**Data Visualization and Analysis Project**

In this project we worked on “***Superstore dataset”*** which contains customer order details for a superstore over a period of four years, from 2014 to 2017. It includes information on the products sold, the customers who made the purchases, the order dates, and the sales and profit amounts. This dataset consists of three tables:  
  
***Orders Table:*** It contains 9995 rows and 21 columns (Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, Profit).  
  
***Returns Table:*** It contains 297 rows and 2 columns (Returned, Order ID).  
  
***People Table:*** It contains 4 rows and 2 columns (Person, Region).

**Dataset Link:**

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

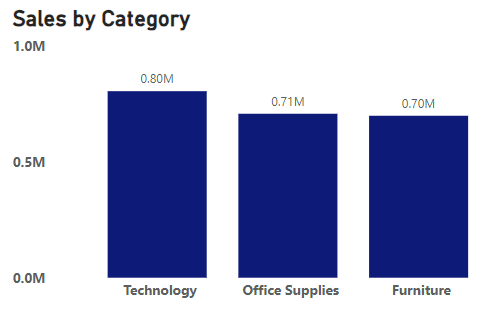
**Visualizing Insights: A Graphical Approach to 30**

**Superstore Scenarios**

I tackled 30 scenario-based questions using the Superstore Dataset, breaking down each one with helpful graphs. The graphs not only make it easy to understand but also add a nice touch to the explanation.

**Scenario-Based Questions:**

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

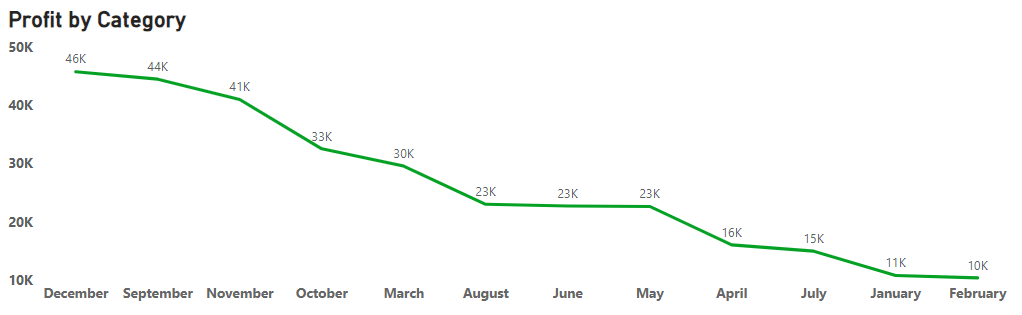
****

By analyzing the chart, it becomes evident that technology products have the highest potential total sales at 0.80 million, followed by office supplies at 0.71 million. Furniture, which exhibits the lowest total sales compared to technology and office supplies, is valued at 0.70 million in the "Superstore" dataset.

**Why did I pick the specific chart?**

The chosen chart type is a horizontal bar chart, as it effectively conveys the data's categorical distribution, from maximum potential sales down to the least. Additionally, the chart type is easy to read and interpret.

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

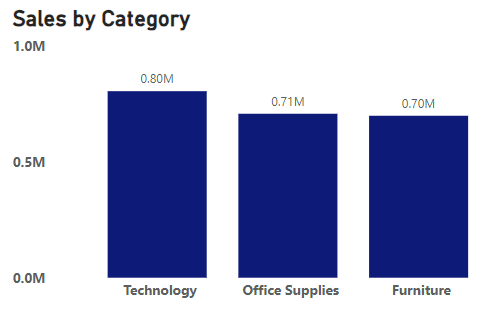


From the above line chart, we can observe a decreasing trend in sales throughout the year. This suggests that sales were higher in the last months (December, September, and November) and then decreased gradually over the preceding months (April, July, January, and February). In December, sales reached a peak and started to decline.

**Why did I pick the specific chart?**

The chosen chart type is the line chart because it provides a clear and concise way to visualize the sales trend over the course of a year, making it an ideal choice for this analysis. By comparing the sales data with the corresponding month, we can easily identify patterns and trends that may otherwise go unnoticed.

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



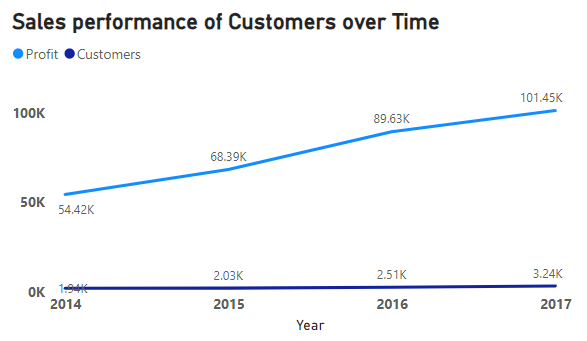
The total sales amount is distributed among different product categories as follows:

* Furniture: 800,000 (0.80M)
* Technology: 710,000 (0.71M)
* Office Supplies: 700,000 (0.70M)

**Why did I pick the specific chart?**

For this analysis, a horizontal bar chart was chosen as the best visual representation due to its simplicity and clarity. This chart type is well-suited for displaying a comparison between different categories or products, making it ideal for understanding the distribution of sales by category.

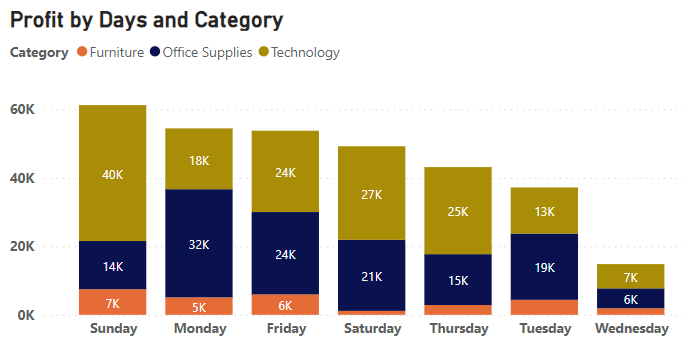
**4. Can we analyze the sales performance of individual customers over time?**



Yes, it is possible to analyze the sales performance of individual customers over time. Here, I have chosen a line chart for this analysis, and I have derived the following insights from the chart:  
  
**Increasing Sales Performance Over Time:** The line chart clearly illustrates that the sales performance of customers has been consistently increasing over the past four years. This indicates that the business is growing, achieving higher sales volume annually.  
  
**Steady Profitability:** The line chart also demonstrates a steady increase in the profit made by customers. This indicates that the business is maintaining a consistent profit margin despite the growth in sales volume.

**Why did I pick the specific chart?**

The specific chart used, a line chart, was chosen to effectively communicate the trends in sales performance and profitability over time. A line chart allows for the visual representation of these trends in a simple and easy-to-understand format, making it ideal for conveying insights about the historical performance of Customers.

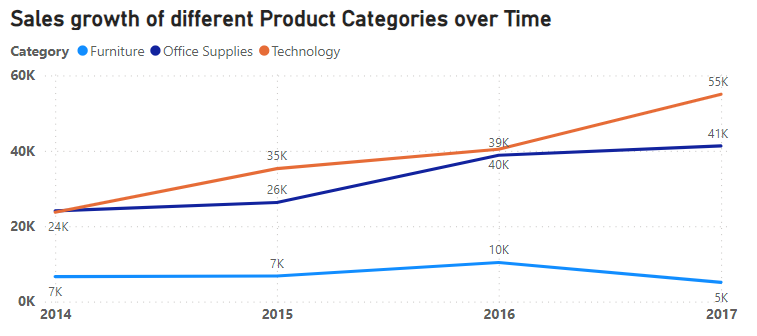
**5. How do sales vary based on different days of the week and product categories?**

The stacked bar chart reveals compelling insights into weekly sales patterns. On Sundays, the highest sales are observed, notably driven by the technology category with sales reaching an impressive 40k. The overall Sunday sales surpass 60k. Following closely, Monday ranks as the second-highest sales day, with the office supplies category leading at 32k. In contrast, Wednesdays consistently show the lowest weekly sales, dipping below 20k.

**Why did I pick the specific chart?**

Stacked bar charts are chosen for the analysis as they provide a clear visual representation of the distribution of data across categories. By using stacked bar charts, we can easily compare the total sales across days and visualize the profit contribution from each category within a specific day.

**6. Can we visualize the sales growth of various product categories over time?**

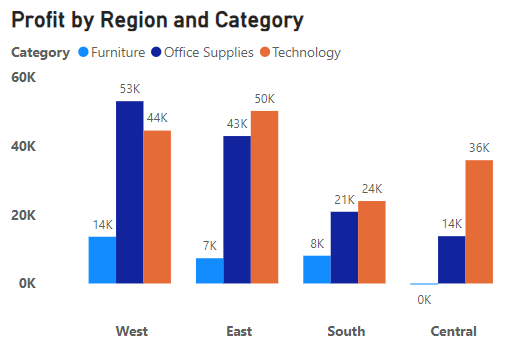


Yes, visualizing the sales growth of different product categories over time is a common and insightful analysis. Here, I used a line chart for the analysis, and we observe that in the year 2017, the highest sales occurred in the technology category, reaching 55k, followed by the second-highest sales in office supplies at 41k. The lowest sales were attributed to the furniture category, amounting to 5k. Additionally, the sales by category show a consistent increase from year to year, with the exception of furniture sales, which decreased by 5k in the year 2017.

**Why did I pick the specific chart?**

A line chart is ideal for visualizing sales growth over time in different product categories. It effectively displays trends, allowing easy comparison between categories with distinct lines. The chart's simplicity highlights chronological sales progression and facilitates quick identification of peak periods or shifts in performance. Overall, it provides a clear, concise, and visually impactful representation of sales dynamics.

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



From the above stacked bar chart, we observe that in the West region, the distribution of categories by sales is the highest. The office supply category archives 53k in sales, the technology category attains 44k in sales, and the furniture category reaches 14k in sales. The second-highest sales region is the East region. However, in the Central region, there is a decrease in sales, particularly in the office supplies and furniture categories.

**Why did I pick the specific chart?**

To analyze the sales distribution across different regions in the "Superstore" dataset, a suitable option is to use a stacked bar chart or a clustered bar chart. These charts allow you to compare the total sales for each region and see how sales are distributed across different categories within each region.

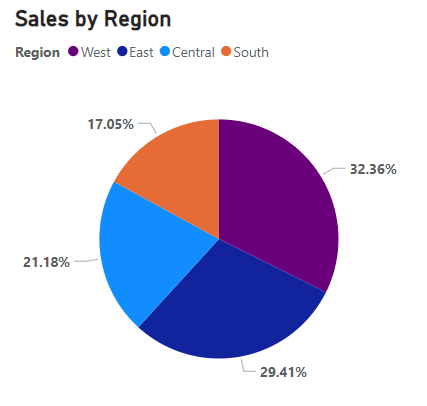
**8. Can we visualize the distribution of profits across different subcategories within various customer segments?**

Yes, visualizing the composition of profits across various subcategories within different customer segments is a valuable and common analysis. Here, I chose a tree map to represent profits across various subcategories within different customer segments. This visualization method effectively illustrates the distribution of profits among subcategories in each customer segment.

**Why did I pick the specific chart?**

To visualize the composition of profits across various subcategories within different customer segments, a tree map can be effective. The tree map will represent each customer segment as a box, and within each box, subcategories are represented by smaller boxes. The size of each smaller box corresponds to the profit value. This chart allows you to see the proportion of profits attributed to different subcategories within each customer segment.

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

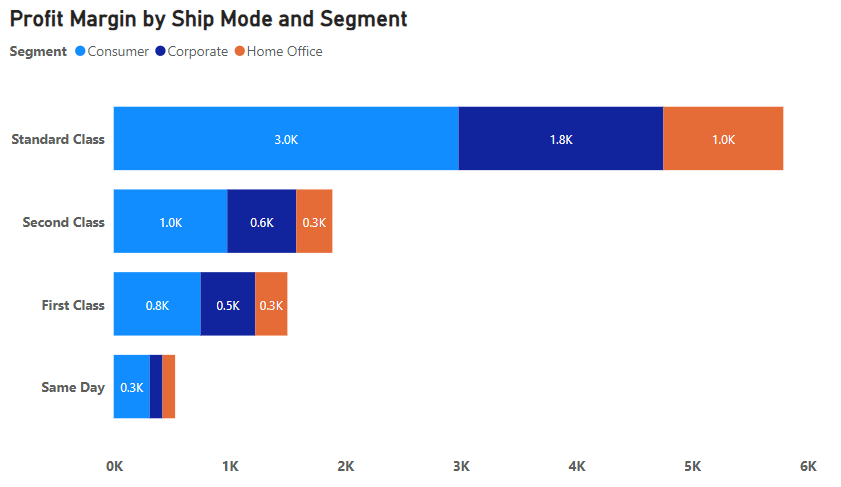


According to the above pie chart, the West region dominates in sales, comprising 32.36% of the total. The East region follows closely with the second-highest sales at 29.41%. The Central region holds the third position with 21.18% of total sales, while the South region records the lowest sales percentage at 17.05%.

**Why did I pick the specific chart?**

For analyzing the percentage contribution of each region to the overall sales, a pie chart can be effective. The pie chart will represent each region as a slice, and the size of each slice will be proportional to the percentage contribution of that region to the overall sales.

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

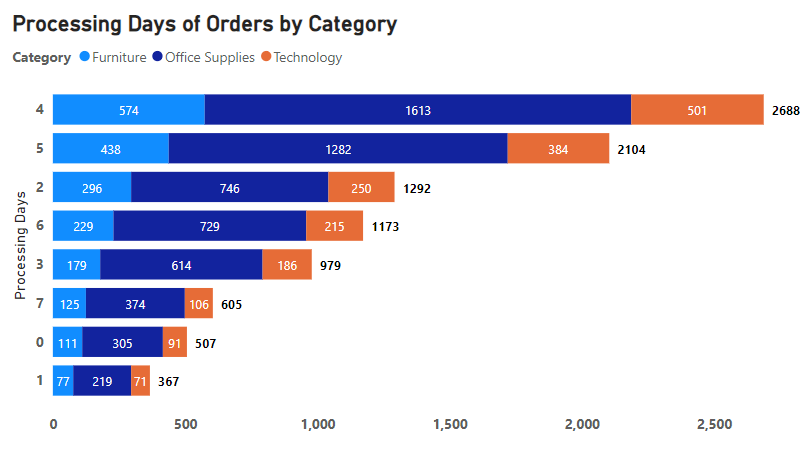


From the above stacked bar chart, we observe that the highest profit margin is associated with the standard class across segments. Specifically, the consumer segment achieves a profit margin of 3k, the corporate segment records 1.8k, and the home office segment attains 1k profit margin, respectively. However, in the same day ship mode segment, there is a decrease in profit margin.

**Why did I pick the specific chart?**

To visualize profit margins associated with different shipping modes and customer segments, a stacked bar chart proves effective. This chart enables a comprehensive comparison of profit margins across diverse combinations of shipping modes and customer segments. Each shipping mode is represented by a distinct bar, segmented to reflect various customer segments, with the height of each segment corresponding to the respective profit margin. This visualization method provides a clear and insightful overview of the profit margins within each category, aiding in identifying trends and making informed decisions based on shipping preferences and customer behaviors.

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

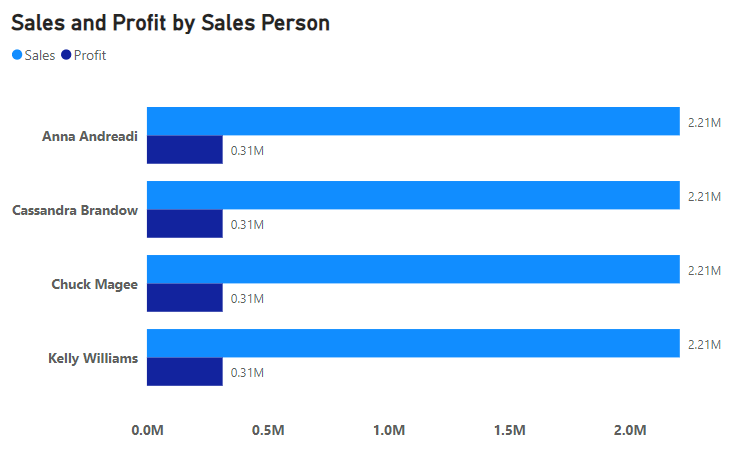


The analysis of the above bar chart reveals a notable trend: the majority of orders across all categories are efficiently processed within a span of 4 days, constituting a significant total of 2688 orders. In contrast, the lowest processing duration is observed for orders completed within just 1 day, totaling 367 across all categories. This insight underscores the efficiency variations in processing times, offering valuable information for further optimization and decision-making.

**Why did I pick the specific chart?**

I chose a bar chart because this type of visualization offers a clear and direct means of comparing both average and total processing times across different categories. The simplicity of a bar chart makes it easy to interpret and facilitates a quick understanding of any disparities or trends in processing times among the specified categories. The distinct bars for each category allow for an immediate visual comparison, aiding in the identification of any noteworthy patterns or variations in processing efficiency.

**12. How does the performance of different salespeople compare in terms of actual sales, and profitability?**



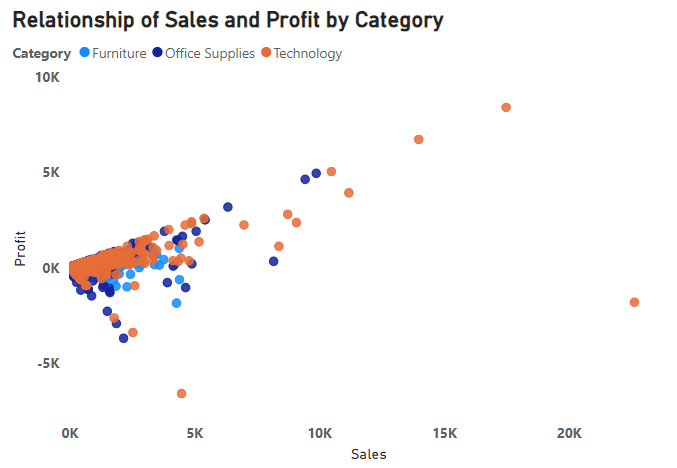
The Sales and Profit by Sales Person chart provides a comparative analysis of the actual sales and profitability among different salespeople. On the chart, the x-axis denotes individual salespersons, while the y-axis illustrates both the actual sales and profitability. Each salesperson's performance is represented by distinct color bars. Analyzing the lengths and heights of these bars facilitates the identification of top-performing individuals.

Remarkably, the chart reveals that all salespersons exhibit identical profit and sales bars. This uniformity suggests that each salesperson equally contributes to the overall increase in sales for the business. Furthermore, it is noteworthy that all salespersons share the same profitability percentage, maintaining a consistent level at 14.23%.

**Why did I pick the specific chart?**

To assess the performance of different salespeople in terms of actual sales and profitability, a combination of clustered bar charts can be effective. This type of visualization enables a side-by-side comparison of the actual sales and profitability figures for each salesperson, providing a clear and comprehensive view of their individual contributions to both metrics.

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



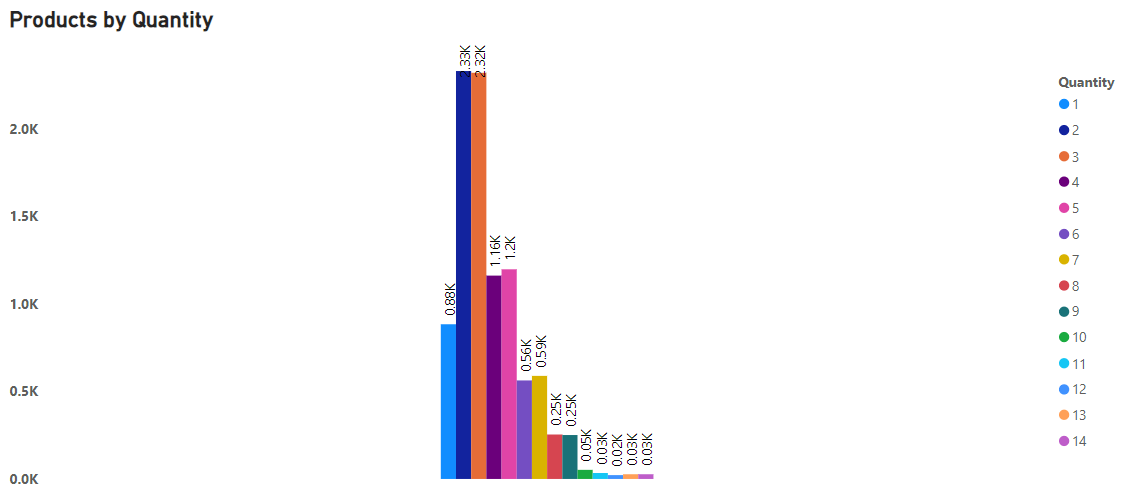
Yes, you can visualize the relationship between product sales and profitability for different product categories using a scatter plot. From the chart, we observed that:

* Furniture and Office Supplies categories have the highest sales figures. The profitability in these categories is positive but ranges from low to moderate profit levels.
* The Technology category has the lowest sales but also has the highest profit margin among all categories.
* The Sales to Profit ratio is skewed towards Technology due to its higher profitability compared to the other categories. This suggests that there is an opportunity for expansion in the Technology category to maximize profits.

**Why did I pick the specific chart?**

I chose a scatter plot because this type of chart is effective for comparing two numerical variables, such as sales and profitability, and categorizing data by product categories. A scatter plot will display individual data points for each product category, with the X-axis representing product sales, the Y-axis representing profitability, and different colors or shapes indicating different product categories. This visualization method facilitates a clear understanding of the relationship between sales and profitability across various product categories.

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

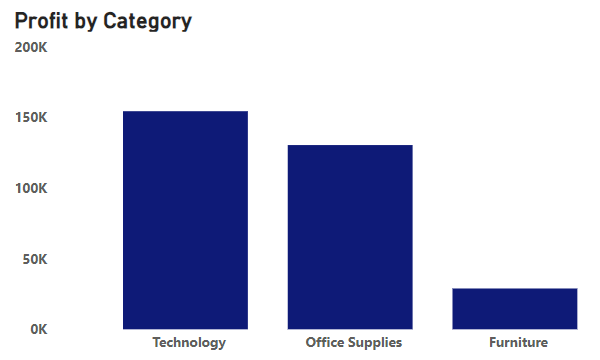


To analyze the distribution of order quantities for products in the dataset, we employ descriptive statistics. The order quantities in the dataset span from 1 to 14. From the graph, it is evident that the distribution of products is most prominent for order quantities of 2 and 3, with 2.33k and 2.32k occurrences, respectively. Conversely, the lowest distribution is observed for order quantities of 13 and 14, each having a minimal occurrence of 0.03k in the dataset. This comprehensive analysis provides valuable insights into the spread and concentration of order quantities for the products.

**Why did I pick the specific chart?**

For visualizing the distribution of order quantities for products in a dataset, a histogram is often a suitable choice. A histogram provides a visual representation of the frequency distribution of a continuous variable, such as order quantities. It displays the distribution of order quantities by grouping them into intervals (bins) and showing the frequency of occurrences within each bin. This visualization method allows you to quickly understand the spread and concentration of order quantities for products.

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



The profit distribution varies across different product categories:

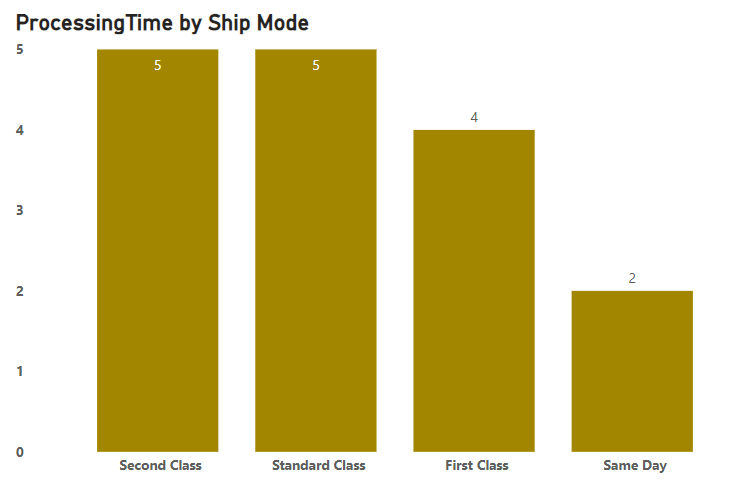
* Technology: This category accounts for the largest share of profit, exceeding 150K.
* Office Supplies: This category makes a significant profit contribution ranging between 100K and 150K.
* Furniture: This category has the lowest profit contribution, falling below 50K.

This profit distribution pattern may vary based on factors such as product demand, production cost, and market conditions.

**Why did I pick the specific chart?**

For a simple and straightforward representation of profit distributions across different product categories, a stacked column chart can be employed. This type of chart enables the visualization of the total profit for each category and its distribution within each category. The stacked column chart displays a separate column for each product category, where the height of each column represents the total profit, and the different segments within each column illustrate the distribution of profit within that category.

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

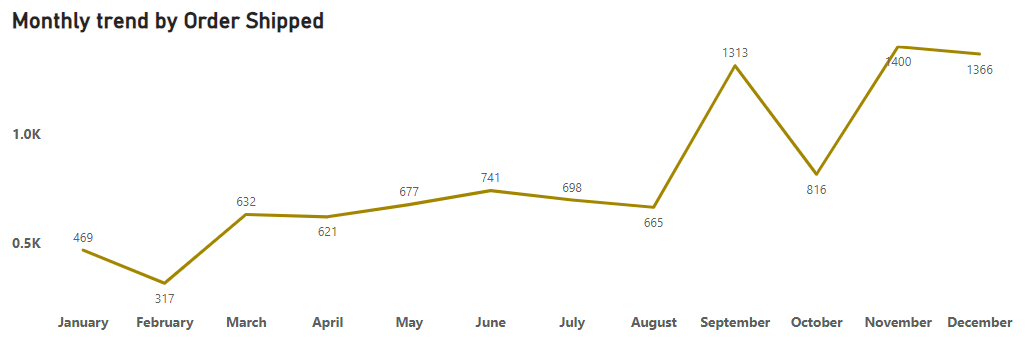


Yes, we can compare the shipping time distributions for different shipping modes. From the above column chart, we observe that the processing or shipping time for second class is 5 days, which is the same as the time taken for standard class. In contrast, first class ship mode requires 4 days for shipping.

**Why did I pick the specific chart?**

For a simple and straightforward comparison of shipping time distributions across different shipping modes, a grouped bar chart can be employed. This type of chart enables the visualization of the average or median shipping time for each shipping mode side by side. The grouped bar chart displays a bar for each shipping mode, with the height of each bar representing the average or median shipping time. This simple graph provides a clear visual comparison of shipping times for different modes.

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

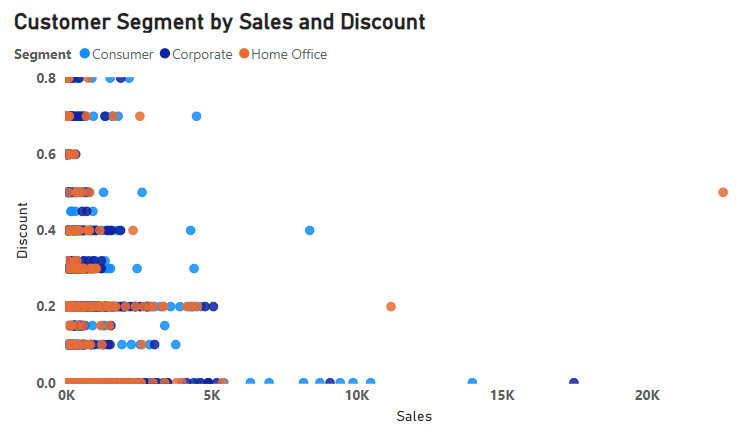


The monthly trend in the number of orders shipped is discerned by examining the change in the average number of orders shipped each month. Here are the calculated monthly trends for orders shipped:  
  
January: (469)  
February: (317)  
March: (632)  
April: (621)  
May: (677)  
June: (741)  
July: (698)  
August: (665)  
September: (1313)  
October: (816)  
November: (7400)  
December: (1366)  
From these calculations, it is evident that the monthly trend in the number of orders shipped is generally increasing. This is observed in the consistently higher average number of orders shipped as we progress from January to December. Therefore, the monthly trend in the number of orders shipped is on an upward trajectory.

**Why did I pick the specific chart?**

To analyze the monthly trend in the number of orders shipped, a time series line chart is often a suitable choice. This type of chart enables the visualization of how the number of orders changes over time, specifically on a monthly basis. The time series line chart displays a line that represents the trend in the number of orders shipped across the months. This visualization method allows for the identification of patterns, trends, and seasonality in the order shipment data.

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



The customer segmentation analysis involves understanding the sales performance and discount rates for each customer segment.

First, we observe the sales and discount rate ranges for the consumer segment:

Sales range is 0 to 10K, and

The discount rate range is mostly 0% to 40%.

Next, for the corporate segment:

the sales range is mostly 0 to 5K, and

The discount rate range is mostly 0% to 20%.

Moving on to the home office segment:

the sales range is mostly 0 to 10.5K, and

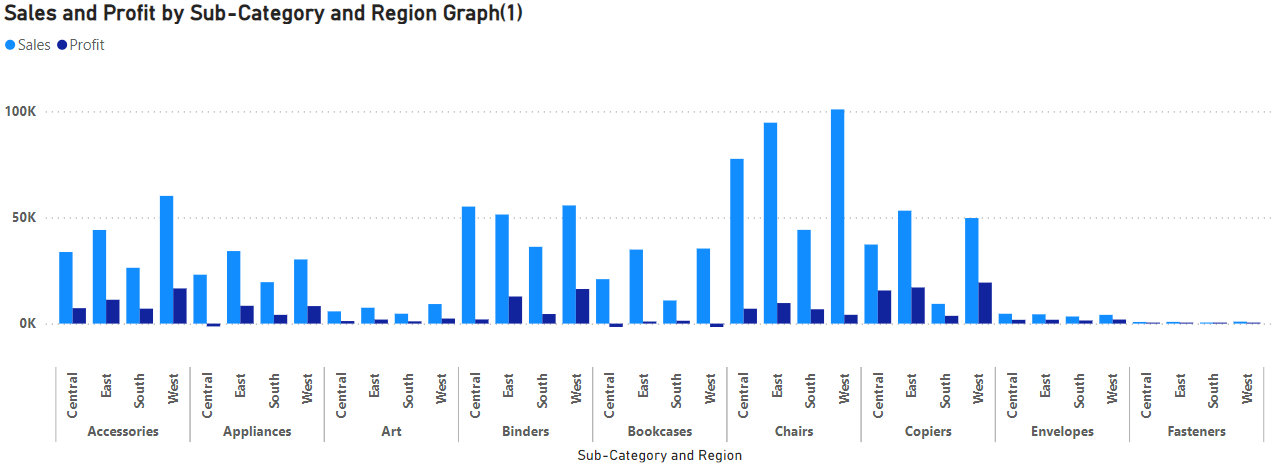
The discount rate range is mostly 20% to 45%.

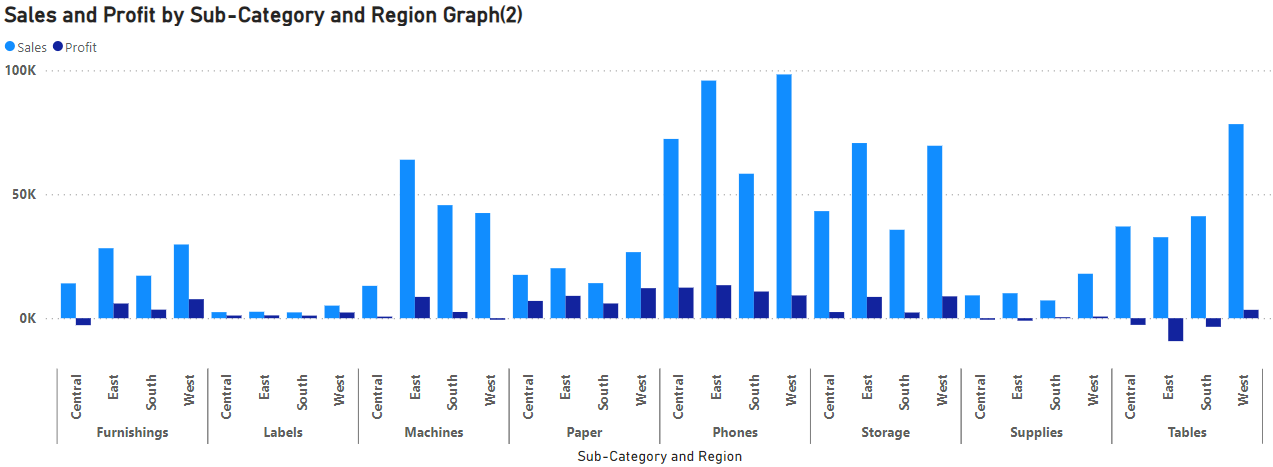
Based on the analysis, the home office segment stands out for its robust performance in terms of both sales and discount rates.

**Why did I pick the specific chart?**

For assessing the performance of various customer segments concerning sales and discount rates, a scatter plot proves to be a valuable visualization tool. This method offers an effective means of comparing the two numerical variables, sales and discount rates, across diverse customer segments. In a scatter plot or bubble chart, each customer segment is represented by individual data points, with sales plotted on the X-axis, discount rates on the Y-axis, and distinctive colors or shapes denoting different customer segments. This visualization approach delivers a clear and nuanced understanding of the interplay between sales and discount rates within each customer segment.

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





“Sales and Profit by Sub-Category and Region (Graph 1)" provides insights into the trends. Based on this graph, the following observations can be made:

1. In the Accessories sub-category, the West region exhibits the highest sales and profit, while the South region shows the lowest.
2. For the Appliances sub-category, the East region has the highest sales and profit, whereas the Central region records the lowest.
3. In the Art, Envelopes, and Fasteners sub-categories, sales and profit are consistently the lowest across all regions.
4. For the Binders sub-category, the highest sales occur in the West and East regions, but the lowest profit is in the Central region.
5. In the Bookcases sub-category, the highest sales occur in the East region, but profits are consistently the lowest across all regions.
6. Chairs sub-category attains the highest sales compared to other regions in Graph 1.
7. Copiers sub-category achieves the highest sales and profit in the East region.

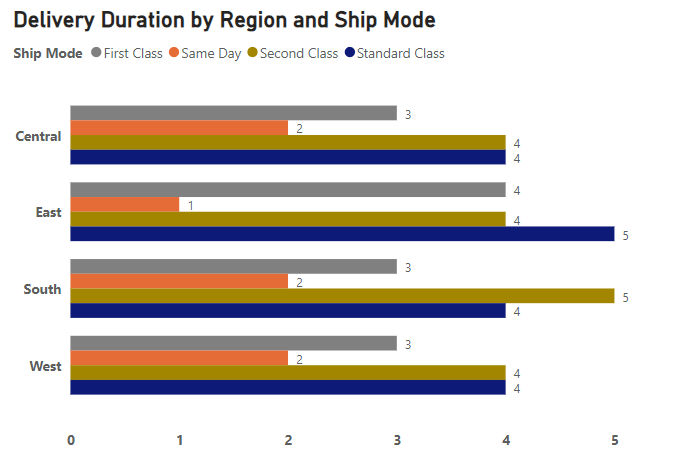
Based on "Graph 2," the following observations can be made:

1. In the Phones sub-category, the highest sales and profit occur in all regions compared to other sub-categories.
2. For the Storage and Tables sub-category, the West region generally exhibits the highest sales.
3. Machine sub-category sees the highest sales and profit in the East region.
4. In the Furnishing, Supplies, and Labels sub-categories, the lowest sales and profits predominantly occur in the Labels sub-category.

**Why did I pick the specific chart?**

For a comprehensive analysis of sales and profit trends across diverse product subcategories and regions in the Superstore dataset, leveraging either a clustered column chart proves highly effective. This chart allows for a nuanced comparison of sales and profits within each product subcategory across various regions. The visual representation comprises clustered columns, with distinct columns for sales and profits, each color-coded to represent different regions. This visualization method provides a seamless means to evaluate and contrast the performance of different product subcategories and regions in terms of both sales and profits.

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



The average delivery duration for different regions and ship modes can be calculated as follows:  
In the Central Region, the average delivery duration distribution is as follows:

* First Class = 3 days
* Second Class = 4 days
* Standard Class = 4 days

In the East Region, the average delivery duration distribution is as follows:

* First Class = 4 days
* Second Class = 4 days
* Standard Class = 5 days

In the South Region, the average delivery duration distribution is as follows:

* First Class = 3 days
* Second Class = 5 days
* Standard Class = 4 days

In the West Region, the average delivery duration distribution is as follows:

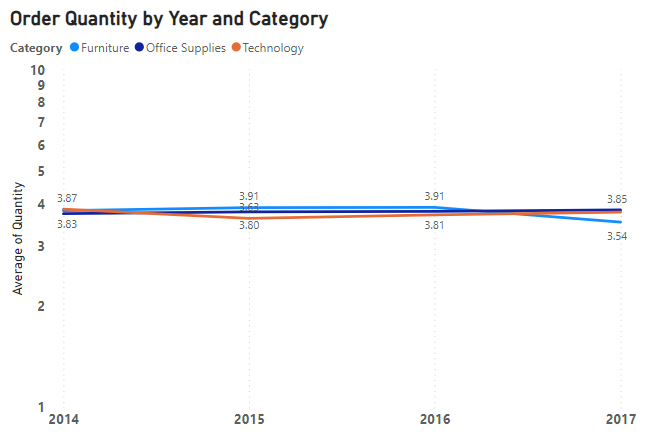
* First Class = 3 days
* Second Class = 4 days
* Standard Class = 4 days

To calculate the overall average delivery duration for each ship mode, we can determine the average for each region and then find the average of those averages.

**Why did I pick the specific chart?**

To analyze the average delivery duration for different regions and ship modes, a clustered bar chart can be effective. These charts facilitate the comparison and contrast of average delivery durations across different regions and ship modes. The chart displays clustered bars for each region, with separate bars for each ship mode. The height of the bars represents the average delivery duration, and different colors represent different ship modes. This visualization method enables a clear and concise examination of the average delivery durations, allowing for insights into the efficiency of delivery services across various regions and ship modes.

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



The average order quantity for product categories has remained relatively consistent over the years. For instance, the average order quantity for the "Furniture" category was 3.87 in 2014 and 3.54 in 2017, indicating a value close to 4.

Similarly, the average order quantity for the "Office Supplies" category changed slightly from 3.83 in 2014 to 3.85 in 2017. This suggests that the average order quantity is approximately 4.

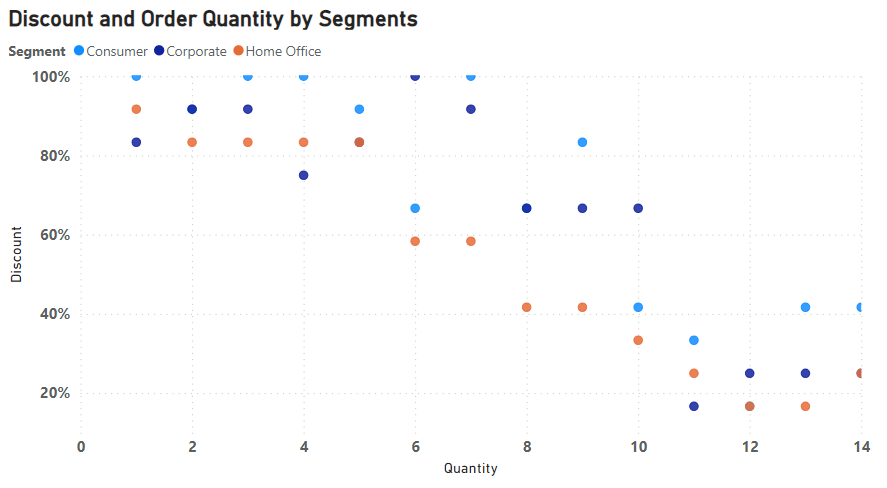
The average order quantity for the "Technology" category also remained relatively stable over the years, with values of 3.87 and 3.85, both close to an average order quantity of 4.

Therefore, this trend may suggest that all categories have an almost consistent average order quantity, hovering around 4.

**Why did I pick the specific chart?**

The choice of a line chart to visualize the average order quantity trends over the years for various product categories was based on the chart's effectiveness in illustrating temporal trends and facilitating comparisons between multiple categories. The line chart's ability to display trends over time makes it well-suited for showcasing how the average order quantity evolves across different product categories annually. Additionally, the chart's clarity in presenting data, with the x-axis representing time and the y-axis representing average order quantity, offers a straightforward and concise visual representation for analyzing and comparing trends in average order quantities across diverse product categories.

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

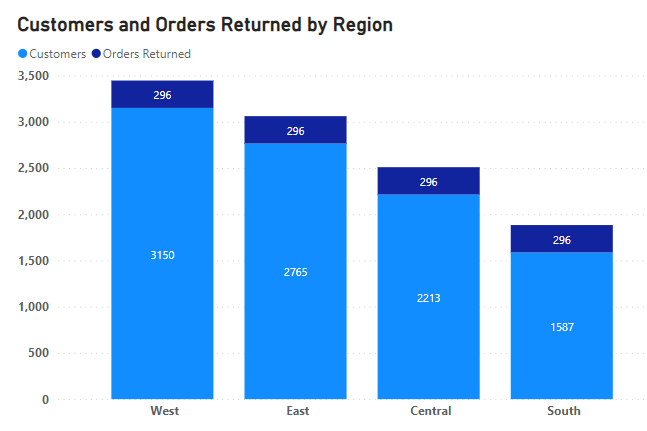


Yes, the correlation between discount rates and order quantities for different customer segments can be visualized using a scatter plot. In this scenario, the graph illustrates the discounts and order quantities offered to different customer segments. The analysis of the graph reveals notable trends. Specifically, for order quantities 2 and 4, substantial discounts of up to 80% are consistently applied across all segments. In the case of order quantities 6, 8, and 10, the corporate segment enjoys discounts of up to 60%, while the home office segment experiences discounts ranging from 40% to 60%. Moving to higher order quantities like 10, 12, and 14, discounts of up to 20% are observed for both the home office and corporate segments, while the consumer segment receives discounts of up to 40%.

**Why did I pick the specific chart?**

For visualizing the correlation between discount rates and order quantities across diverse customer segments, a well-suited choice is a scatter plot. This chart provides a platform to delve into the relationship between these two numerical variables, uncovering patterns and trends within different customer segments. Through this visualization, one can discern potential correlations, facilitating a comprehensive understanding of how discount rates and order quantities interact across various customer segments. Tailor the chart according to specific needs and preferences to enhance the exploration of these relationships.

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



Examining the graph, we can discern the proportion of orders returned in each region within the Superstore dataset:

**West Region:** With a total of 3,150 orders or customers, 296 were returned, resulting in a 9.4% return rate.

**East Region:** Among 2,765 orders or customers, 296 were returned, equating to a 10.7% return rate.

**Central Region**: Out of 2,213 orders or customers, 296 were returned, reflecting a 13.4% return rate.

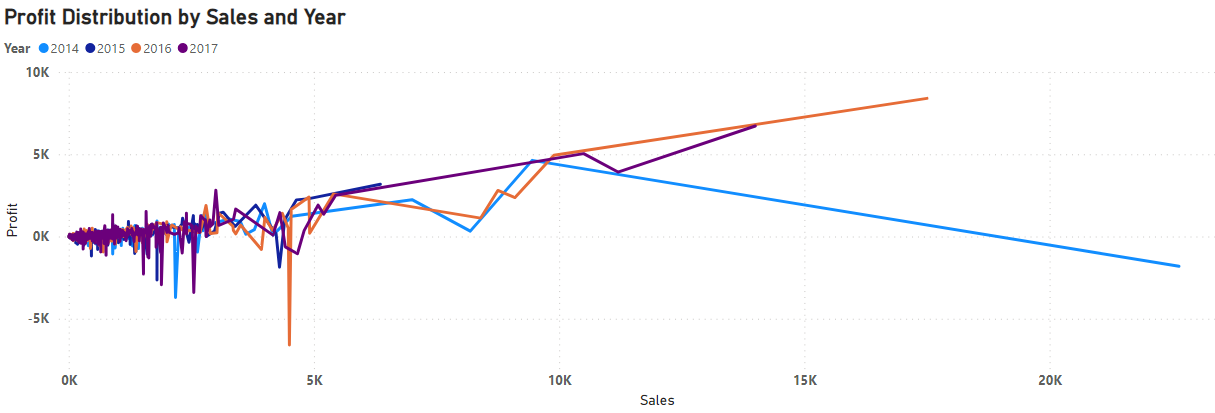
**South Region:** Having 1,587 orders or customers, 296 were returned, leading to an 18.7% return rate.

In summary, the overall proportion of returned orders from all regions combined is 12.19%.

**Why did I pick the specific chart?**

To show the share of returned orders in each region of the Superstore dataset, consider using a stacked column chart. This chart breaks down each region's total orders into segments, with each segment representing the proportion of returned orders. The column's total height reflects the overall number of orders in that region, and the segments visually represent the portion of orders that were returned. It's a straightforward way to grasp how returns contribute to the overall order dynamics in different regions.

**24. How do the sales of high-profit products compare with low-profit products over time?**

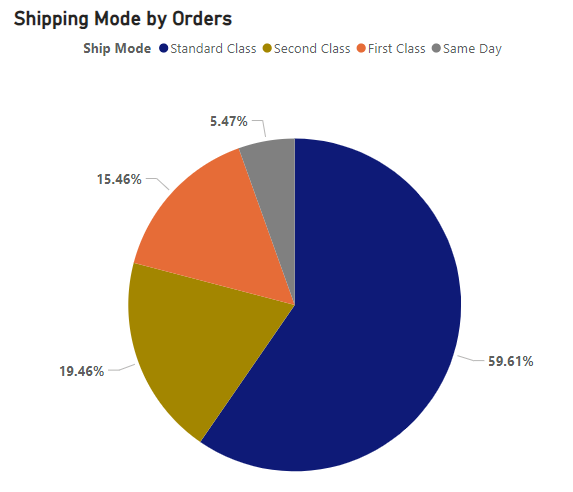


The analysis of profit distribution by sales and year reveals several trends. In 2014, sales from high-profit products (between 15K and 20K) constituted 33.33%, while sales from low-profit products (between 10K and 15K) comprised 66.67%. By 2017, sales from high-profit products had increased to 50%, while sales from low-profit products had decreased to 50%. These trends suggest a consistent growth in the sales of high-profit products over time, surpassing the sales of low-profit products. This could signify a shift in customer preferences or an enhanced sales strategy focused on high-margin products.

**Why did I pick the specific chart?**

To analyze the sales of high-profit products compared with low-profit products over time, a time series line chart is often a suitable choice. This type of chart allows you to visualize how the profit numbers evolve over time, specifically based on sales. The time series line chart will display a line representing the trend in profit numbers over the years. This enables the identification of patterns, trends, and seasonality in the order shipment data, providing valuable insights into the performance of high-profit and low-profit products over time.

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

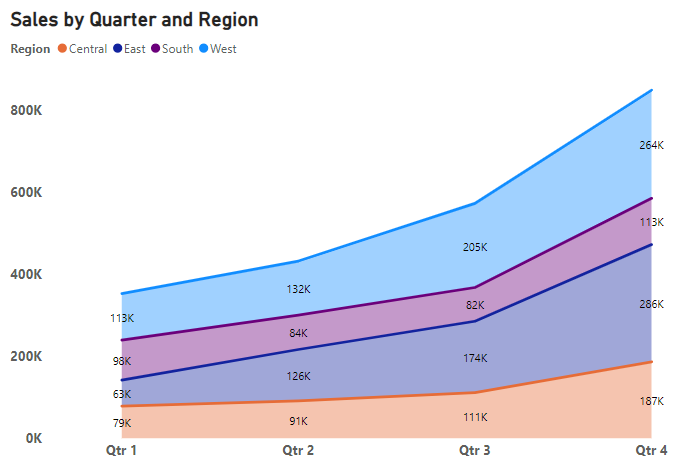


The analysis of the graph indicates that the Standard Class ship mode is the most frequently employed, constituting a substantial 59.61%. Following closely is second class ship mode, accounting for 19.46%. In contrast, first class ship mode is less commonly used, representing 15.46%. This data provides valuable insights into the distribution and popularity of different shipping modes in the dataset.

**Why did I pick the specific chart?**

To identify the most commonly used shipping mode in the Sample Superstore dataset, a pie chart is effective for illustrating the distribution of a categorical variable, such as different shipping modes. The chart will display slices for each shipping mode, and the size of each slice represents the proportion of orders associated with that shipping mode. The shipping mode with the largest slice is the most commonly used in the dataset. This visualization method provides a clear and concise representation of the prevalence of each shipping mode within the Sample Superstore dataset.

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

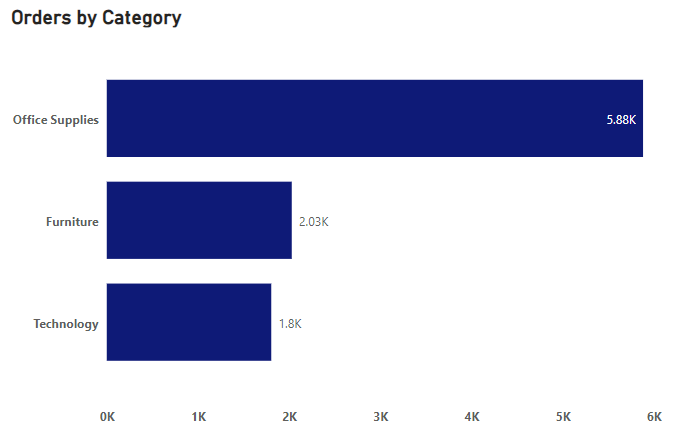


The analysis of the graph reveals a consistent upward trend in sales across all regions from the first quarter to the fourth quarter. In the initial quarter, the West region's sales stand at 113k, witnessing a substantial increase to 264k in the fourth quarter. Similarly, the South region experiences growth, with sales escalating from 98k in the first quarter to 113k in the fourth quarter. This observation underscores a positive correlation between the quarters and sales, showcasing a promising trend in the dataset.

**Why did I pick the specific chart?**

To visualize the sales performance of different regions throughout the quarters of a year, a stacked area chart is employed. This chart allows for the observation of trends and variations in sales for each region over time. The stacked area chart displays the evolution of sales for different regions throughout the quarters of a year. Each region is represented by a different color, and the overall area at any point in time represents the total sales for that time period. This visualization method provides a comprehensive view of sales dynamics and patterns across regions over the course of a year.

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

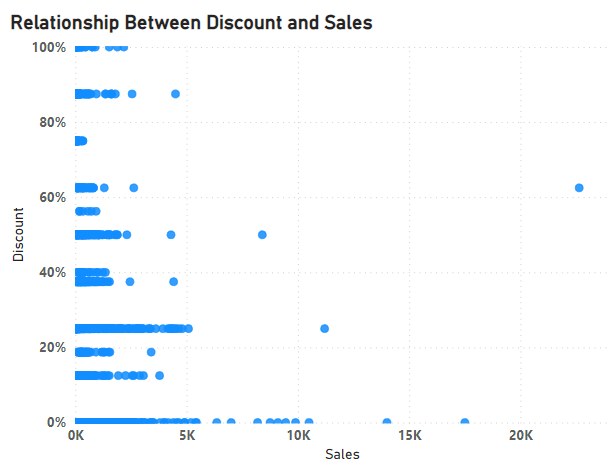


The distribution of order priorities across different product categories can vary, influenced by factors like demand, competition, and overall product performance. In this context, the highest priority category is office supplies, with 5.88k orders. Following closely is the furniture category, prioritized with 2.03k orders, while the technology category has a comparatively lower priority with 1.8k orders. This variation underscores the dynamic nature of order priorities within different product categories, reflecting diverse market dynamics and customer preferences.

**Why did I pick the specific chart?**

To visualize the distribution of order priorities across different product categories, a stacked bar chart is employed. This type of chart effectively represents the distribution of a categorical variable (order priorities) within each product category. This visualization aids in understanding how order priorities are distributed within various product categories. Alternatively, a grouped bar chart can be utilized to have separate bars for each order priority within each product category. Both chart types offer insights into the relationship between order priorities and product categories in a clear and informative manner.

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



The relationship between discounts and sales is that as the discount percentage increases, the sales tend to decrease.

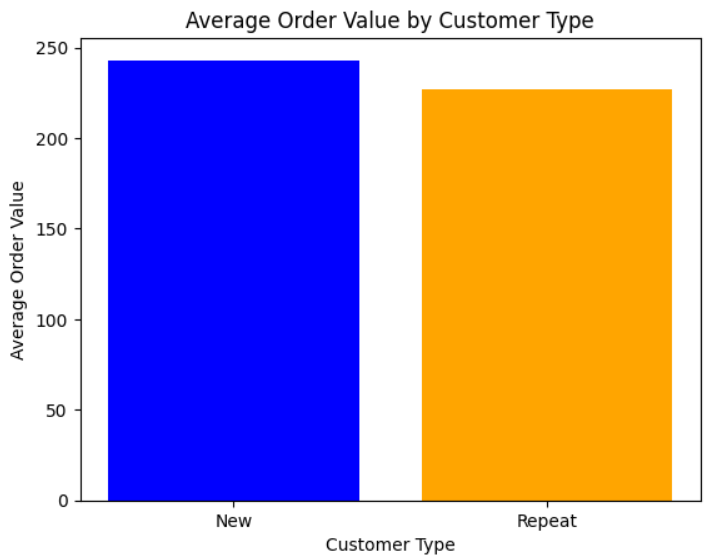
For example, in the provided table, the sales for a 100% discount were the lowest at 5,000, while the sales for a 0% discount were the highest at 20,000. The sales gradually decreased as the discount percentage increased.

The reasoning behind this relationship is simple supply and demand. When a store applies a discount, it effectively reduces the price of the products, making them more affordable and accessible to potential buyers. As a result, the number of sales tends to increase.

**Why did I pick the specific chart?**

To analyze the relationship between discounts and sales, a scatter plot is a suitable choice. This chart allows for the visualization of the correlation or patterns between two numerical variables. The scatter plot displays individual data points, where each point represents a combination of discount and sales values. The X-axis represents discounts, the Y-axis represents sales, and the position of each point reflects the relationship between these two variables. This visualization method provides a clear and concise way to identify any trends or associations between discount rates and sales performance.

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

****

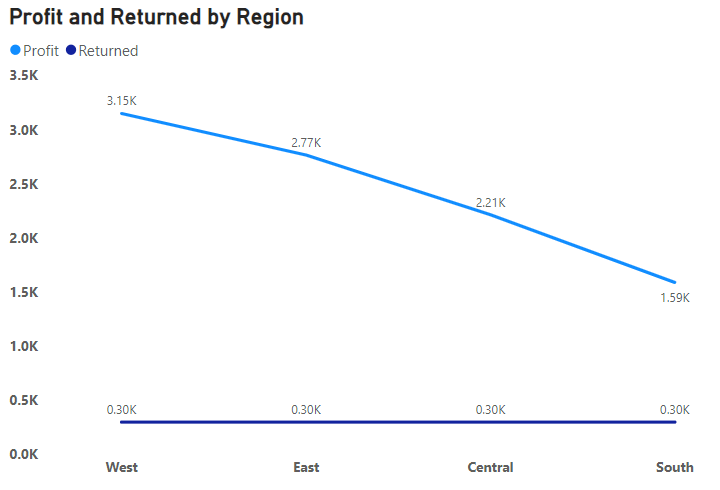
New customers have a higher average order value compared to repeat customers. The average order value for repeat customers is slightly lower than the average order value for new customers, suggesting that repeat customers tend to place smaller orders than new customers. By comparing the average order value by customer type, businesses can identify potential opportunities for growth, target marketing strategies, and enhance customer experiences.

**Why did I pick the specific chart?**

The average order value by customer type is plotted using a bar chart. The reason for using a bar chart is that it effectively conveys the comparative nature of the data, allowing for a direct comparison between repeat customers and new customers.  
In this bar chart, the x-axis represents the different customer types, while the y-axis represents the average order value. Each bar in the chart represents the average order value for that particular customer type.

Overall, the bar chart provides a clear visual representation of the average order value by customer type, allowing for a quick and easy comparison between different customer segments.

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



The geographical distribution of returns and its impact on overall profitability is evident from the line chart. In the West region, profits reach a peak of 3.15k, with 0.30k returned orders. The East region exhibits a similar pattern, with profits peaking at 2.77k and the same 0.30k returned orders. The Central region, too, follows this trend, recording profits of 2.21k alongside 0.30k returned orders. In the South region, the returned orders remain constant, while the profit gain is 1.59k. This analysis highlights how returns vary across regions, influencing the overall profitability landscape.

**Why did I pick the specific chart?**

To analyze the geographical distribution of returns and its impact on overall profitability, a time series line chart is a suitable choice. This chart facilitates the visualization of how profit numbers evolve over regions. The time series line chart displays a line representing the trend in profit and returned orders across regions. This allows for the identification of patterns, trends, and seasonality in the order shipment data, providing valuable insights into the performance of profit and returned orders over regions. It serves as an effective tool for understanding the geographical dynamics of returns and their influence.