the sales performance of individual customers , the overall pattern of sales by the customers , and also the top customers in terms of sales performance.

Insights:

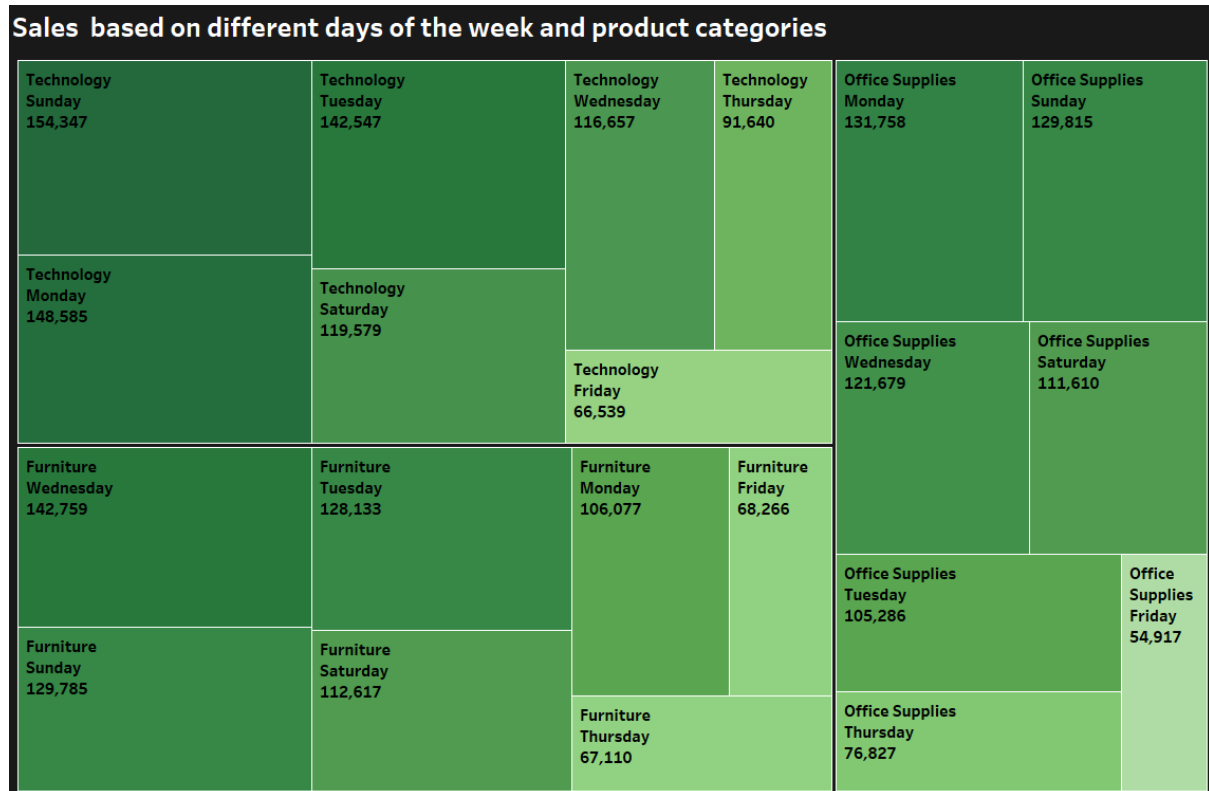The Circle View Chart reveals that the overall sales amount of individual customers are around **0** to **5000.**
Top customer in 2021 - **Sean Miller**
Top customer in 2022 - **Peter Fuller**
Top customer in 2023 - **Tamara Chand**
Top customer in 2024 - **Raymond Buch**

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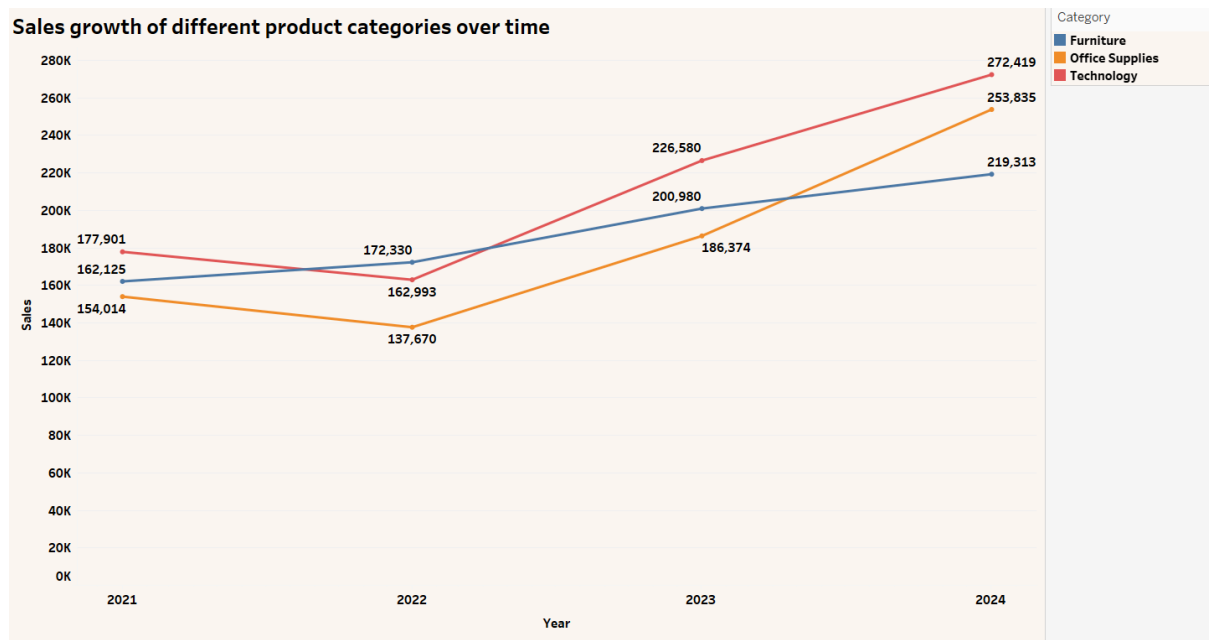

Chart Type: Heat Map

Explanation:

A heat map visually represents the relationship between days of the week and product categories by using colour gradients. Darker shades indicate higher sales.

Insights:

The heat map shows that **Technology** sales spike on **Sunday , Furniture** sales spike on wednesday , and **Office Supplies** perform well on Monday.

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

Sales growth of different product categories over time

| Category |
| Furniture |
| Office Supplies |
| Technology |

Data points from chart:
- 2021: 177,901 (Technology), 162,125 (Furniture), 154,014 (Office Supplies)
- 2022: 172,330 (Furniture), 162,993 (Technology), 137,670 (Office Supplies)
- 2023: 226,580 (Technology), 200,980 (Furniture), 186,374 (Office Supplies)
- 2024: 272,419 (Technology), 253,835 (Office Supplies), 219,313 (Furniture)

Chart Type: Line Chart

Explanation:

A line chart allows us to compare sales growth for multiple categories over time, with different lines representing different categories.

Insights:

The chart shows steady sales growth in **Furniture** , **Technology** and **Office Supplies** sales reduced in **2022** but after 2022 sales growth increased at almost similar rates.

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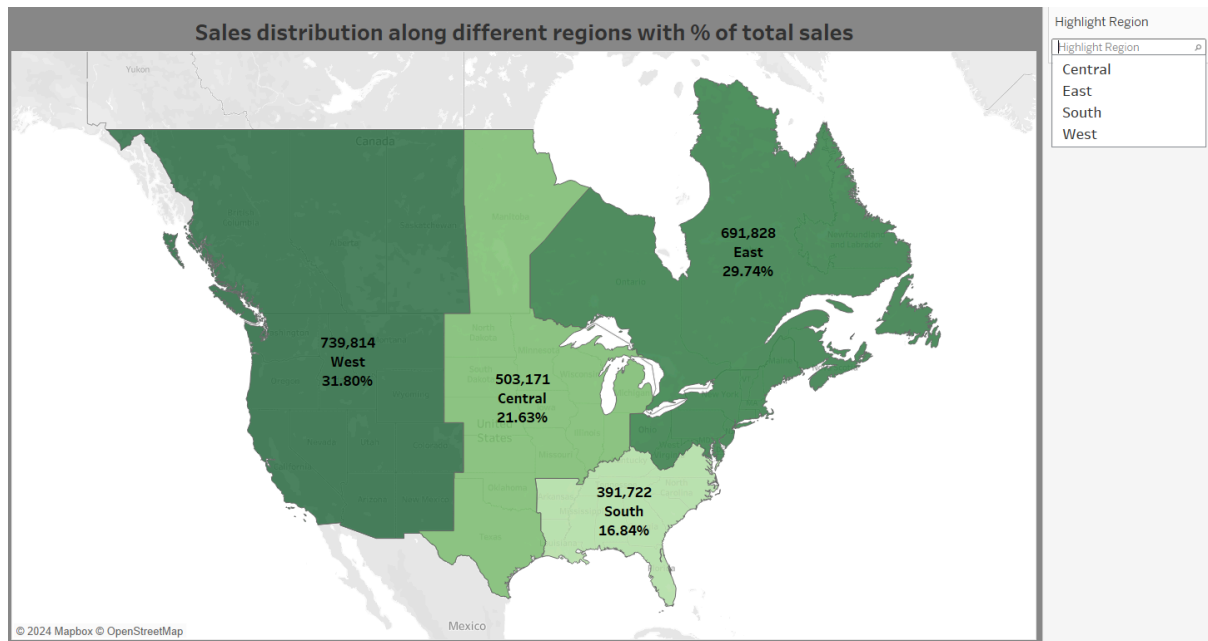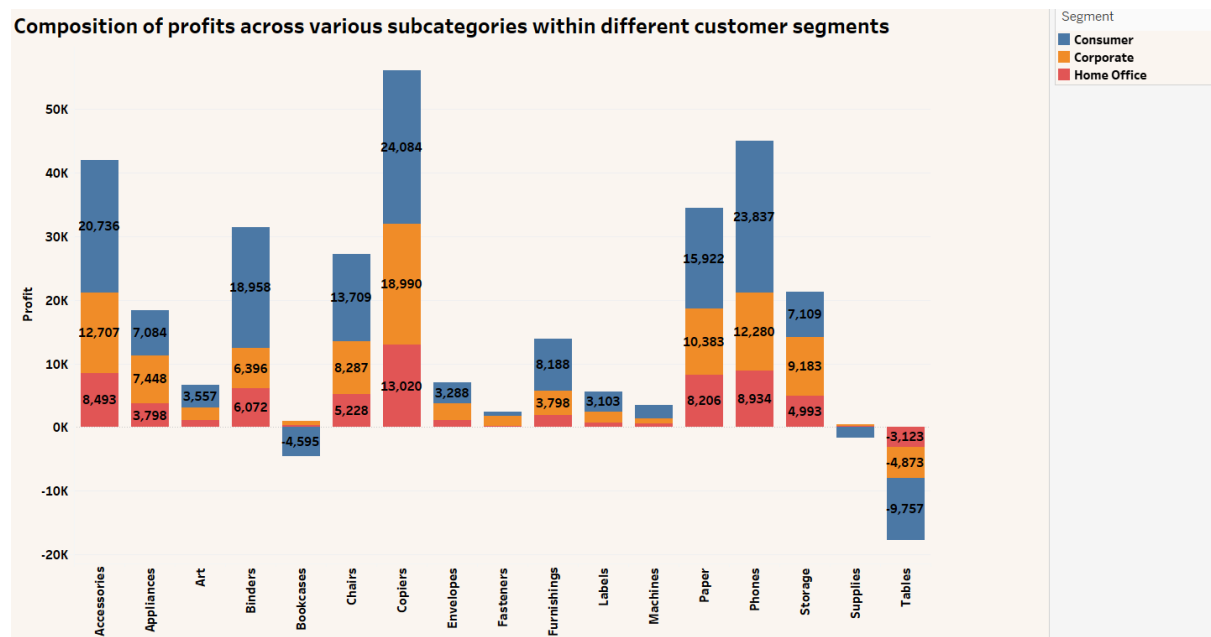

Chart Type: Map

Explanation:

A Map shows the spread of sales for each region, and dark colour represents the more sales and lighter colour represents lower sales . I also used a calculated field to show the percentage of total sales in the region . This allows for a comparison of sales distributions across regions.

Insights:

The **West** region has the highest sales, having 31.8 % of total sales.Then comes the **East** region with 29.74 % and **Central** region with 21.63 % . The **South** region has lower sales.

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


Chart Type: Stacked Bars

Explanation:

Stacked Bar chart is useful for showing the distribution of different subcategories to total profit within each customer segment. Each subcategory's total profits is represented by a section of the bar.

Insights:

Phones, **Chairs**, and **Accessories** are the most profitable subcategories across all segments. **Tables** and **Bookcases** are the only subcategories facing significant losses, particularly in the **Consumer** segment.
Consumer dominates as the largest segment in terms of total profit, followed by **Corporate** and **Home Office**.

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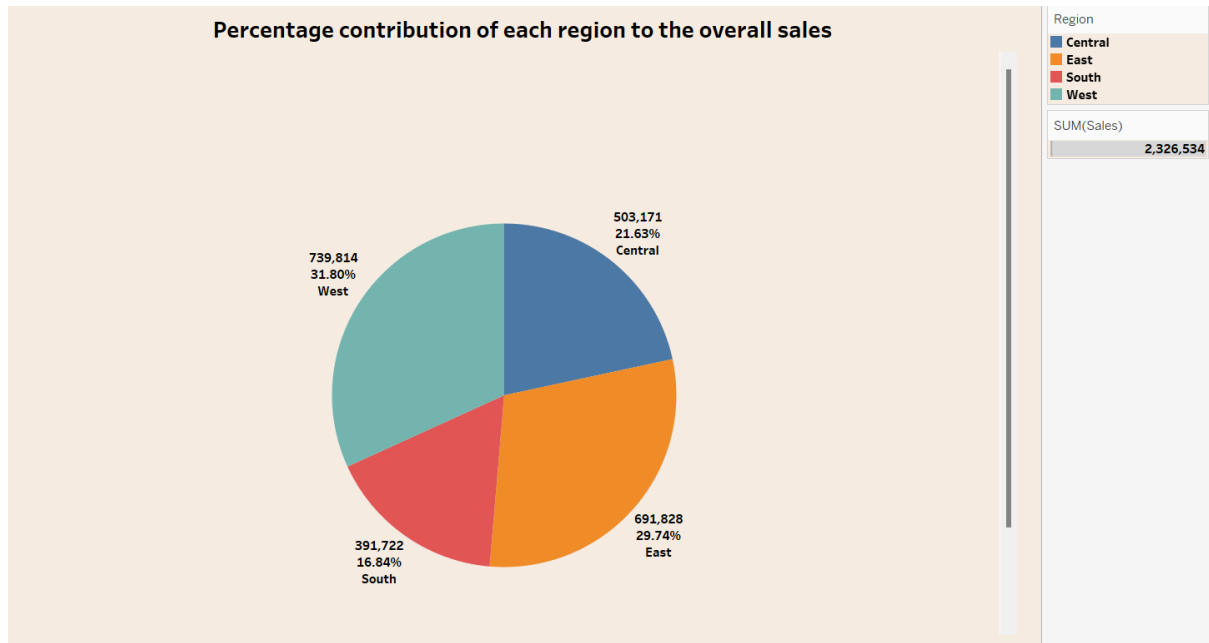

Chart Type: Pie Chart

Explanation:

A pie chart visually displays the proportion of total sales attributed to each region.

Insights:

The **West** region contributes the largest share of sales, followed by the **East** and **Central** regions. The **South** region contributes the least, indicating potential for growth.

10. Can we visualise the profit margins associated with different shipping modes and customer segments?
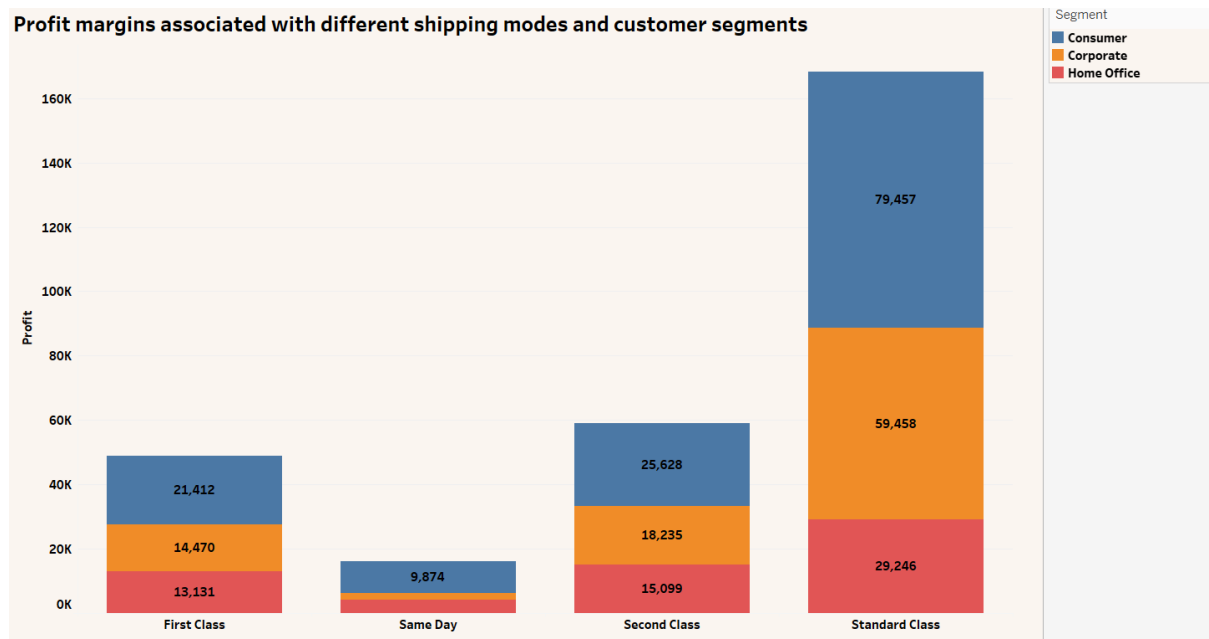
Chart Type: Stacked Bar Chart

Explanation:

A stacked bar chart effectively shows profit margins by both shipping mode and customer segment.

Insights:

Standard Class dominates in terms of **profit**, especially in Consumer segments and **Same Day** shipping mode has the lowest profit margins especially in the corporate segment.

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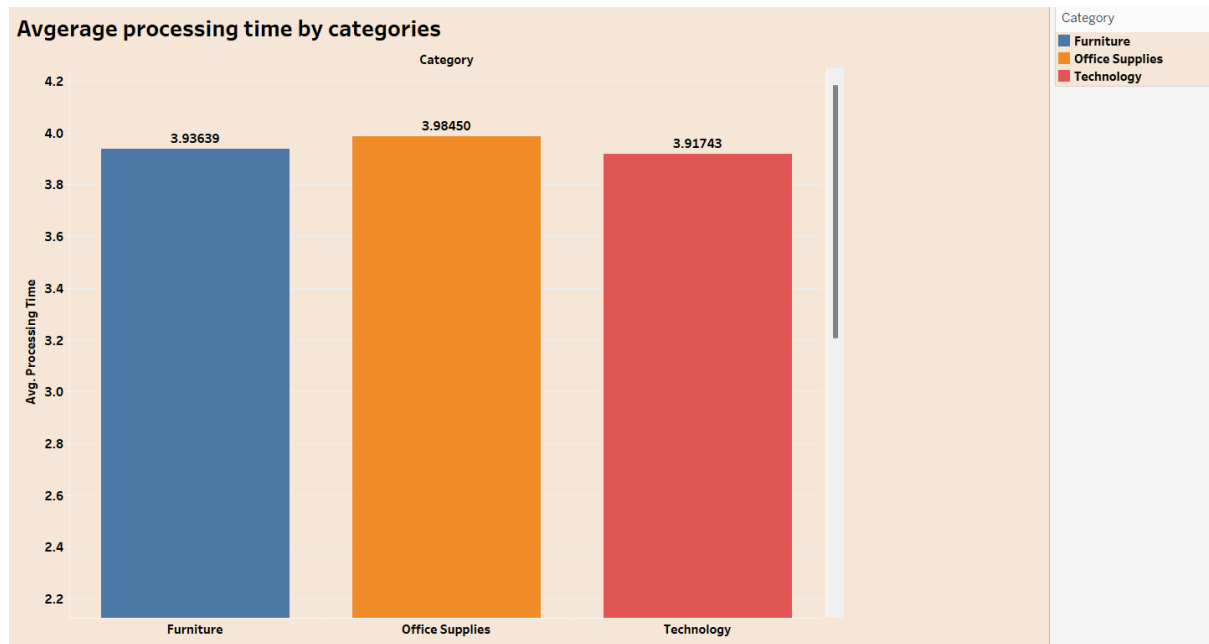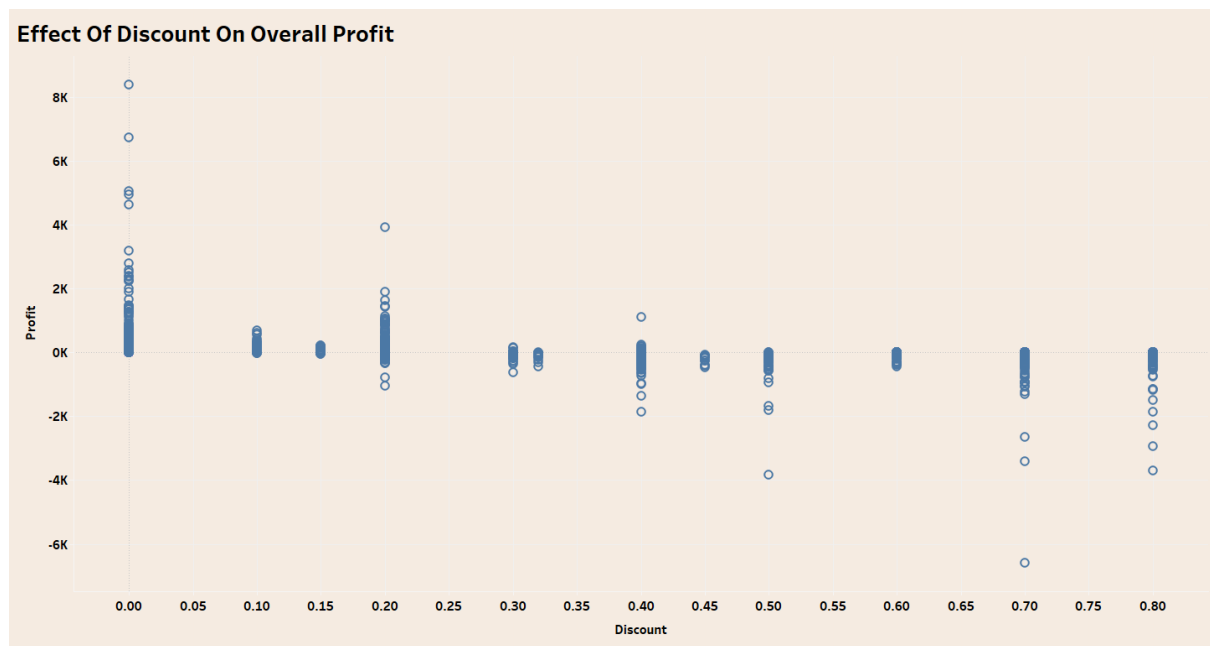

Chart Type: Side-By-Side Bar Chart

Explanation:

A Side-By-Side Bar chart effectively shows Average processing time in different product categories.
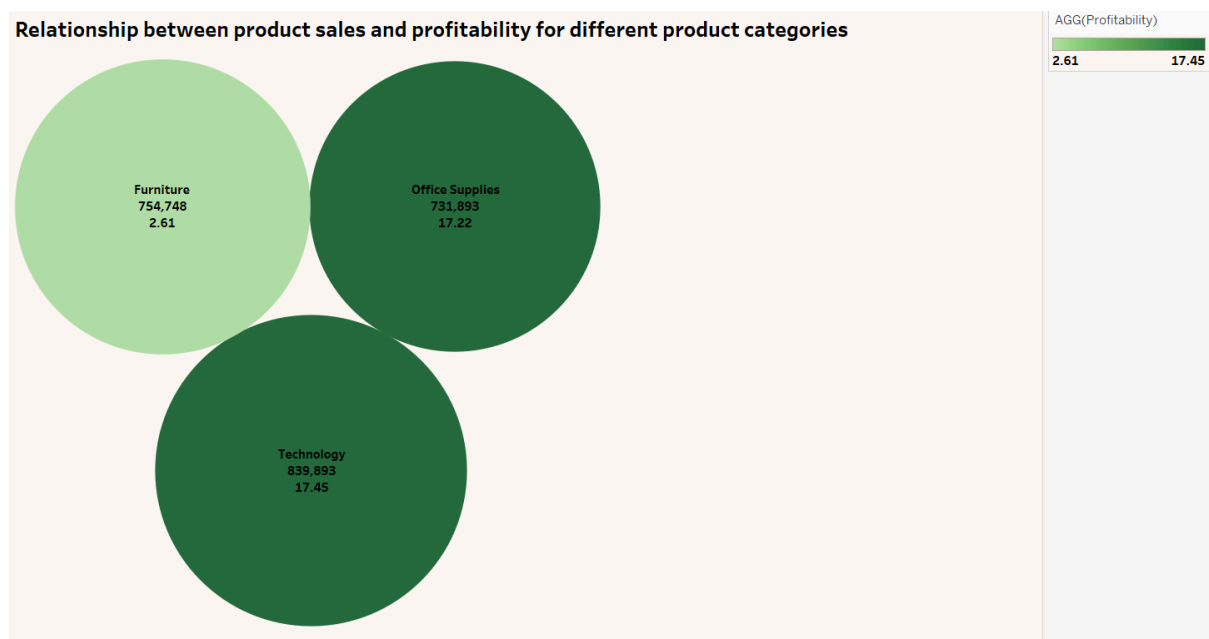
Insights:

The average processing time is overall similar in all three categories which is around 3.9 .

12. How do discounts affect overall profit?

Effect Of Discount On Overall Profit

Chart Type: Scatter Plot

Explanation:

A scatter plot allows us to assess the relationship between discount rates and profit.

Insights:

The chart reveals a negative correlation between discounts and profit. As discount rates increase, profits generally decrease, suggesting that high discounts may erode profit margins.

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

Chart Type: Packed Bubble Chart

Explanation:

A Packed bubble chart is useful for showing the relationship between sales in terms of size of the bubble and profitability in terms of colour intensity.

Insights:

Technology and **office supplies** have overall same sales and profitability but **Furniture** have the lowest profitability in **superstore**.

14. What is the distribution of order quantities for products in the dataset?
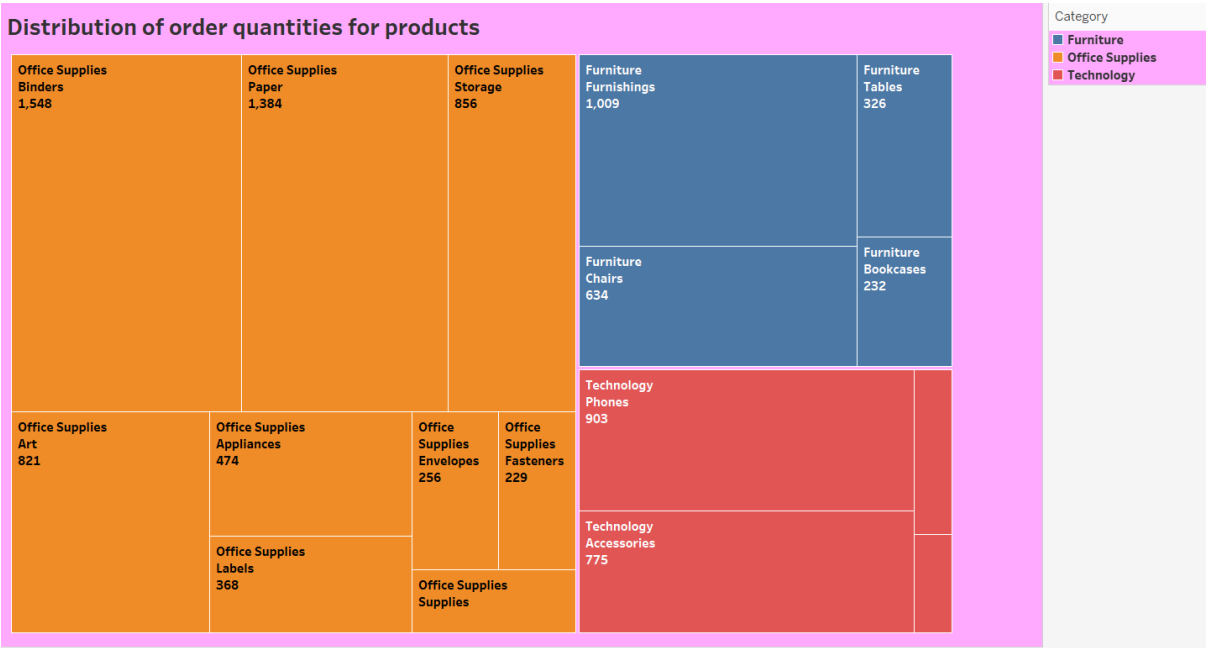
Chart Type: Treemap

Explanation:

A Treemap is used to show the distribution of order quantities in terms of products.

Insights:

Office Supplies have the highest orders in which **Binders** subcategory is dominating.

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

Profit distributions across different product categories
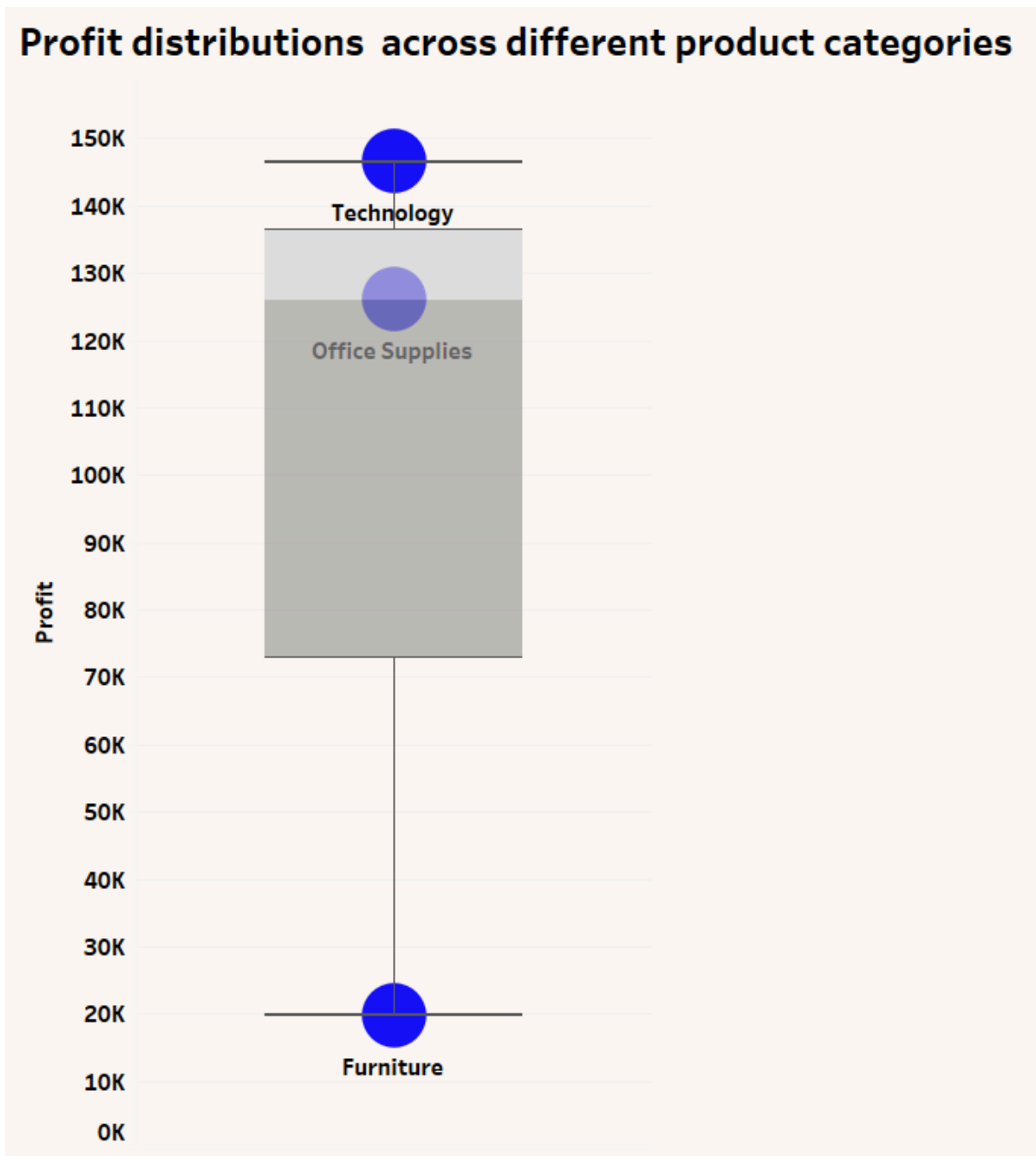
Chart Type: Box Plot

Explanation:

Box plots show the range and spread of profit across different categories.

Insights:

Technology and **Office Supplies** are far more profitable compared to **Furniture**, with Technology having the highest median profit.
Furniture shows low profits and minimal variability, indicating that it consistently performs worse than the other categories.
Office Supplies has a wider distribution, meaning its profits vary more widely within the category.

16. Can we compare the shipping time distributions for different shipping modes?
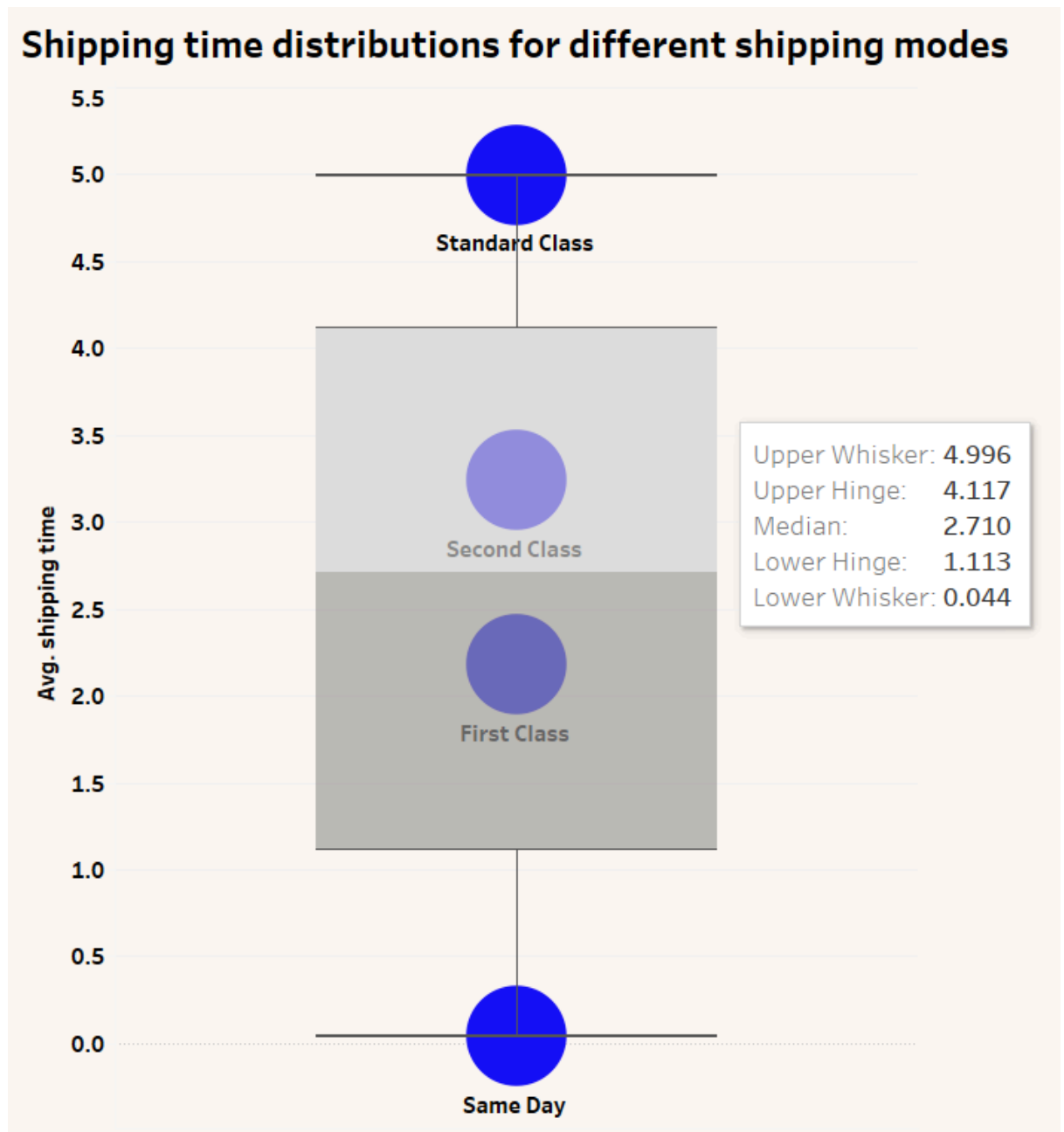
Chart Type: Box Plot

Explanation:

A box plot is ideal for showing the distribution of shipping times for different shipping modes.

Insights:

Same Day shipping is consistently fast with virtually no variation.
First Class and **Second Class** shipping offer faster delivery times than Standard Class, but with some variability, especially in Second Class.
Standard Class takes the longest and has the widest range in shipping times, making it the slowest and least consistent option.

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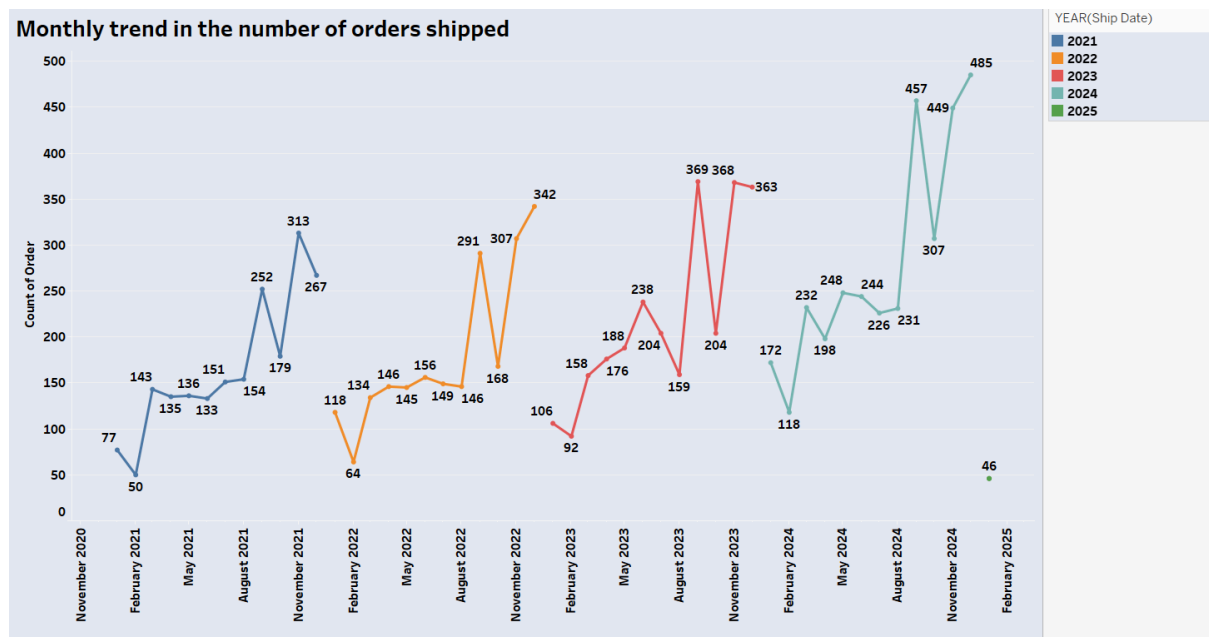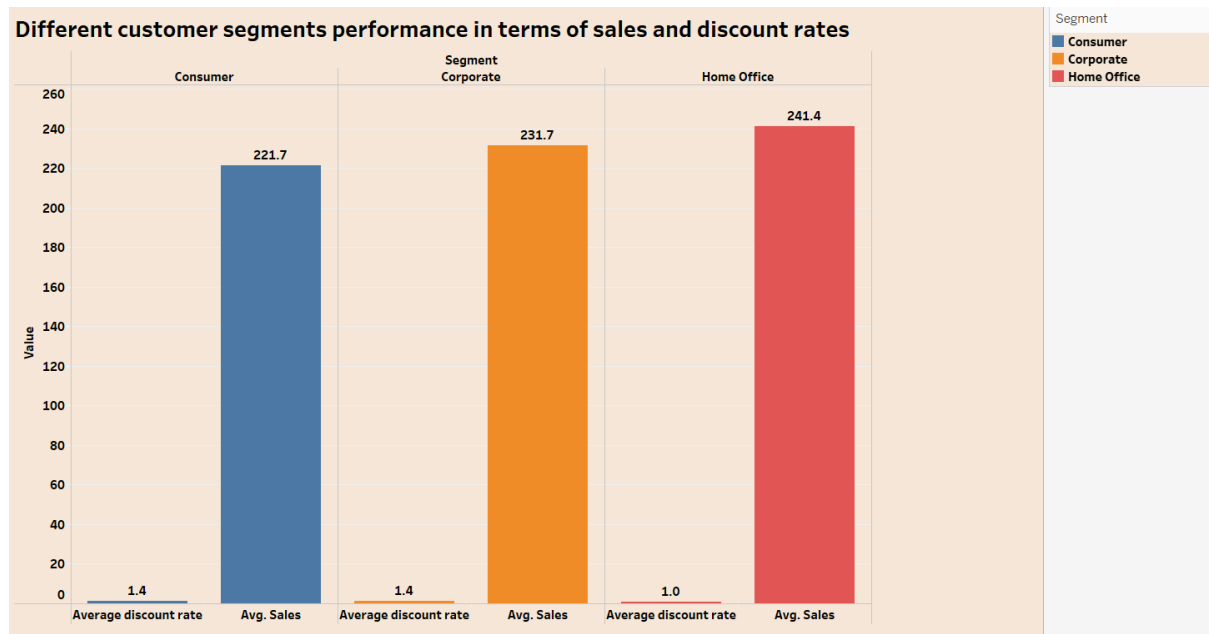

Chart Type: Line Chart

Explanation:

A line chart effectively shows the number of orders shipped each month, revealing trends.

Insights:

Orders tend to peak in **November and December**, corresponding to holiday seasons. The first half of the year has lower levels of shipping.

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



Chart Type: Side-by-Side Bar Chart

Explanation:

A side-by-side bar chart allows us to compare the sales and discount rates for different customer segments.

Insights:

Despite having low average discount rates , **Home Office** got the highest sales in all three segments but there is not a huge difference between their discount rates.

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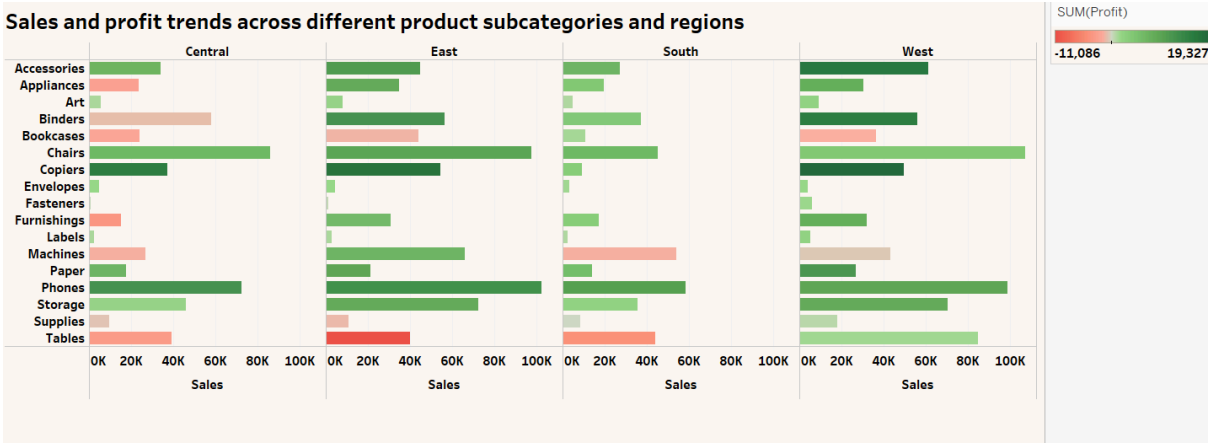

Chart Type: Horizontal Bars

Explanation:

A Horizontal Bars chart allows us to track sales and profit trends across different subcategories with regions.

Insights:

Technology subcategories like **Phones** and **Accessories** drive the highest profits, while **Office Supplies** maintain the steady sales across all regions. However, profitability issues in **Furniture**, particularly in the Central and South regions, suggest the need for more cost-efficient strategies in these areas.

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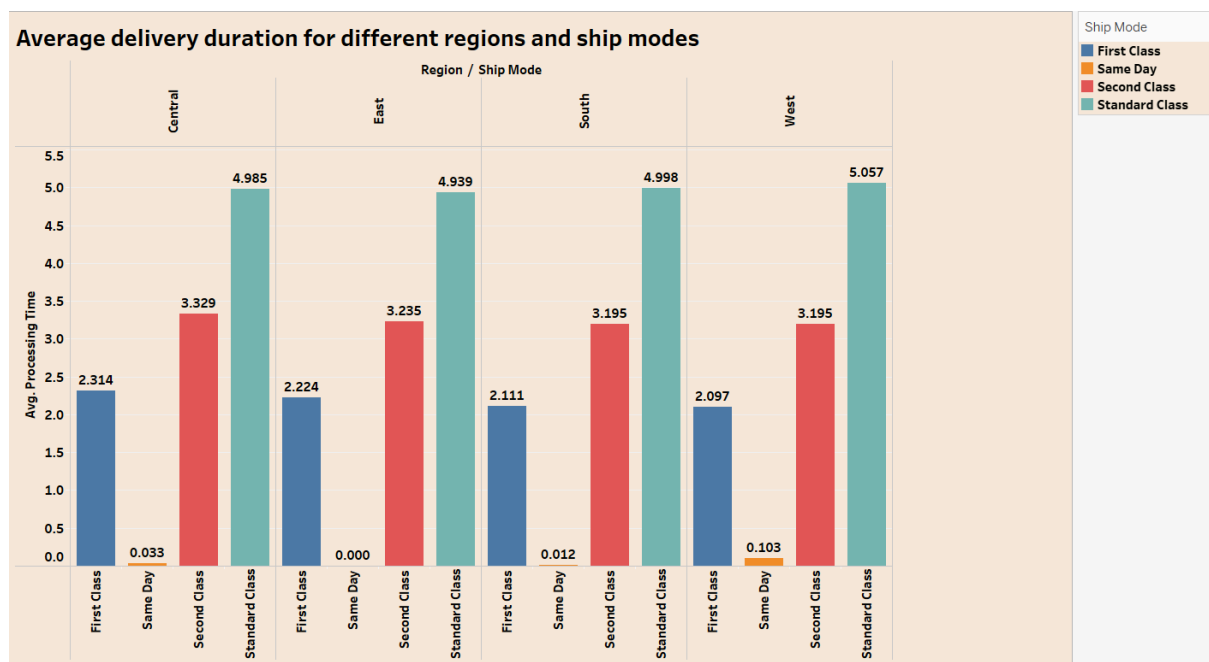

Chart Type: Side-By-Side bar

Explanation:

A Side-By-Side bar allows us to easily see variations in delivery duration across regions and shipping modes.

Insights:

Same Day delivery is the quickest across all regions, while **Standard Class** takes the longest. The **West** region consistently shows slightly longer delivery times across all modes.

21. How has the average order quantity changed over the years for various product categories?
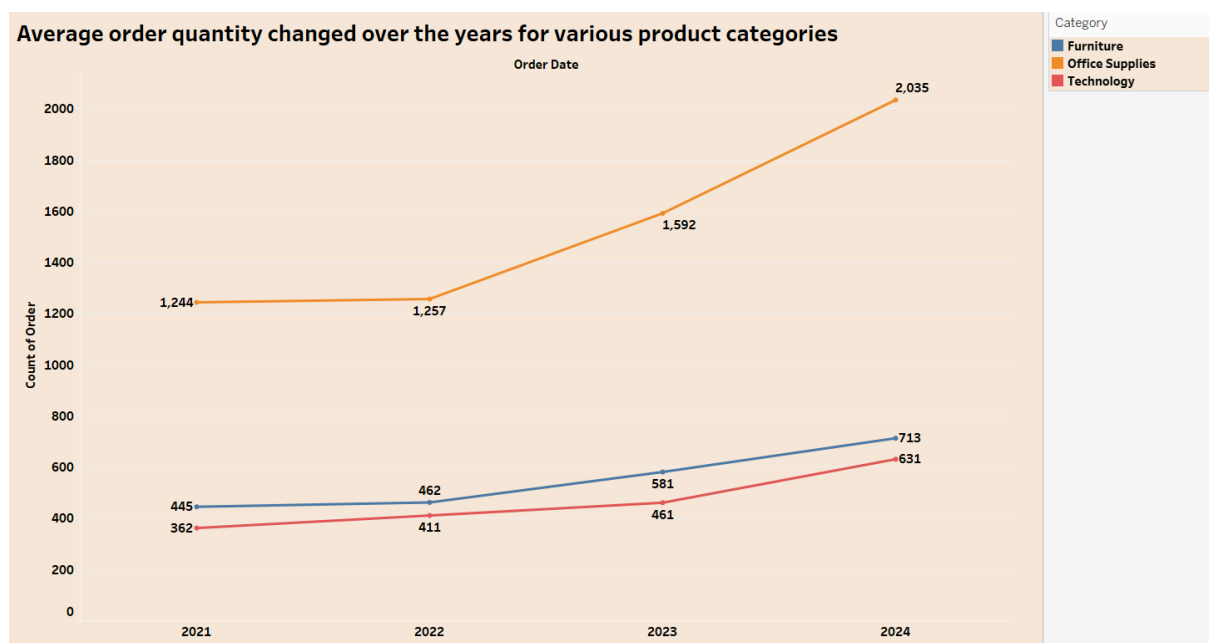
**Chart Type:** Line Chart

Explanation:

Line charts are ideal for tracking changes over time, showing the evolution of average order quantities for different product categories.

Insights:

Office Supplies orders have seen consistent growth in quantity over the years, while **Furniture** and **Technology** show smaller fluctuations.

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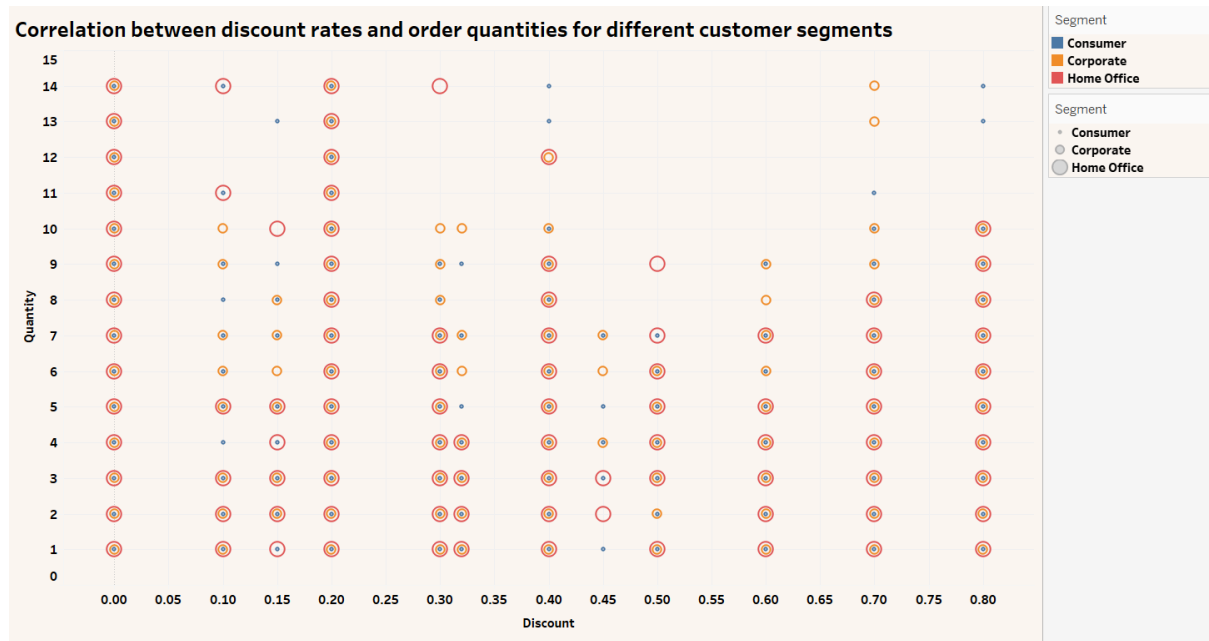

Chart Type: Scatter Plot

Explanation:

A scatter plot shows how discount rates correlate with order quantities, with different colours or shapes representing customer segments.

Insights:

Higher discounts are generally associated with **larger orders**, particularly in the **Corporate** and **Home Office** segments.

23. What is the proportion of orders returned in each region within the Superstore dataset?
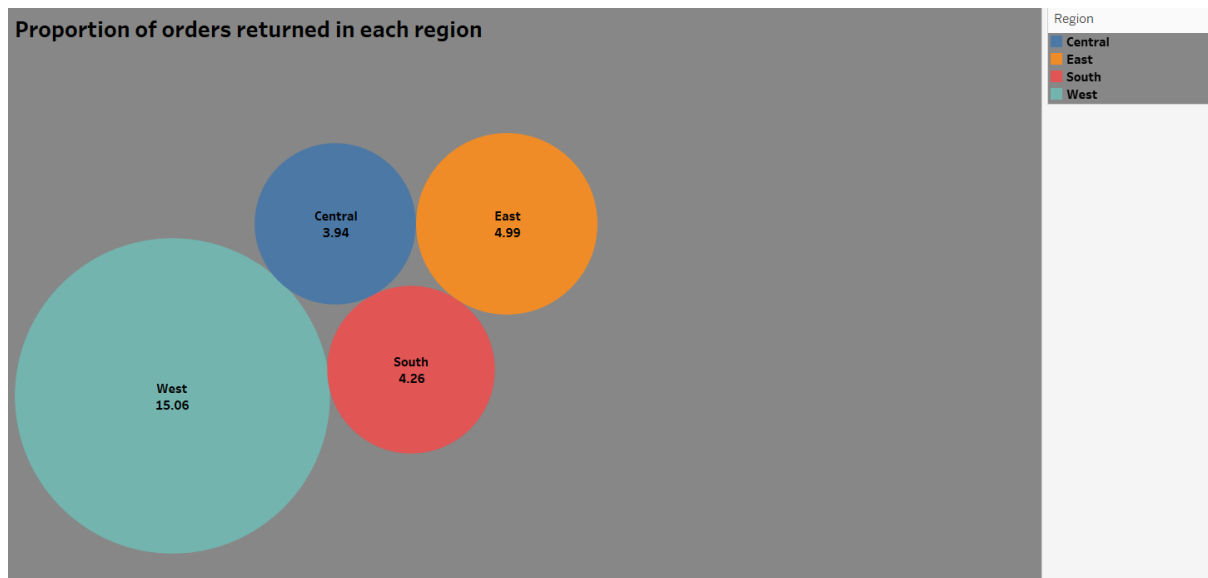
**Chart Type:** Packed Bubble

Explanation:

A Packed Bubble visually represents proportions of returns by region.

Insights:

The **West** region typically has the highest proportion of returns, followed by the **East** and **Central** regions. The **South** region has the fewest returns. This suggests that there may be operational or product quality issues more prevalent in certain regions.

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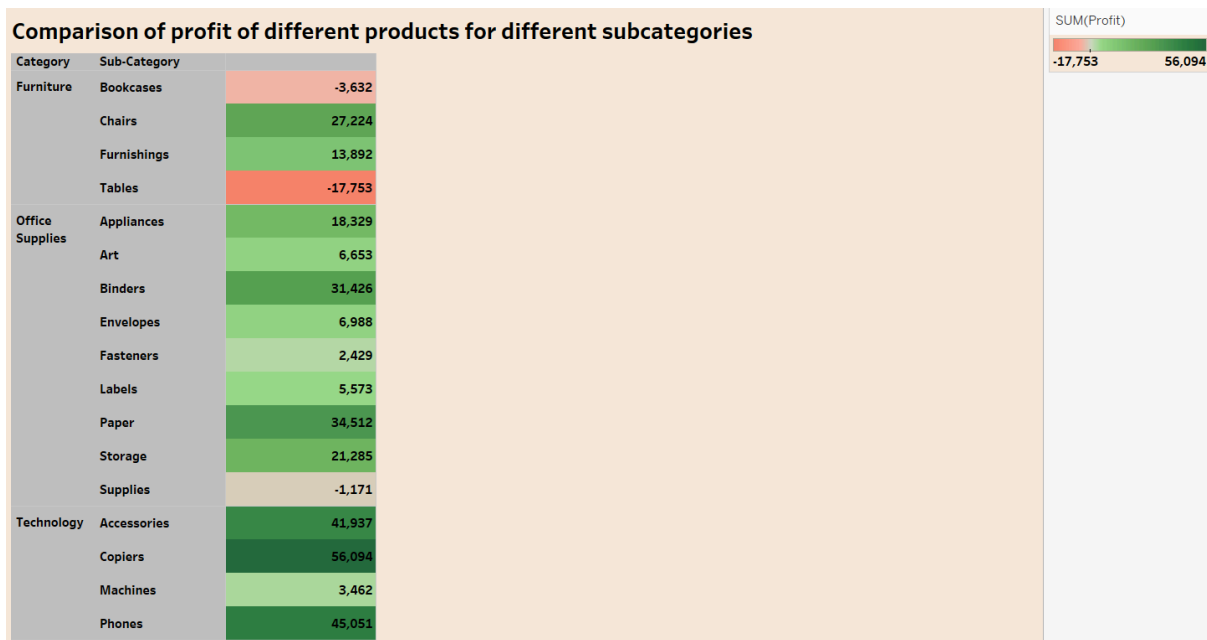

Chart Type: Highlight Table

Explanation:

A Highlight Table allows for comparison across multiple subcategories.

Insights:

Technology subcategories like **Copier, Phones** and **Accessories** have significantly to highest profit. In contrast, certain **Furniture** subcategories such as **Tables** often generate smaller or even negative profits due to higher costs or operational inefficiencies.

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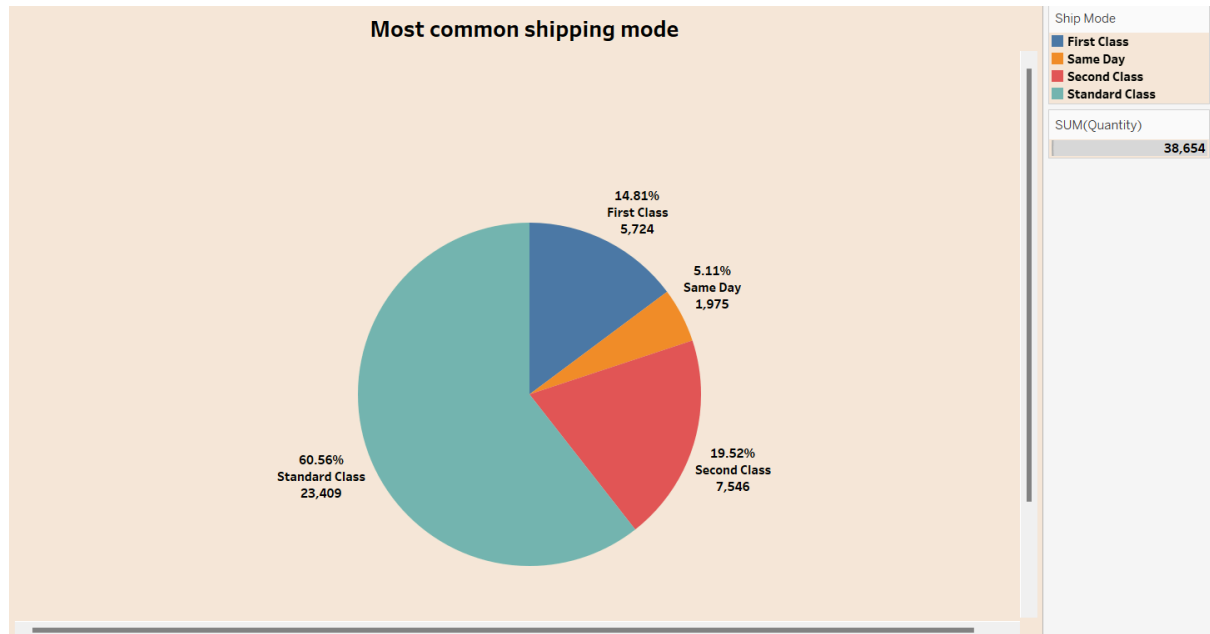

Chart Type: Pie Chart

Explanation:

A simple Pie chart is used to compare the frequency of use for different shipping modes.

Insights:

Standard Class is the most commonly used shipping mode, accounting for a majority of the shipments, followed by **Second Class** and **First Class**. This could indicate that customers prefer slower, cost-effective shipping options, while **Same Day** is used much less frequently, possibly due to higher costs.

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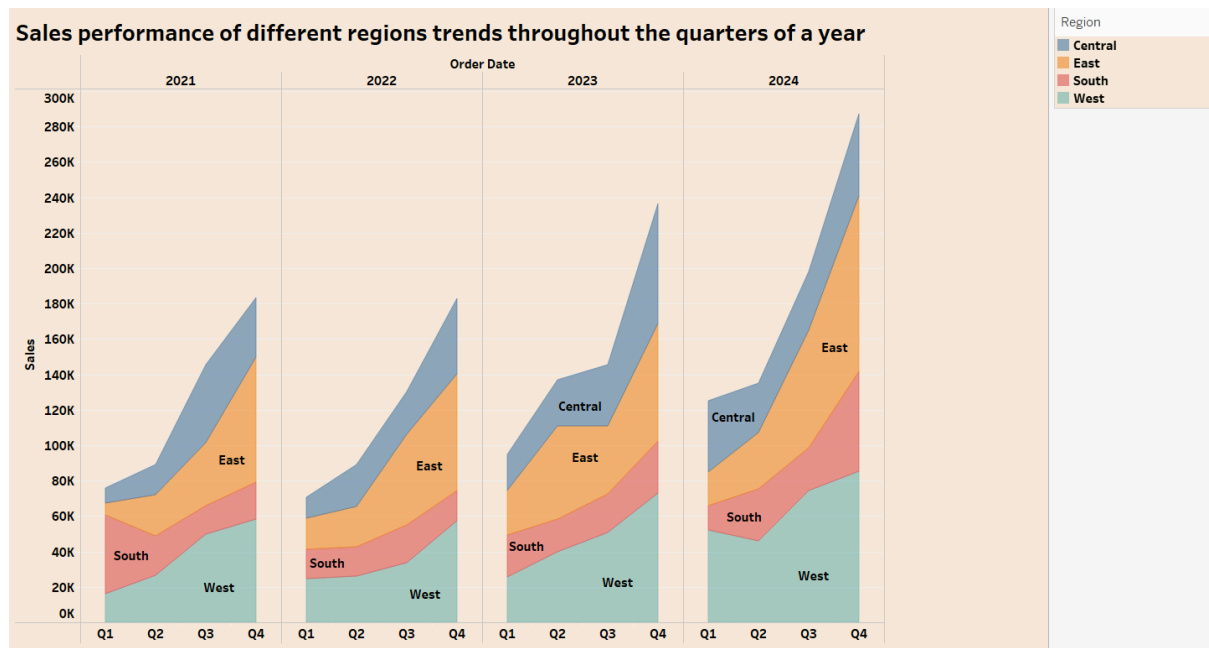

Chart Type: Stacked Area Chart

Explanation:

A Stacked area chart can track sales performance over the quarters for each region. The stacked area chart shows both individual region trends and the overall sales growth.

Insights:

Sales in the **West** and **East** regions show stronger growth in Q3 and Q4, likely due to holiday sales. The **Central** and **South** regions remain more stable throughout the year but see smaller spikes in Q4. Seasonal trends seem to play a significant role in driving sales across regions.

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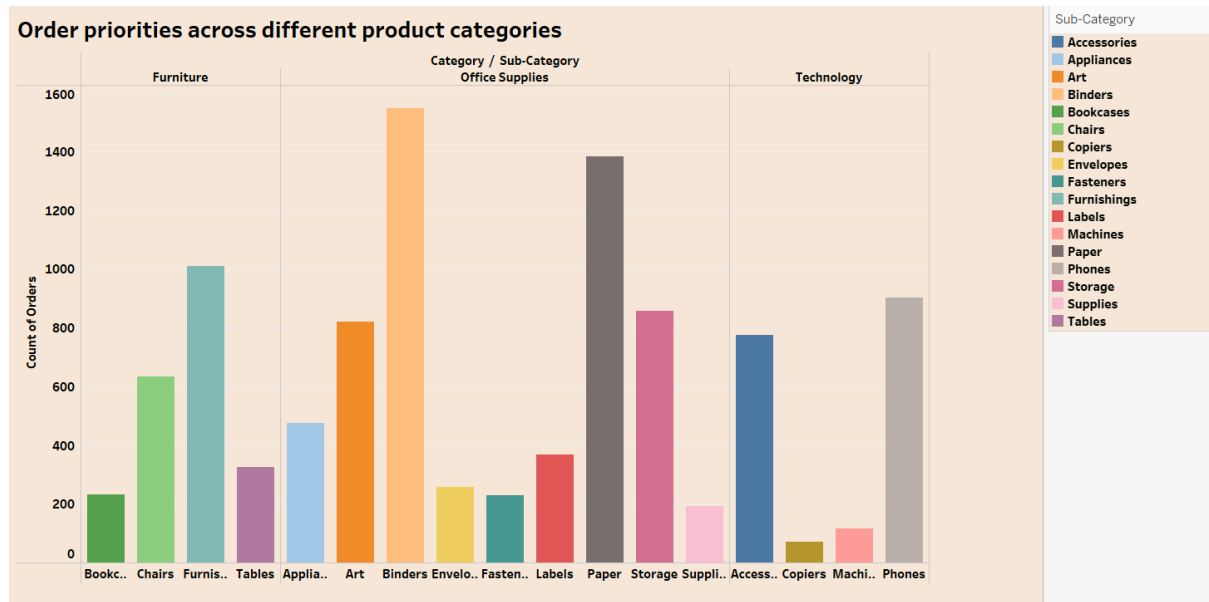

Chart Type: Side-By-Side Bar Chart

Explanation:

A Side-By-Side Bar Chart chart allows for easy comparison of order priorities within different product categories.

Insights:

Technology often has **low-priority** orders, while **Office Supplies** tend to have a greater proportion of **high-priority** orders. This could be due to the urgent need for **Office Supplies** products, while **Technology** products can be planned ahead for restocking.

28. What is the relationship between discounts and sales?
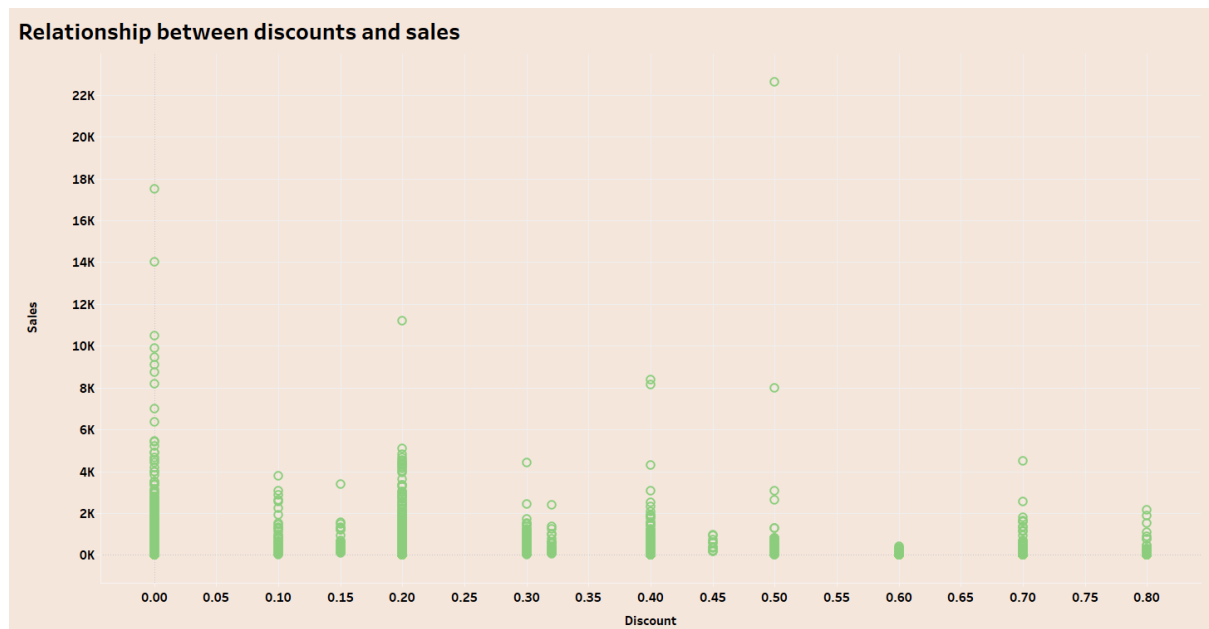

Chart Type: Scatter Plot

Explanation:

A scatter plot visualises the relationship between discounts (x-axis) and sales (y-axis), with each point representing an individual order.

Insights:

There is a **positive correlation** between discounts and sales, meaning that higher discounts often result in increased sales. However, the relationship is not linear—after a certain point, very high discounts do not lead to proportionally higher sales, indicating diminishing returns on discounting.

29. How does the average order value differ between repeat customers and new customers?
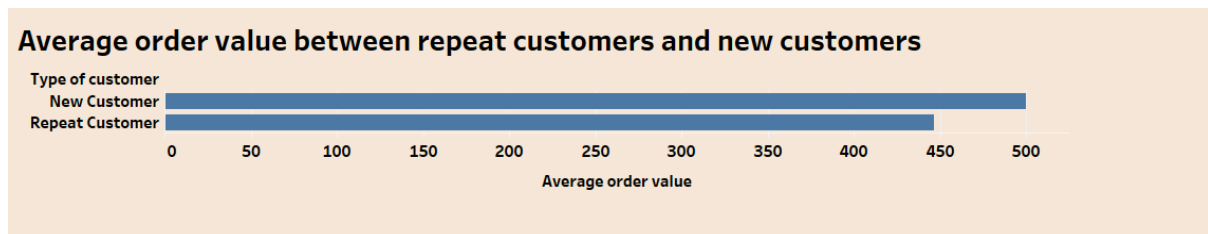
Chart Type: Horizontal Bars
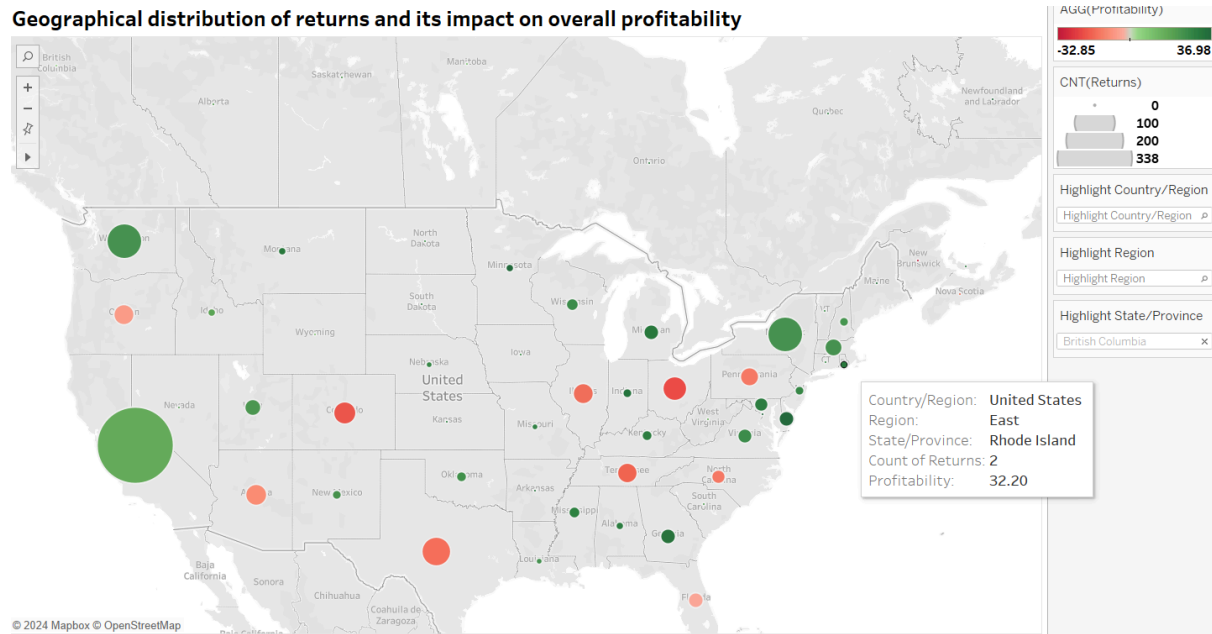
Explanation:

A Horizontal Bars effectively compares the average order values between repeat and new customers.

Insights:

Repeat customers' overall average order value is slightly lesser than **New customers**. We can assume that the **New customers** tend to spend more money than the **Repeat customers**.

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

Chart Type: Symbol Map with Colour Shading.

Explanation:

Symbol Map with Colour Shading allows us to visualise where returns are concentrated geographically. The green and red diverging colour represents the profitability respectively.

Insights:

Returns are more frequent in the **West** and **East** regions, which are also the regions with the highest sales. **Profitability is negatively affected** in these regions but in some **Central** and **South** regions the number of returns are low but also the profitabilities are also low, may be due to lower sales in these regions.