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 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 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 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.
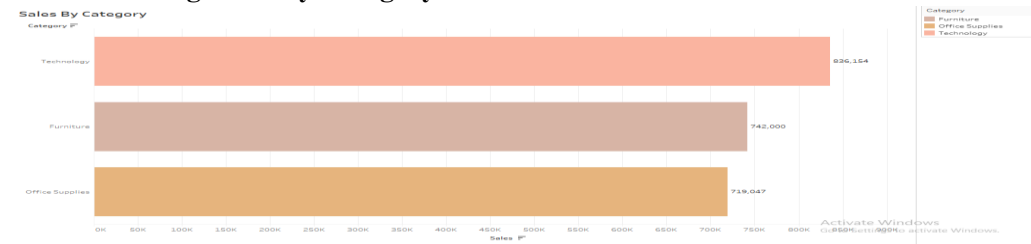
Good luck!

Questions:

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

   Reason: Horizontal bar charts are great for comparing categories, especially when dealing with longer category names. Their readability and straightforward layout make it easier to see differences and trends at a glance.

   **Chart Showing Sales By Category-**

   Here's a detailed analysis of the chart:

   - **Technology**: This category has the highest sales, totaling 836,154. It is represented by the longest bar, coloured in light pink.
   - **Furniture**: The sales for this category are 742,000, making it the second highest. The bar representing Furniture is shorter than Technology but longer than Office Supplies, coloured in a light brownish hue.
   - **Office Supplies**: This category has the lowest sales among the three, with 719,047. The bar for Office Supplies is the shortest and is coloured in light orange.

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

   Reason: Using a line chart to analyse monthly sales amounts over the course of a year is a great approach. A line chart effectively shows trends and patterns over time, making it easy to visualise how sales fluctuate month by month.

   Chart Showing Monthly Sales Over Year -
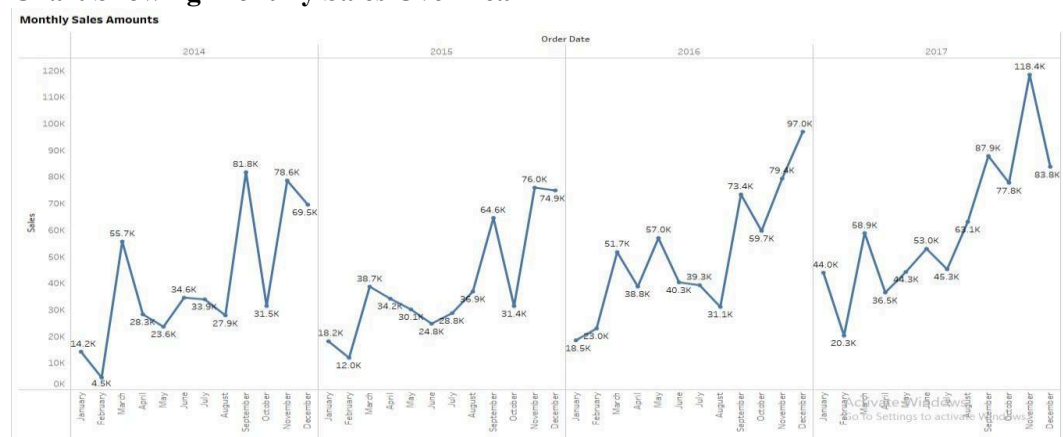
Here are the key insights from the chart:

- Each year shows distinct seasonal fluctuations with peaks typically in the later months.
- Sales appear to generally increase year over year, with the highest overall peak in November 2017 (118.4K).
- The data shows patterns that might indicate higher sales during certain months, suggesting seasonal or promotional influences.

This overview highlights the sales trends and notable peaks and troughs within the data, offering insight into seasonal sales patterns and year-over-year growth.

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

Reason: A bar chart is an excellent choice for visualising the distribution of total sales amounts among different product categories because it provides a clear, straightforward comparison of sales figures across categories, making it easy to identify the highest and lowest performing segments.

Chart Showing Total Sales By Category and Sub-Category -

Here are the key insights from the chart:

- **Phones** in the Technology category have the highest sales among all sub-categories.
- **Chairs** and **Storage** are the leading sub-categories in Furniture and Office Supplies, respectively.
- There is a wide variance in sales within each category, indicating differing product popularity.

This overview highlights the distribution of total sales across different product categories and sub-categories, providing insight into the leading and lesser-performing segments.

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

Reason: Using a line chart for each individual customer over the years is a good choice for visualising sales performance. It allows you to track trends and patterns

in sales over time for each customer, making it easier to identify growth, decline, or seasonality in their purchasing behaviour.

Chart Showing Sales Performance of Individual Customers Over Time-
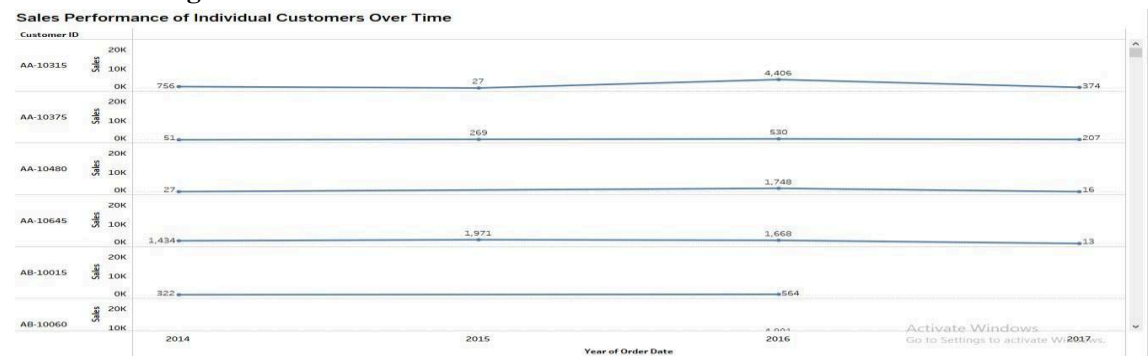
The analysis of **sales performance** for individual customers from **2014** to **2016** reveals distinct trends:

- **Customer AA-10315** saw a significant sales increase in 2015, peaking at around 4,406 units, followed by a sharp decline to 374 units in 2016.
- **Customer AA-10375** showed a gradual rise from 51 in 2014 to 530 in 2015, then decreased to 207 in 2016.

These insights underscore the varying sales trajectories among customers, highlighting opportunities for targeted strategies and the importance of monitoring and adapting to individual customer behaviours.

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

Reason: A bar chart is the right choice for displaying how sales vary based on different days of the week and product categories because it allows for easy comparison of sales figures across multiple categories and days, making patterns and differences immediately visible.

**Chart Showing Sales Per Day of the Week For Each Category -**
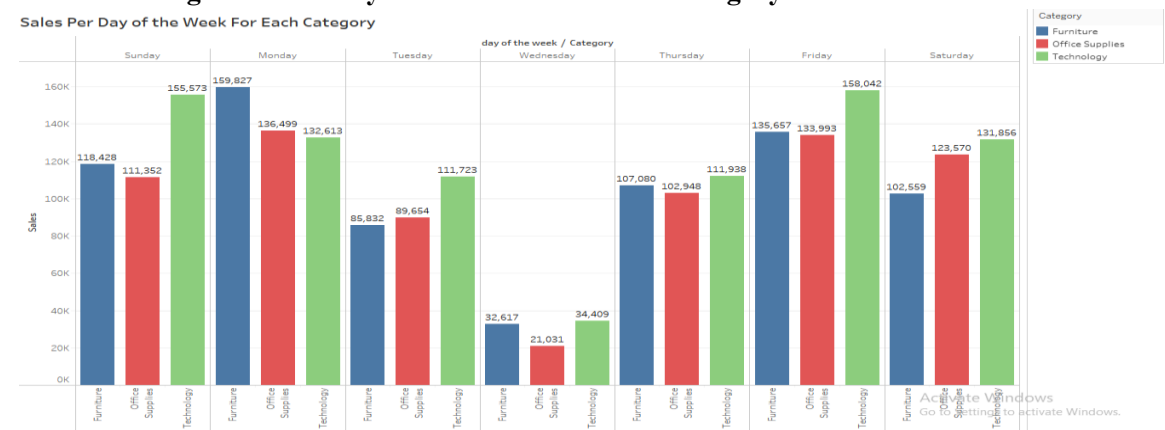
Here are the key insights from the chart:

- **Technology** consistently has the highest sales across most days, particularly on weekends.
- **Monday** and **Friday** are strong sales days for all categories, with Technology peaking on these days.
- **Wednesday** has the lowest sales across all categories, indicating mid-week is a slow period.
- **Sunday** sees higher sales in Technology, with Furniture and Office Supplies also performing well.
- **Tuesday** and **Thursday** show moderate sales, with Tuesday being slightly lower.
- **Saturday** maintains strong sales, especially in Technology and Office Supplies, while Furniture sees a slight dip compared to other strong days.

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

   Reason: A line chart is appropriate for visualising sales growth of different product categories over time because it effectively shows trends and patterns across multiple categories with the use of distinct lines and a legend for clear differentiation.

   **Chart Showing Sales Growth By Product Categories Over Time -**
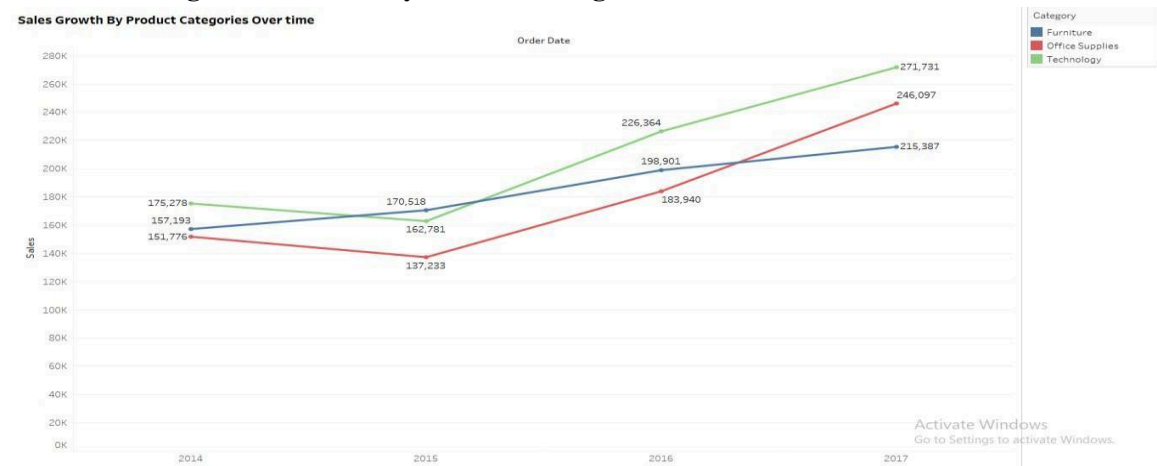
   Here are the key insights from the chart:

   - **Technology** is the category that has seen the most growth over time. Sales of technology products have grown from **137,233** in **2014** to **271,731** in **2017**.
   - Furniture and Office Supplies have also seen sales growth, but not to the same extent as Technology. Sales of furniture products have grown from **151,776** in **2014** to **198,901** in **2017**. Sales of office supplies have grown from **120,000** in **2014** to **162,781** in **2017**.

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

Reason: A bar chart is appropriate for visualising sales distribution across different regions in the "Superstore" dataset because it allows for easy comparison of sales volumes across discrete categories (regions).

Chart Showing Sales Across Different Regions -

Sales Across Different Regions
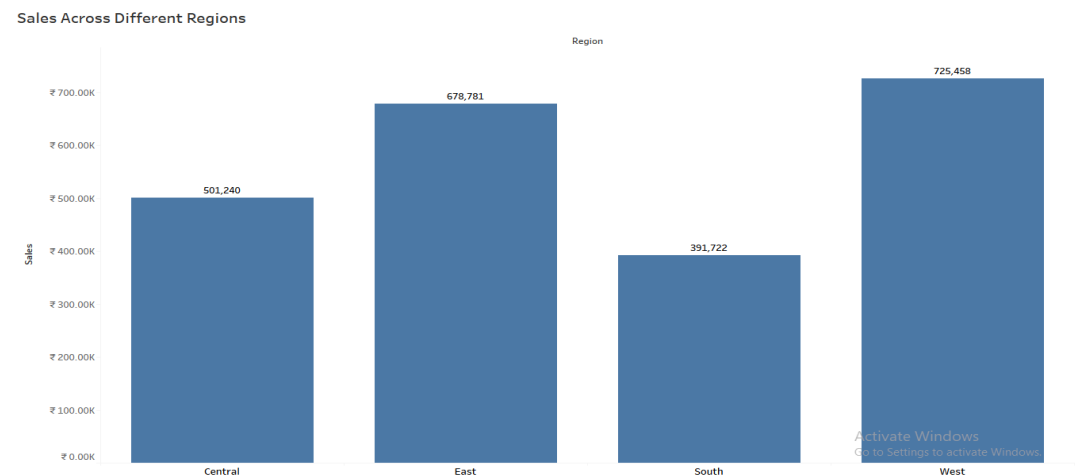
Here are the key insights from the chart:

- **West** leads with highest sales (725,458)
- **East** follows closely (678,781)
- **Central** in third place (501,240)
- **South** has lowest sales (391,722)

The sales data reveals significant regional disparities. The **West** and **East** regions are the clear frontrunners, with the West slightly outperforming the East. Central and South lag behind, with the South showing the weakest performance. This pattern suggests a need for targeted strategies to boost sales in underperforming regions, particularly the South, while maintaining the strong performance in the West and East.

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

Reason: A grouped bar chart is a suitable choice for visualising the composition of profits across various subcategories within different customer segments because it facilitates clear comparison and breakdown of profits within each segment.

Chart Showing Profit Composition By Sub-Category and Customer Segment -

Profit Composition by Subcategory and Customer Segment
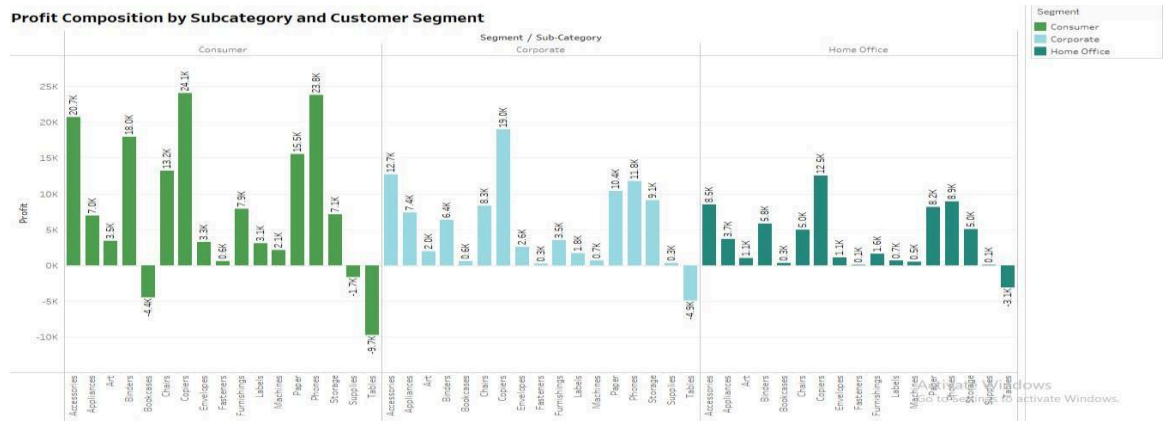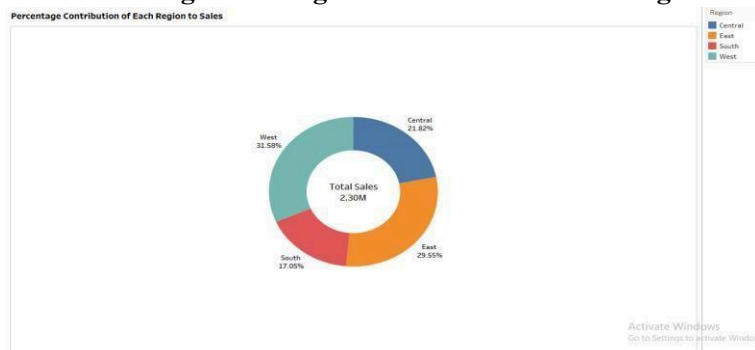
Here's a brief analysis:

- **Corporate** segment appears more stable, with positive profits across most categories. Copiers and Machines stand out as particularly profitable.
- **Home Office** segment has fewer categories but generally shows positive profits, with Copiers being notably profitable.
- **Copiers** and **Machines** tend to be highly profitable across all segments.
- **Furniture items** (Chairs, Tables) show mixed results, with some losses in the Consumer segment.
- The **Corporate** segment seems to have the most consistent positive profits across sub-categories.

This chart highlights the varying profitability of different product categories across customer segments, indicating areas of strength and potential improvement for each market segment.

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

Reason: A Donut chart is a suitable choice for visualising the percentage contribution of each region to overall sales because it effectively shows the proportional distribution of parts to a whole, making it easy to compare relative sizes at a glance.

\

Chart Showing Percentage Contribution of Each Region to Sales-

Percentage Contribution of Each Region to Sales

The pie chart provides a clear visualisation of the percentage contribution of each region to the overall sales, totaling 2.30 million.

Here is a brief analysis:

- **West Region**: The West region contributes the largest share, accounting for 31.58% of total sales. This indicates that the West is the most significant contributor to overall sales.
- **East Region**: The East region follows with a 29.55% contribution. This shows that the East is also a substantial contributor, just slightly behind the West.
- **Central Region**: The Central region contributes 21.82% to the total sales. While this is a significant portion, it is notably less than the contributions from the West and East regions.
- **South Region**: The South region has the smallest contribution, with 17.05%. This indicates that the South is the least significant contributor among the four regions.

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

Reason: A bar chart is a suitable choice for visualising the profit margins associated with different shipping modes and customer segments because it clearly displays comparisons between distinct categories, allowing for easy interpretation and analysis of differences in profit margins.

Chart Showing Profit Margins by Shipping Mode and Customer Segment-
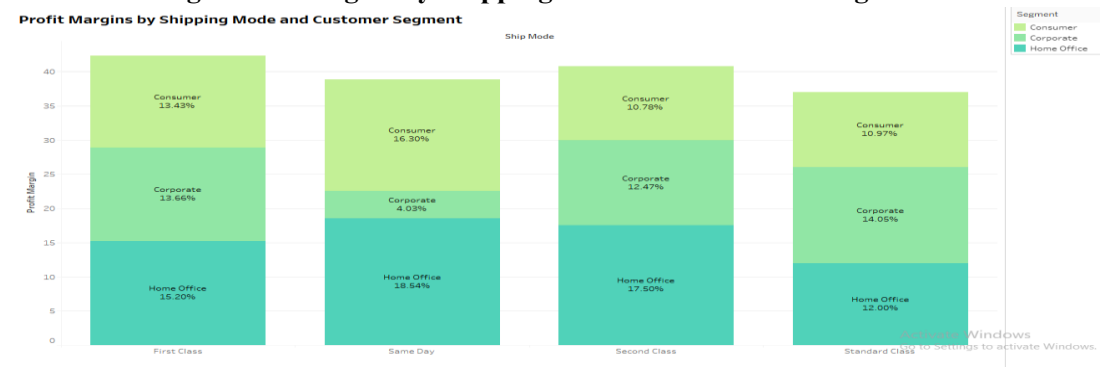
Here are the key insights from the chart:

- **Consumer** segment generally has higher margins across most shipping modes.
- **Home Office** segment shows significant variation, performing best with **Same Day shipping** but worst with **Standard Class**.
- **Corporate segment** has the most consistent margins across shipping modes, except for the low **Same Day** shipping margin.

The data suggests that pricing and profitability vary considerably depending on the combination of shipping mode and customer segment, indicating potential for optimization in pricing or service offerings.

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

Reason: To visualise the processing times for different product categories, a bar chart would be suitable. Bar charts are effective for comparing discrete categories, such as different product categories and their respective processing times. They allow you to easily see variations in processing times across categories.

Chart Showing Average Order Process Time By Product Categories-

Average Order Process Time By Product Categories

Office Supplies
3.98 days

Technology
3.92 days

Furniture
3.92 days

Here are the key insights from the chart:

- The chart depicts the **average order processing times** for three product categories:
- Office Supplies: 3.98 days
- Technology: 3.92 days
- Furniture: 3.92 days
- Processing times show minimal variation across categories, with less than a 0.1-day difference between them.
- **Office Supplies** have the longest processing time, although marginally.
- **Technology** and **Furniture** share identical processing times at **3.92** days each.

These findings suggest a consistent and standardised approach to order handling across different product categories, indicating efficient operational management with minimal deviation in processing efficiency.

12. How do discounts affect overall profit?

Reason: Using a bar chart to compare profits with and without discounts is a suitable choice. It allows for a clear visual comparison of how discounts influence overall profitability

Chart Showing Impact of Discounts on Profitability -
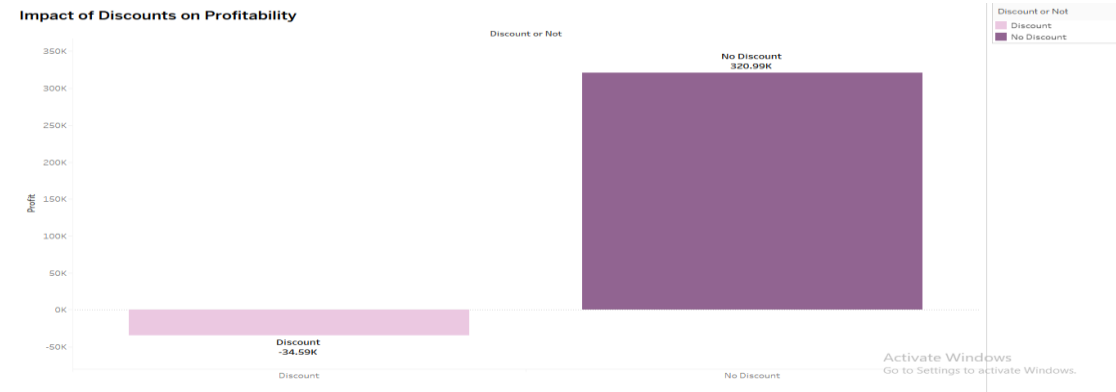
This chart illustrates the impact of discounts on profitability:

- **No Discount:** Generates a profit of $320,990.
- **Discount:** Results in a loss of $34,590.
- The difference in profitability between discounted and non-discounted sales is substantial.
- Discounts appear to significantly erode profitability, turning what would be a profitable sale into a loss.
- The chart suggests that the company's discount strategy may need review, as it's currently negatively impacting overall profitability.

This data indicates that the current discount approach is not sustainable and may be harming the company's financial performance.

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

Reason: Using a scatter chart with categories as a legend is appropriate for visualising the relationship between product sales and profitability across different product categories.

Chart Showing Relationship Between Sales and Profitability Across Categories-
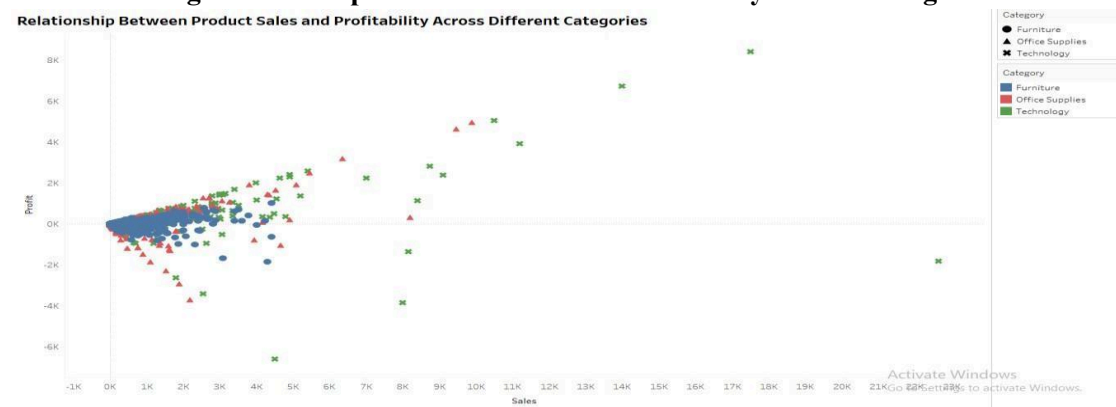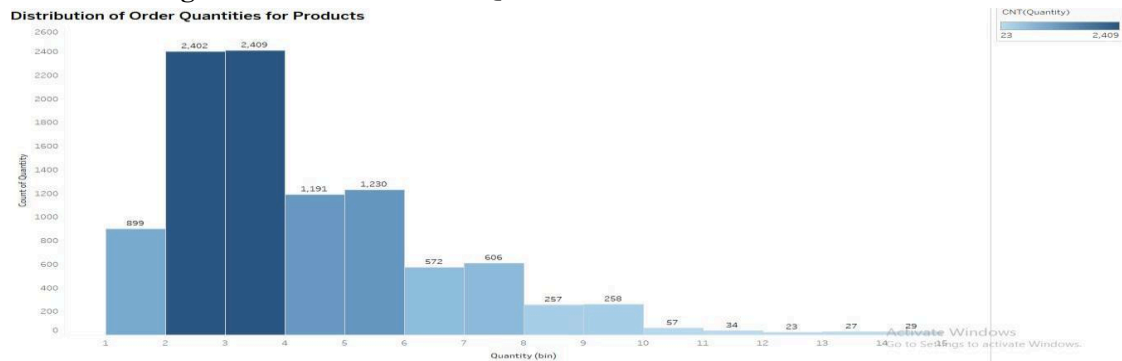

Here are the key insights from the chart:

- Positive correlation: Higher sales generally correspond to higher profits across all categories.
- Category differences:
  - Technology: Wide range of sales and profits, including top-profit items.
  - Office Supplies: Lower sales with moderate profits.
  - Furniture: Lower to moderate sales with varied profitability.
- Some products show negative profits, particularly at lower sales levels.
- Certain technology products exhibit exceptionally high sales and profits.
- Most products fall within lower sales ranges (below 5K), with varying profitability.

These findings highlight varied category performance, with technology offering high potential returns alongside greater variability.

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

Reason: Using a histogram chart to visualise the distribution of order quantities for products in the dataset is a good choice. Histograms are effective for showing the frequency distribution of numerical data, such as order quantities, by grouping them into bins.

Chart Showing Distribution of Order Quantities for Products -

Key insights from the chart:

- Order quantity distribution: Most orders are for smaller quantities, with a peak at 2-3 items per order.
- Frequency drops sharply for larger quantities: Orders of 8 or more items are much less common.
- Highest frequency: Orders of 2 and 3 items are the most common, with over 2,400 orders each.
- Gradual decline: There's a steady decrease in frequency as order quantity increases from 4 to 7 items.
- Low volume orders: Single-item orders are relatively common (899 orders), but not as frequent as 2-3 item orders.

- Rare large orders: Orders of 10 or more items are infrequent, with each category having fewer than 60 orders.

This distribution suggests customers typically prefer smaller order sizes, which could impact inventory management and shipping strategies.

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

Reason: Using a boxplot chart because it effectively displays the spread and central tendency of profit distributions across different product categories, highlighting medians, quartiles, and potential outliers.

Chart Showing Variation in Profit Distribution Across Categories -


From analysing the chart, a few key insights emerge:

- Technology products show the highest median profit and widest profit range, suggesting both higher potential returns and risks.
- Furniture has the lowest median profit but a relatively wide distribution, indicating variable performance across products.
- All categories have some negative profit items, but Technology seems to have the fewest low-end outliers.
- The profit distributions are not symmetrical, with each category showing different skews and outlier patterns.

These insights provide a quick overview of the relative profitability and variability across the three product categories.

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

Reason: Using a histogram by different ship modes because it allows you to compare the distribution of shipping times for each shipping mode, highlighting any differences or patterns in their delivery durations.

Chart Showing Comparison of Shipping Time Distribution Across Shipping Mode -

Comparison of Shipping Time Distributions Across Different Shipping Modes
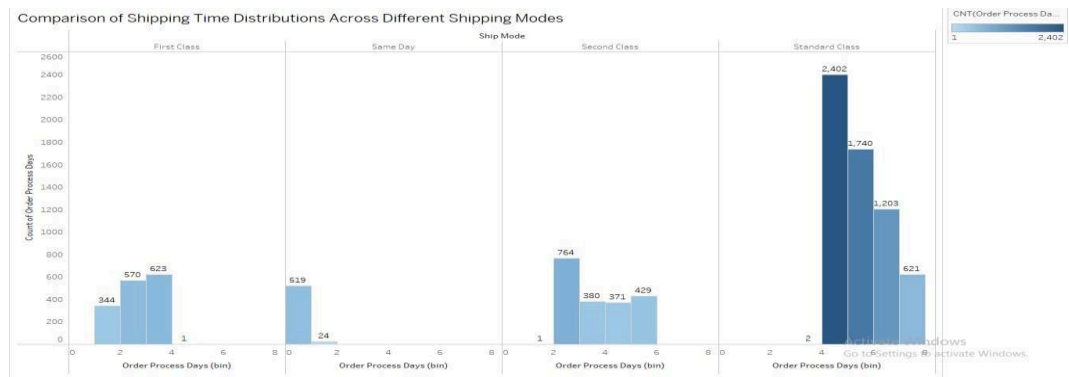
Here are some key observations:

- **First Class**: The majority of orders are processed within 2 to 4 days, with a small number extending up to 6 days.
- **Same Day**: Nearly all orders are processed on the same day, indicating a very quick turnaround.
- **Second Class**: Most orders are processed within 2 to 4 days, with fewer orders taking up to 6 days.
- **Standard Class**: The processing time distribution is broader, with a peak at 2 days and significant counts up to 6 days, indicating a more variable processing time.

Overall, the chart effectively highlights the differences in processing times for each shipping mode, with Same Day being the fastest and most consistent, while Standard Class shows the most variation.

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

Reason: A line chart effectively shows the monthly trend in the number of orders shipped by illustrating changes over time and highlighting patterns or fluctuations in the data.

Chart Showing Monthly Trend in the Number of Orders Shipped -

Monthly Trend in the Number of Orders Shipped
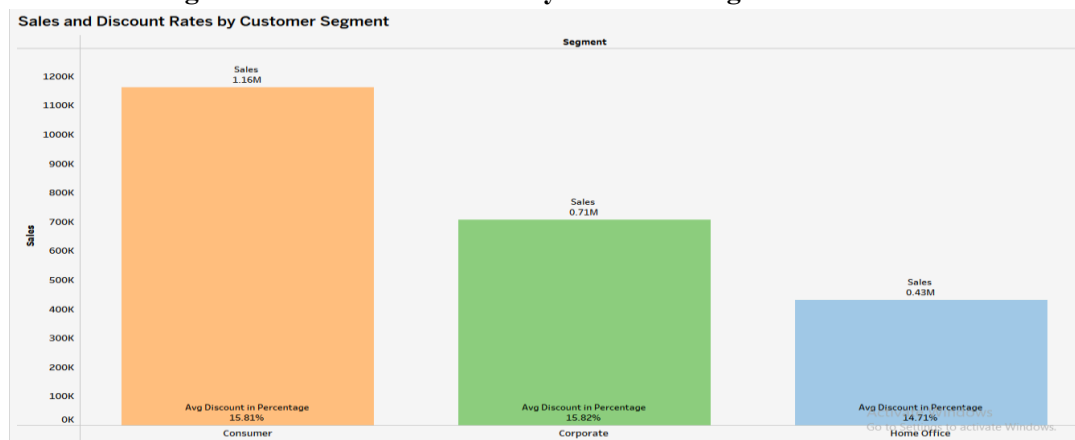
Here are some key observations:

- A low starting point in January with 381 orders, dipping further in February to 300 orders.
- A sharp increase in March, reaching 696 orders.
- Relatively stable numbers from April to August, fluctuating between 668 and 735 orders.
- A notable spike in September and October, peaking at 1,383 and 1,471 orders respectively.
- A slight decline in November and December, with 1,408 orders in December.

These trends suggest seasonal variations and potential peak periods for orders shipped.

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

Reason: Using a bar chart to visualise the performance of different customer segments in terms of sales and discount rates is a correct choice because it allows for clear comparison between segments, making it easy to see differences in sales and discount rates across categories.

Chart Showing Sales and Discount Rates by Customer Segment -
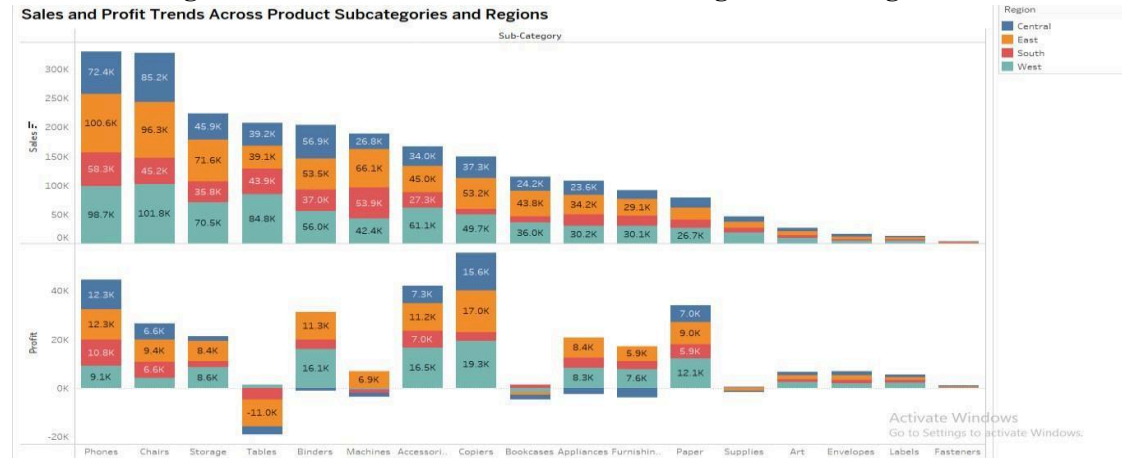

Here are some key observations:

- **Consumer Segment**: Highest sales at 1.16 million with an average discount rate of 15.81%.
- **Corporate Segment**: Sales of 0.71 million with an average discount rate of 15.82%.
- **Home Office Segment**: Sales of 0.43 million with an average discount rate of 14.71%.

The data indicates that while the Consumer segment generates the highest sales, it also receives a comparable discount rate to the Corporate segment. The Home Office segment has the lowest sales and the lowest average discount rate.

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

Reason: A bar chart to visualise sales and profit trends across different product subcategories and regions is an excellent choice because it clearly displays comparisons between multiple categories, making it easy to discern patterns and trends in the data.

Chart Showing Sales and Profit Trends Across Sub-Categories and Regions -
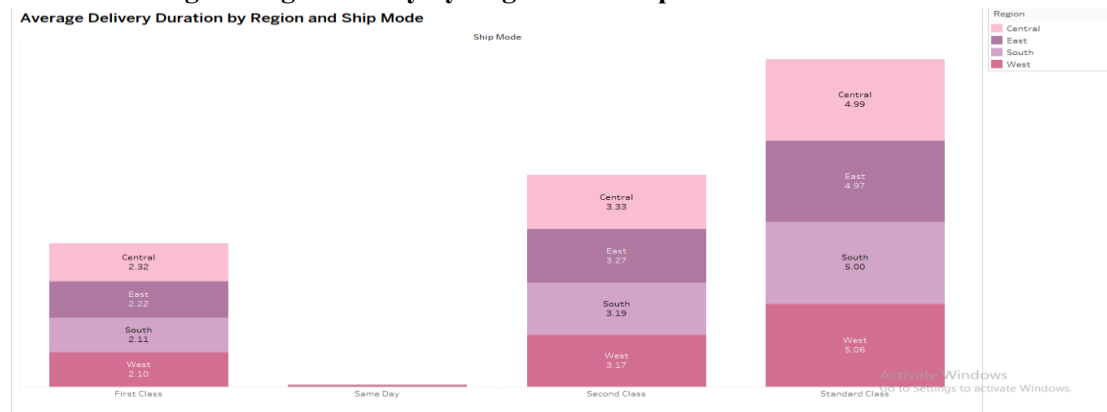

Here are some key observations:
- Phones and chairs are the top-selling subcategories, with the highest sales bars.
- The West region consistently shows strong sales across many subcategories.
- Profit doesn't always correlate with sales. For example, copiers have relatively lower sales but higher profit.
- Tables show negative profit despite moderate sales, indicating potential issues with pricing or costs.
- Some subcategories like fasteners and labels have very low sales and profits compared to others.
- The Central and East regions often have smaller contributions to both sales and profit across categories.
- Paper products show a notable discrepancy between sales volume and profit margin.

This overview suggests varying performance across product lines and regions, with some categories being more profitable than others despite lower sales volumes.

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

Reason: Using a bar chart to visualise the average delivery duration for different regions and ship modes is appropriate because it allows for clear comparison of numerical values across categories (regions and ship modes) using bar lengths, making it easy to identify trends and differences visually.

Chart Showing Average Delivery by Region and Ship Mode -
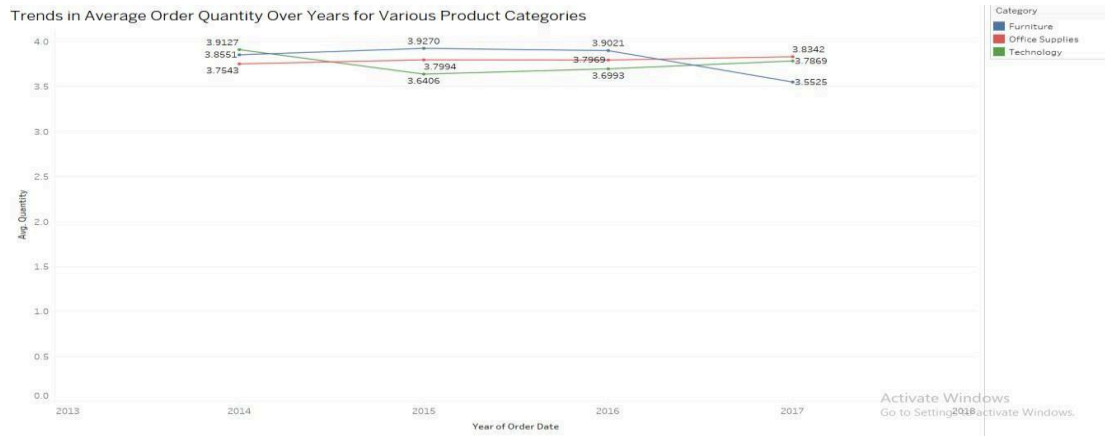

Here are some key observations:
- There are four shipping modes: First Class, Same Day, Second Class, and Standard Class.
- Four regions are compared: Central, East, South, and West.
- Standard Class has the longest delivery times across all regions, ranging from 4.57 to 5.06 days.
- First Class is the fastest mode after Same Day, with durations between 2.10 and 2.32 days.
- Same Day shipping appears to have no data or is instantaneous (shown as a thin line).
- The West region tends to have slightly longer delivery times in most shipping modes.
- Delivery durations increase as the shipping class changes from First to Second to Standard.

This visualisation allows for easy comparison of delivery efficiency across regions and shipping methods.

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

Reason: Using a line chart to visualise changes in average order quantity over the years for different product categories is appropriate. A line chart is effective for showing trends over time, making it suitable for comparing how average order quantities have evolved across various categories from year to year.

Chart Showing Trends in Average Order Quantity Over Years for Categories -

Trends in Average Order Quantity Over Years for Various Product Categories
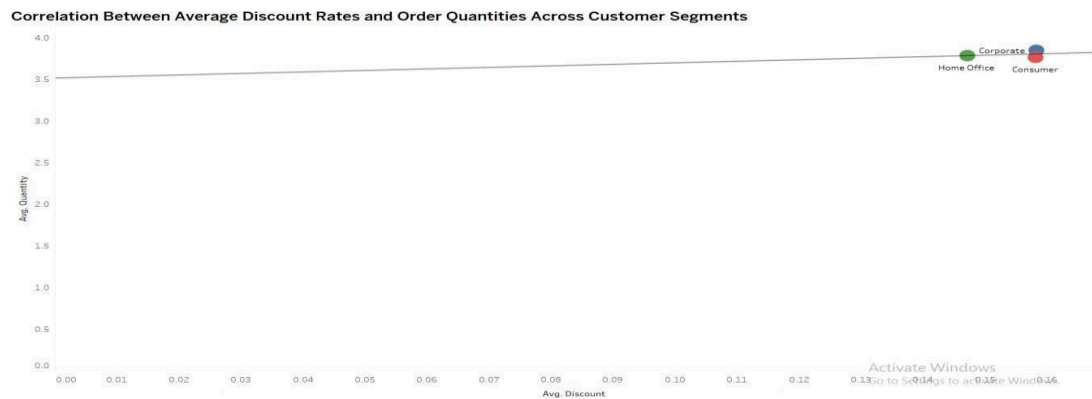
Here are some key observations:

- All three categories show relatively stable trends, with minor fluctuations over the years.
- Technology products consistently have the highest average order quantity, ranging from about 3.8 to 3.9 units.
- Furniture orders show a slight overall decline, starting at 3.9127 units in 2013 and ending at 3.5525 units in 2017.
- Office Supplies maintain the lowest average order quantity, hovering around 3.6 to 3.8 units.
- The lines for different categories intersect at various points, indicating changing relative positions in average order quantities.
- There are no dramatic increases or decreases for any category over the five-year period.
- By 2017, Technology has the highest average order quantity, followed closely by Office Supplies, with Furniture showing the lowest.

This visualisation allows for easy comparison of ordering trends across the three product categories over time.

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

Reason: Using a scatter chart to visualise the correlation between discount rates and order quantities for different customer segments is appropriate because scatter charts effectively show the relationship between two continuous variables (discount rates and order quantities) and can easily incorporate segmentation by customer segments using colour or shape encoding.

Chart Showing Correlation Between Average Discount Rates and Order Quantities -

Correlation Between Average Discount Rates and Order Quantities Across Customer Segments

Here are some key observations from the chart:

- **Average Discount**: The x-axis represents the average discount rate, which ranges from 0.00 to approximately 0.16.
- **Average Quantity**: The y-axis represents the average order quantity, which ranges from 0.0 to 4.0.
- **Trend**: A slight positive correlation is observed between the average discount rate and average order quantity across the customer segments, indicated by the upward sloping trend line.
- **Segment Positioning**:
  - Corporate: Located at a higher average discount rate and slightly higher average quantity.
  - Consumer: Positioned close to the Corporate segment with a similar average discount rate but slightly lower average quantity.
  - Home Office: Situated at a lower average discount rate and lower average quantity compared to Corporate and Consumer segments.

Overall, there is a weak positive correlation between the discount rates and order quantities across the segments

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

Reason: Using a donut chart to visualise the proportion of orders returned in each region within the Superstore dataset is a suitable choice. This is because a donut chart effectively displays the relative proportions of categorical data, allowing for a clear comparison of the returned orders across different regions.

Chart Showing Proportion of Orders Returned by Region -

Proportion of Orders Returned by Region
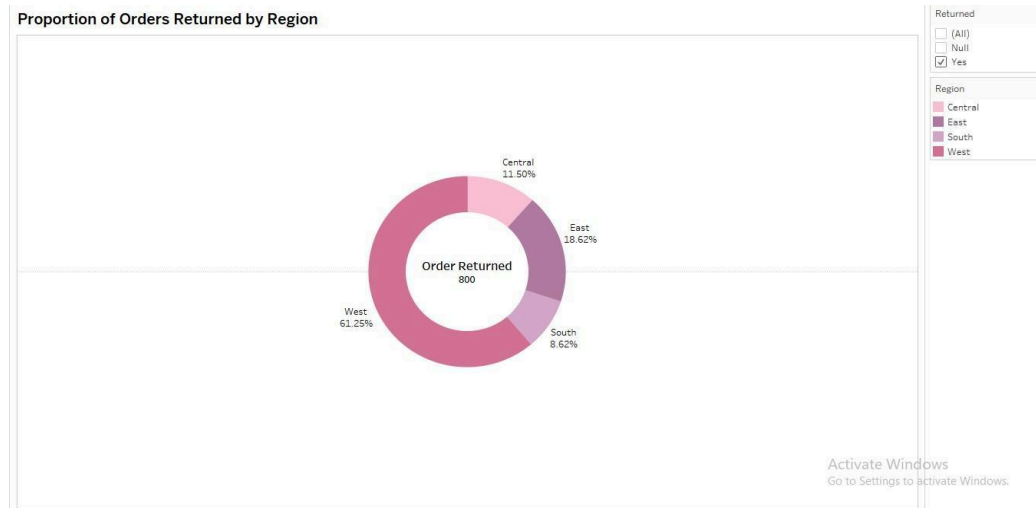
The chart displays the proportion of orders returned by region, with a total of 800 orders returned.

- **West:** Highest proportion of returns at 61.25%.
- **East:** Second highest proportion of returns at 18.62%.
- **Central:** Accounts for 11.50% of returns.
- **South:** Lowest proportion of returns at 8.62%.

The majority of returned orders come from the West region, making up over half of the total returns. The East and Central regions have moderate return rates, while the South region has the least returns.

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

Reason: Using a stacked bar chart to compare the profit of different products for different subcategories is an appropriate choice. It effectively shows the total profit for each subcategory while also allowing for a detailed comparison of the contribution of individual products within those subcategories.

Chart Showing Profit Comparison of Products by Sub-Categories -

Profit Comparison of Products by Subcategories
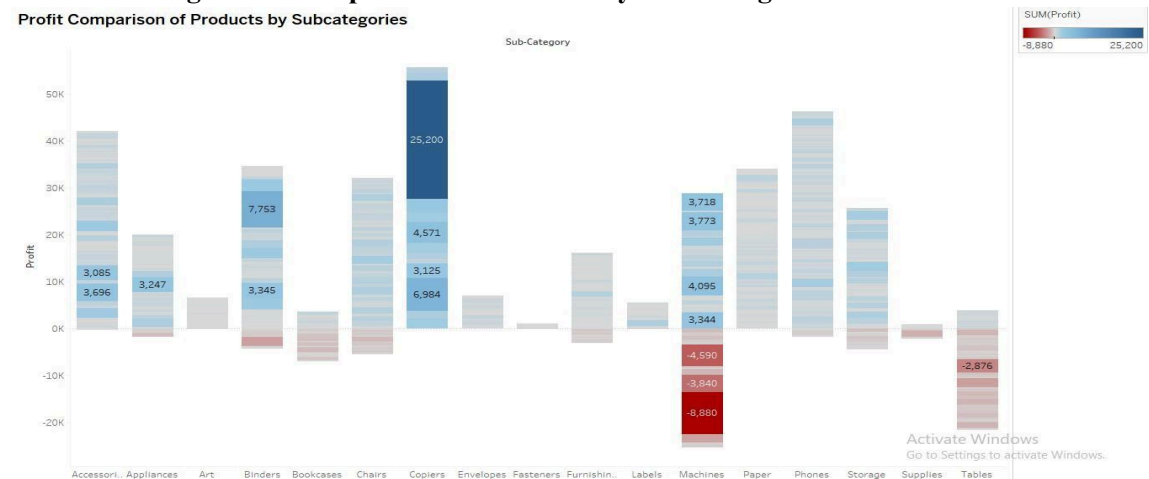
Here are some key observations from the chart:

- **High Profit Subcategories**:
    - **Chairs**: Highest profit at 25,200.
    - **Phones**: Around 40,000.
    - **Copiers**: Approximately 31,000.
- **Moderate Profit Subcategories**:
    - **Accessories, Appliances, Binders, and Furnishings**: Profits ranging from 3,000 to 8,000.
    - **Bookcases, Envelopes, Fasteners, Storage, and Supplies**: Profits around 1,000 to 3,500.
- **Loss-Making Subcategories**:
    - **Machines**: Largest loss at -8,880.
    - **Tables**: Notable loss at -2,876.
    - **Art and Labels**: Losses around -3,000 to -4,000.

While certain subcategories like Chairs, Phones, and Copiers are highly profitable, others such as Machines and Tables incur significant losses. Profits and losses vary widely across different product subcategories.

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

Reason: Using a Treemap to visualise the most commonly used shipping mode in the Sample Superstore data set is a good choice. Treemaps effectively display hierarchical data and allow for easy comparison of the proportions of different categories, making it straightforward to identify the most commonly used shipping mode.

Chart Showing Most Commonly Used Shipping Mode -
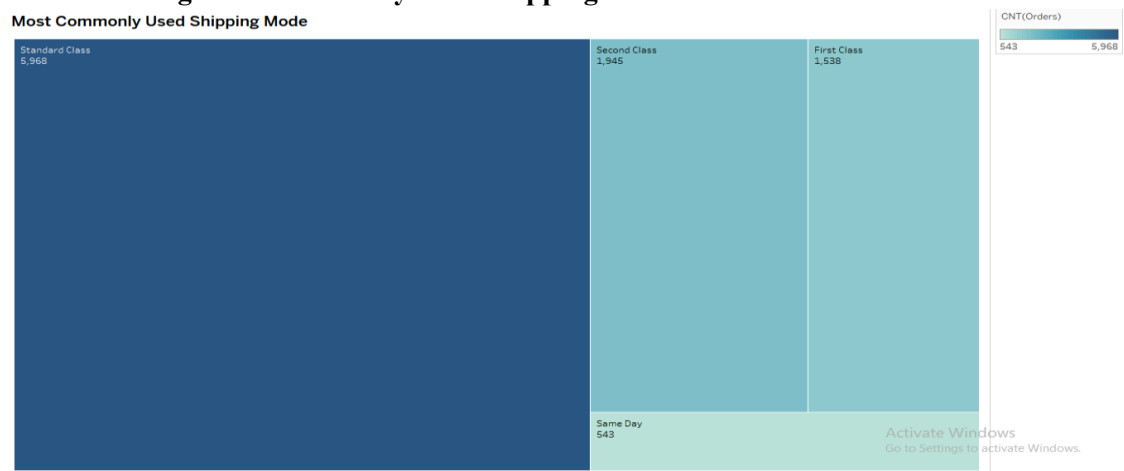
The treemap visualisation displays the most commonly used shipping modes in the dataset.

- **Standard Class**: 5,968 orders (largest segment)
- **Second Class**: 1,945 orders
- **First Class**: 1,538 orders
- **Same Day**: 543 orders (smallest segment)

The chart highlights the dominance of Standard Class shipping among customers.

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

Reason: Using a line chart to visualise the sales performance of different regions throughout quarters of a year is appropriate. A line chart is effective for showing trends and variations over time, making it suitable for comparing how sales evolve across different regions over each quarter of the year.

Chart Showing Quarterly Sales Performance Evolution Across Regions -
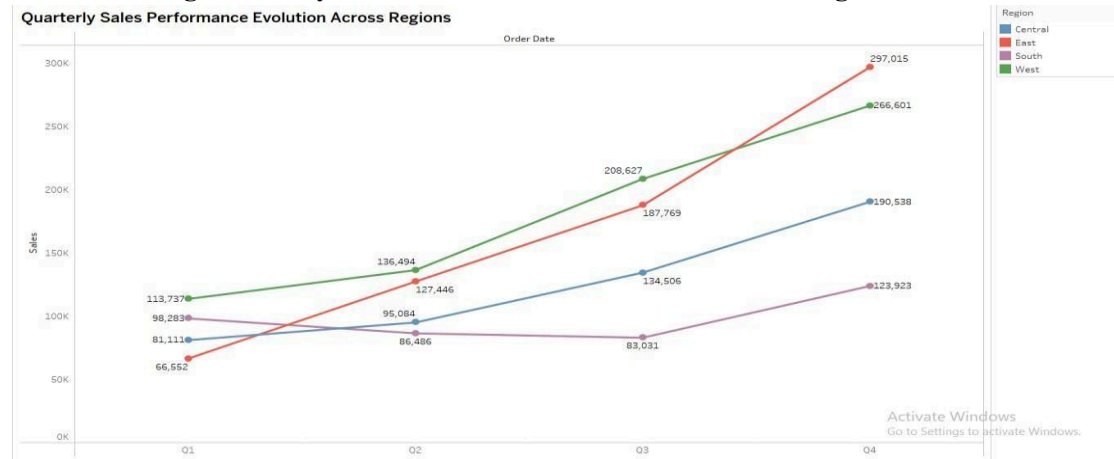
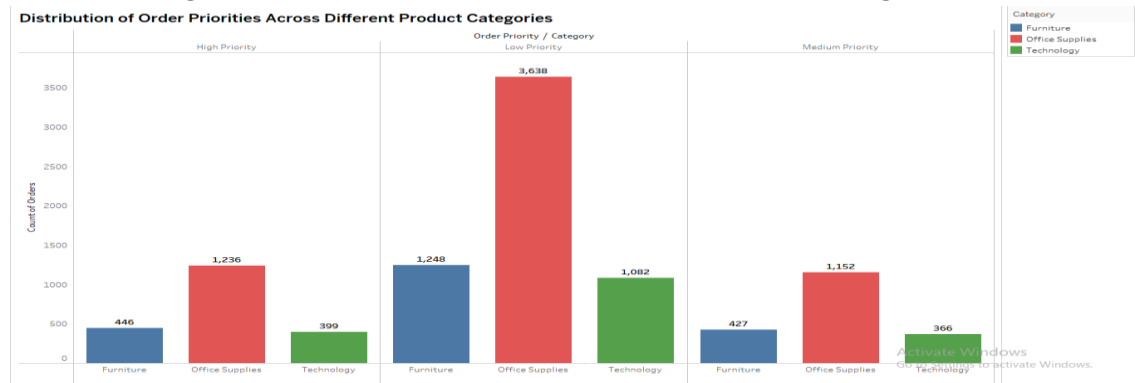Here are some key observations from the chart:

- **East** region shows the most significant growth, ending with the highest sales at 297,015 in Q4.
- **West** region follows, with steady growth throughout the year, reaching 266,601 in Q4.
- **Central** region also shows consistent growth, ending with 190,538 in Q4.
- **South** region displays fluctuating performance, with a dip in Q3 but an overall increase, ending with 123,923 in Q4.

This analysis highlights that the East and West regions lead in sales, while the South region shows the least growth.

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

Reason: Using a bar chart to visualise the distribution of order priorities across different product categories is a suitable choice. Bar charts are effective for comparing categorical data, making it easy to see the frequency or distribution of order priorities within each product category at a glance.

Chart Showing Distribution of Order Priorities Across Different Categories-

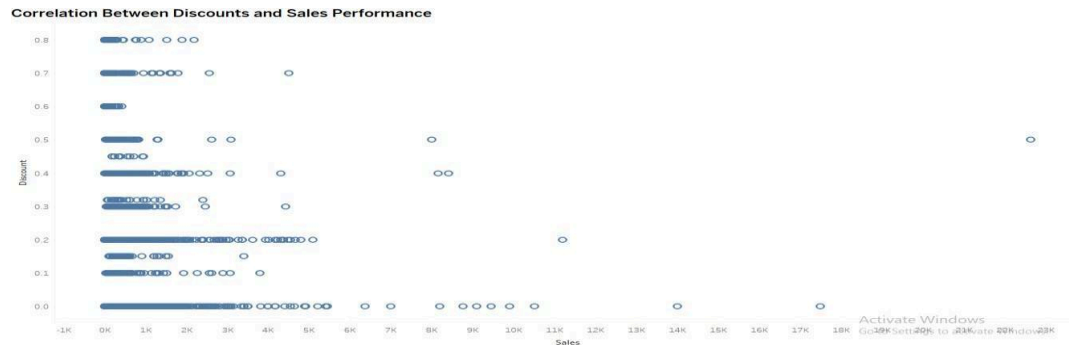Here are some key observations from the chart:

- **High Priority Orders**:
    - Office Supplies: 1,236 orders
    - Furniture: 446 orders
    - Technology: 399 orders
- **Low Priority Orders**:
    - Office Supplies: 3,638 orders (highest)
    - Furniture: 1,248 orders
    - Technology: 1,082 orders
- **Medium Priority Orders**:
    - Office Supplies: 1,152 orders
    - Furniture: 427 orders
    - Technology: 366 orders

This analysis shows that Office Supplies dominate across all priority levels, especially in low priority orders. Furniture and Technology categories have a smaller share of orders across all priority levels.

28. What is the relationship between discounts and sales?

Reason: Using a scatter plot to visualise the relationship between discounts and sales is a good choice because it allows you to clearly see the correlation between the two variables, identify trends, and detect any outliers or patterns in the data.

Chart Showing Correlation Between Discounts and Sales Performance -

Correlation Between Discounts and Sales Performance

Here are some key insights from the chart:

- **Concentration of Discounts**: There is a high concentration of data points at the lower discount levels (0% to 20%). This suggests that most transactions have smaller discounts.
- **Sales Spread**: Sales values are widely spread across all discount levels, with some outliers reaching high sales figures even at lower discount levels. However, higher discounts (above 30%) have fewer data points, indicating they are less common.
- **Higher Discounts and Sales**: At discount levels above 40%, there are only a few data points, and these show a mixed performance in terms of sales. Some transactions achieve high sales, while others do not, suggesting that higher discounts do not consistently drive higher sales.
- **Outliers**: There are a few outliers with high sales figures at various discount levels.
- **Overall Trend**: The overall trend is not immediately clear from the scatter plot, suggesting that the correlation between discounts and sales performance might be weak or nonlinear.

In summary, while lower discounts are more common and widely spread in sales, higher discounts are less frequent and show mixed sales performance. The correlation between discounts and sales performance is not straightforward and may require further analysis to uncover deeper patterns.

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

Reason: Using a bar plot to visualise the difference in average order value between repeat customers and new customers is a good choice because it clearly shows the comparison of these two distinct categories, making it easy to see any differences in their average order values at a glance.

Chart Showing Average Order Value Comparison: Repeat vs. New Customers -

Average Order Value Comparison: Repeat vs. New Customers
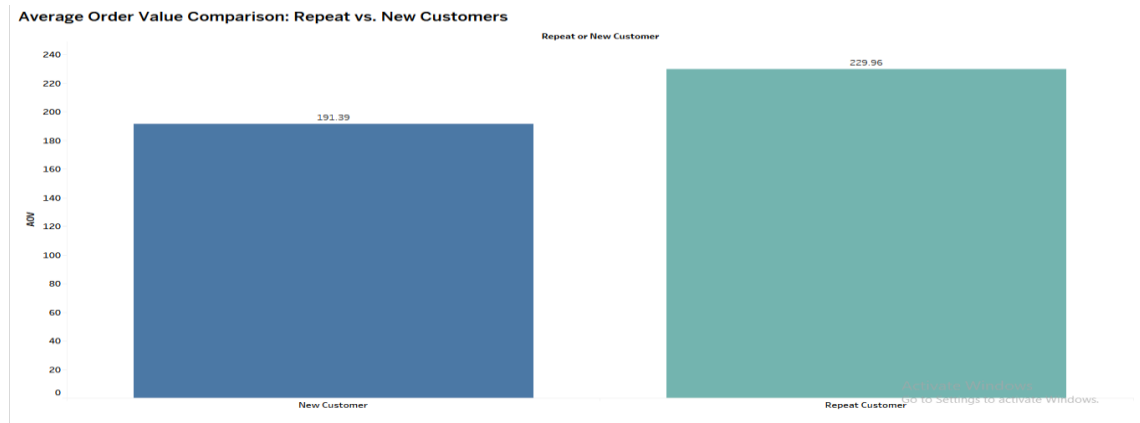
Here are the key insights:

- **Higher AOV for Repeat Customers**: Repeat customers have a higher average order value of 229.96 compared to new customers, who have an average order value of 191.39.
- **Repeat Customer Value**: The significant difference indicates that repeat customers tend to spend more per order than new customers.

This suggests that encouraging repeat business could be a valuable strategy for increasing overall sales.

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

Reason: Using a filled map to visualise the geographical distribution of returns and its impact on overall profitability is appropriate because it effectively illustrates regional variations, making it easy to identify areas with higher return rates and understand their influence on profitability across different locations.

Chart Showing Distribution of Returns and Profitability Impact -

Geographical Distribution of Returns and Profitability Impact
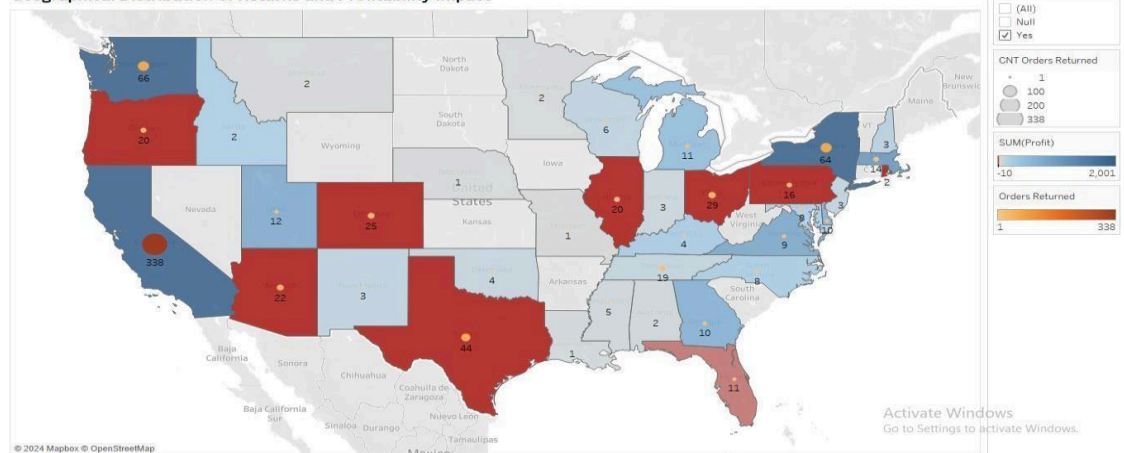
Here are some key insights:

- **Returns by State:**
  - The largest number of returns is from California (338 returns).
  - Texas (44), Illinois (29), and New York (64) also show significant numbers of returns.
  - States like Wyoming, Montana, and others show minimal or no returns.
- **Profitability Impact:**
  - The colour shading indicates the profit impact, with blue showing positive profit and red showing negative profit.
  - California, despite having the highest returns, still shows a positive profit impact.
  - States like Texas and Illinois show a significant number of returns and negative profit impact.
- **Return Magnitude:**
  - The size of the circles represents the number of orders returned. Larger circles indicate more returns.
  - California and Texas have the largest circles, indicating the highest return volumes.
- **Regional Insights:**
  - The West Coast (California, Oregon) and some parts of the Midwest (Illinois, Ohio) have higher returns.
  - The Southeast (Georgia, Florida) and Northeast (New York, Pennsylvania) show a mix of returns but with varying profit impacts.
- **Profit vs. Returns:**
  - States with fewer returns generally show better profitability.
  - States with high returns, such as Texas and Illinois, show poorer profitability.

Overall, this analysis provides a clear picture of how returns are distributed across the U.S. and how they impact profitability, helping in identifying areas that may need attention for improving returns management and profitability.