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: Sample - Superstore Sales (Excel).xls

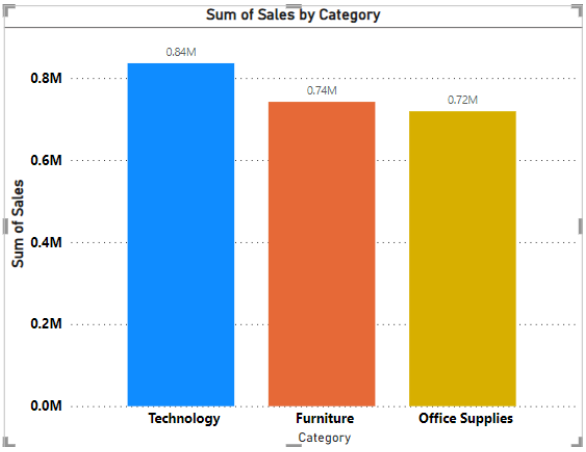
Please keep in mind:

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.

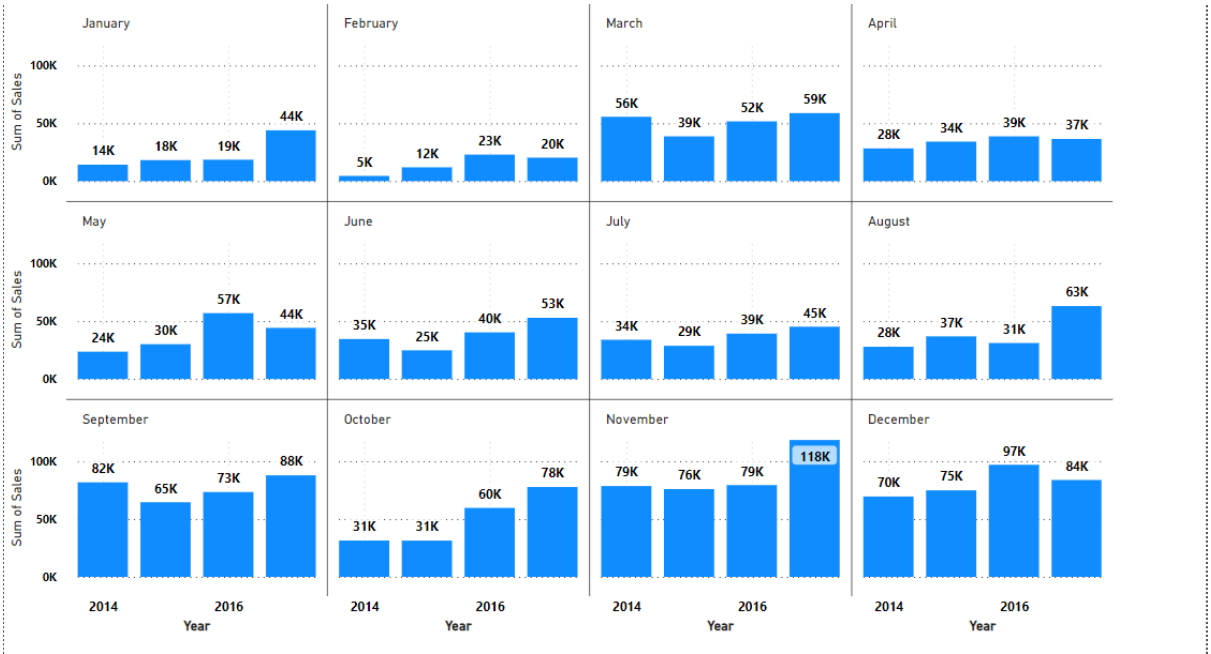
Good luck!

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



To make things simple to grasp and help clients find the best options, I've decided to use a bar plot. With total sales of 839,893, the technology category is the best-selling, while office supplies are the lowest-selling, with total sales of 731,893.

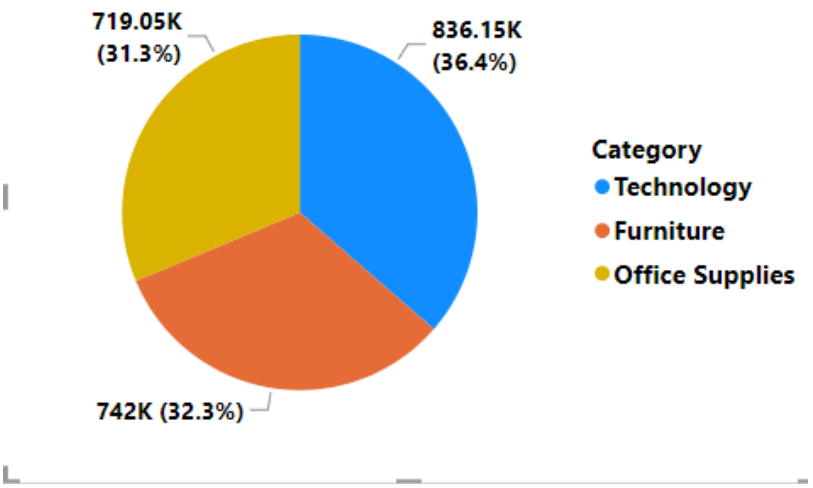
How do the monthly sales amounts change over the course of a year?



I have opted for a line plot since it provides an easy-to-understand visual representation of the pattern over time. 2014 saw the lowest sales in February (5k), and the highest sales in September (82k). In 2015, sales peaked in November at 76k, with February having the lowest sales at 12k. Sales in January of 2016 were 19k, and December saw the highest sales of 97k.

Ultimately, February recorded sales of 20K in 2017, with November having the greatest sales at 118K.

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

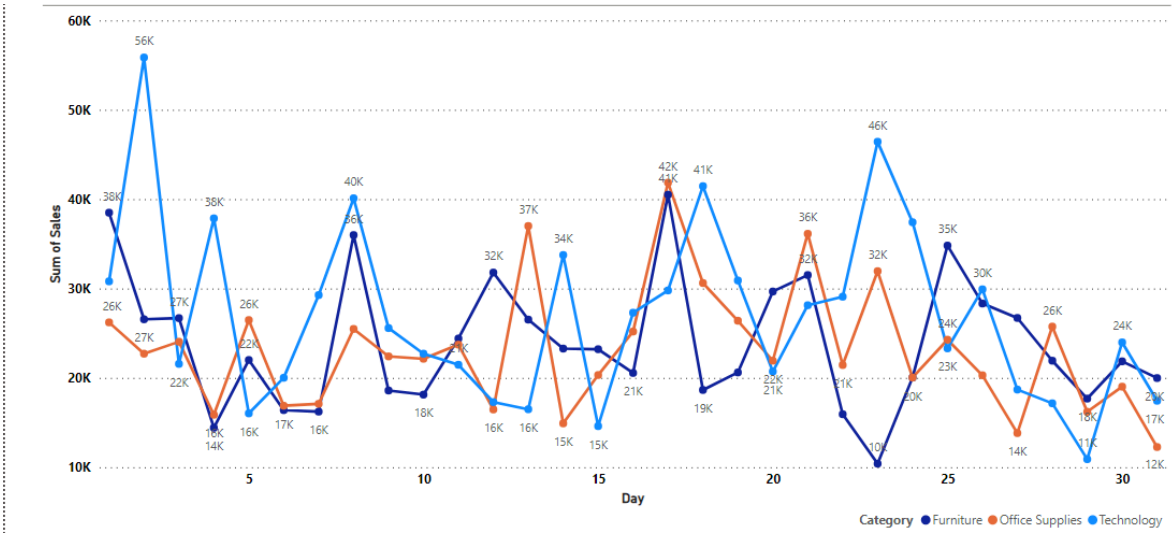


To show how the overall sales amount was distributed across the various product categories, I decided to use a pie plot. It is clear from this plot that the "Technology" category has the largest sales, while the "Office Supplies" category has the lowest. Stakeholders may rapidly understand how sales are allocated among product categories and how much each category contributes to the overall total by using this visualisation, which facilitates decision-making and resource allocation for the firm.

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

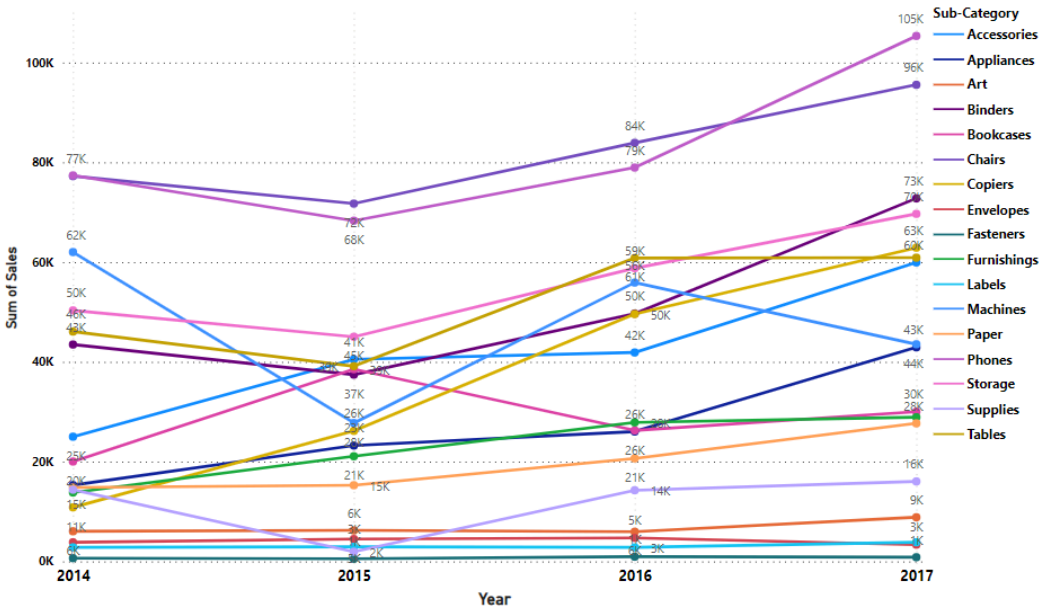
To examine employee performance within each department across multiple areas, I decided to use a tree map. The tree map shows that, with 6k units supplied, the Consumer department in the West area made the most contribution, while the Home Supplies department in the South provided the least, with just 1k units. Making strategic decisions is aided by this visualization, which shows performance differences between departments and geographical areas.

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



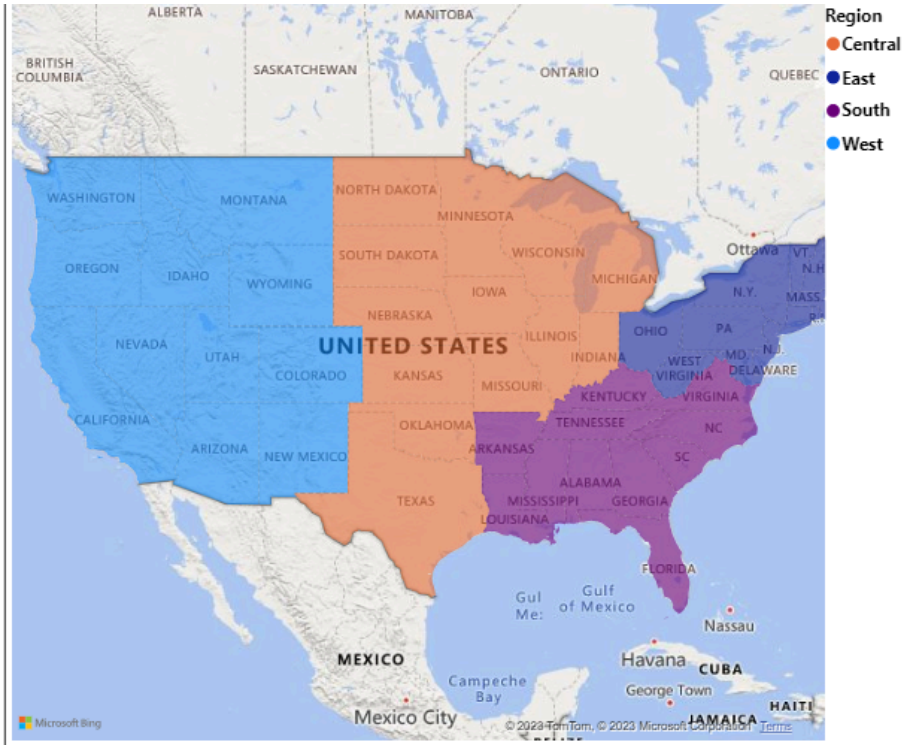
An examination of sales data with a line chart reveals intriguing trends. Wednesdays had the lowest sales in the furniture category, coming in at $14,000, followed by office supply at $16,000 and technology at $22,000. In contrast, the highest sales of furniture occur on Mondays at 36k, followed by those of technology on Saturdays at 34k and office supplies on Saturdays at 32k. Identifying weekly sales trends by product category with the use of this visualization facilitates strategic decision making.

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



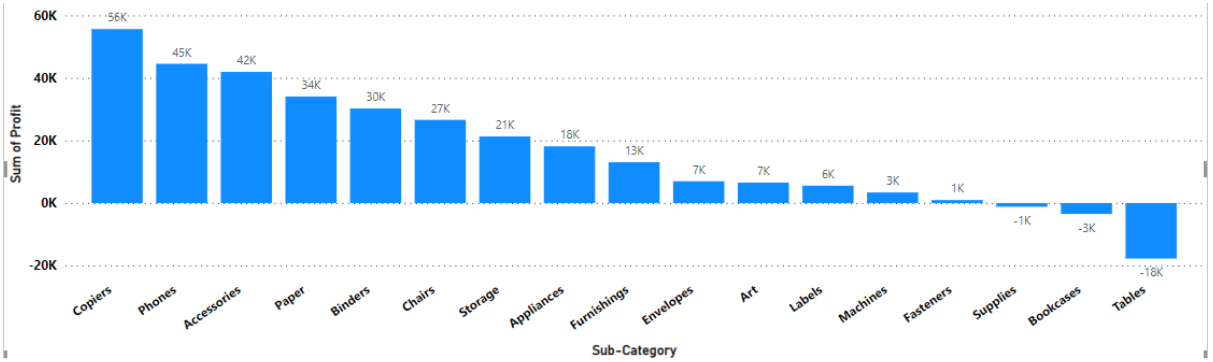
When displaying sales increase over time for different product categories, a line chart works well. I've divided the data into smaller categories in this plot. Notably, the fastest-growing items are Fasteners, whereas the fastest-growing items are Phone Chair and Other Stuff. We can watch the sales trajectory of various product categories with the help of this graphic depiction, which helps us spot trends and make wise business decisions.

How does the sales distribution vary across different regions in the “Superstore” dataset?



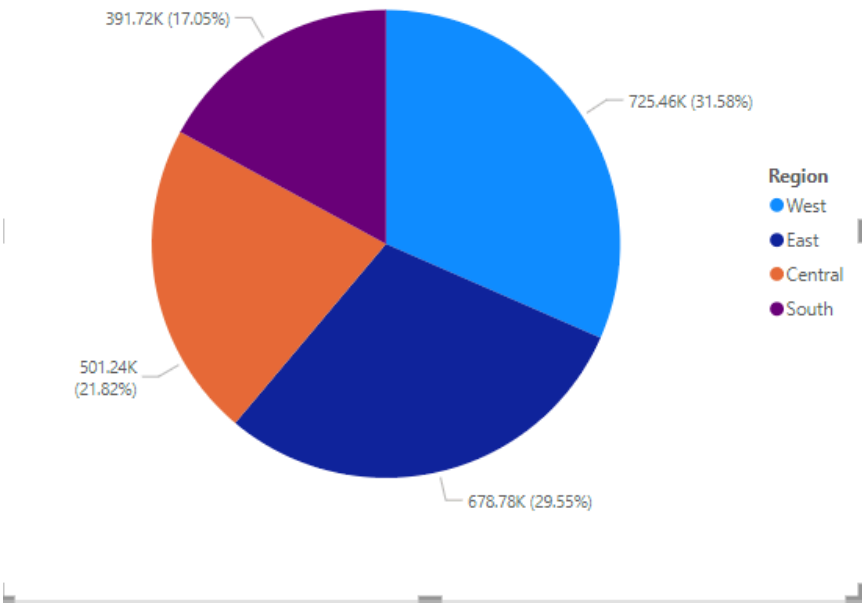
A map plot can be used to display the sales distribution among the various regions in the “Superstore” dataset. With sales of $72,000, the West area leads the representation and indicates notable increase. Comparatively speaking, the South region had the lowest sales (39k), indicating relatively little development in Superstore product sales in that region.

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



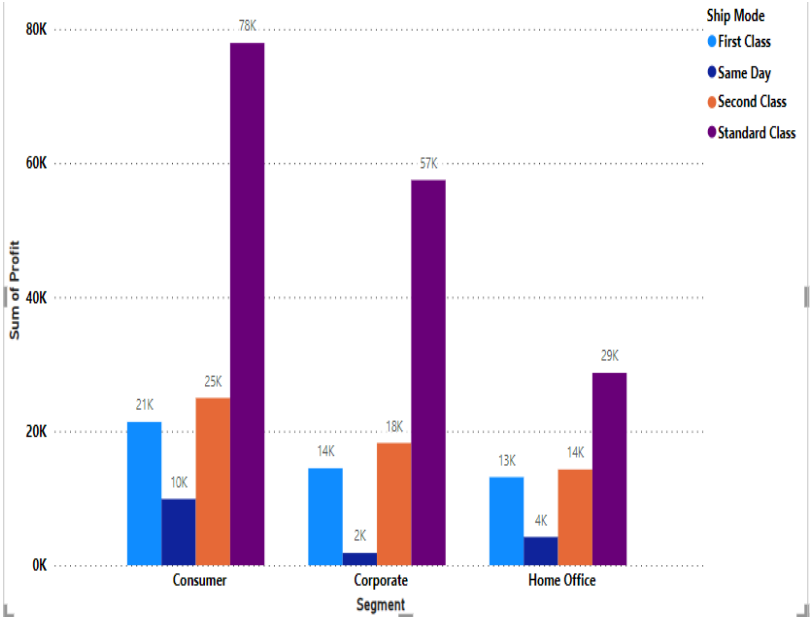
With a bar plot, I can see how profits are distributed among different subcategories inside distinct consumer groupings. It's simple to see which product subcategories in this client segmentation provide the highest profit in this plot. For example, in this superstore data, "Tables" have a negative profit of -18k, suggesting a loss in this subcategory, while "Copiers" earn the maximum profit of 56k.

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



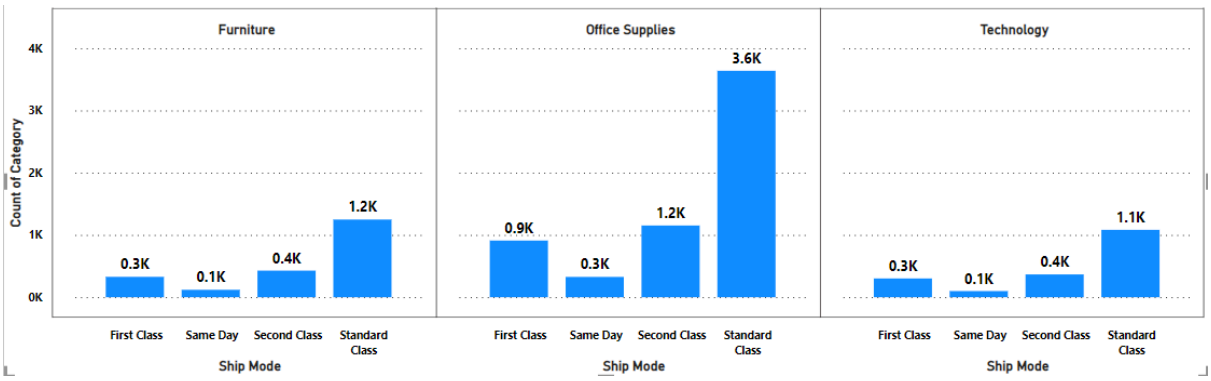
A pie plot is chosen to show the proportion contribution of each region to total sales. With 31.58% of total revenues, the West area clearly contributes the most to this plot, while the South region contributes the least, at 17.05%. The sales distribution among the various regions is efficiently conveyed by this visual representation.

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



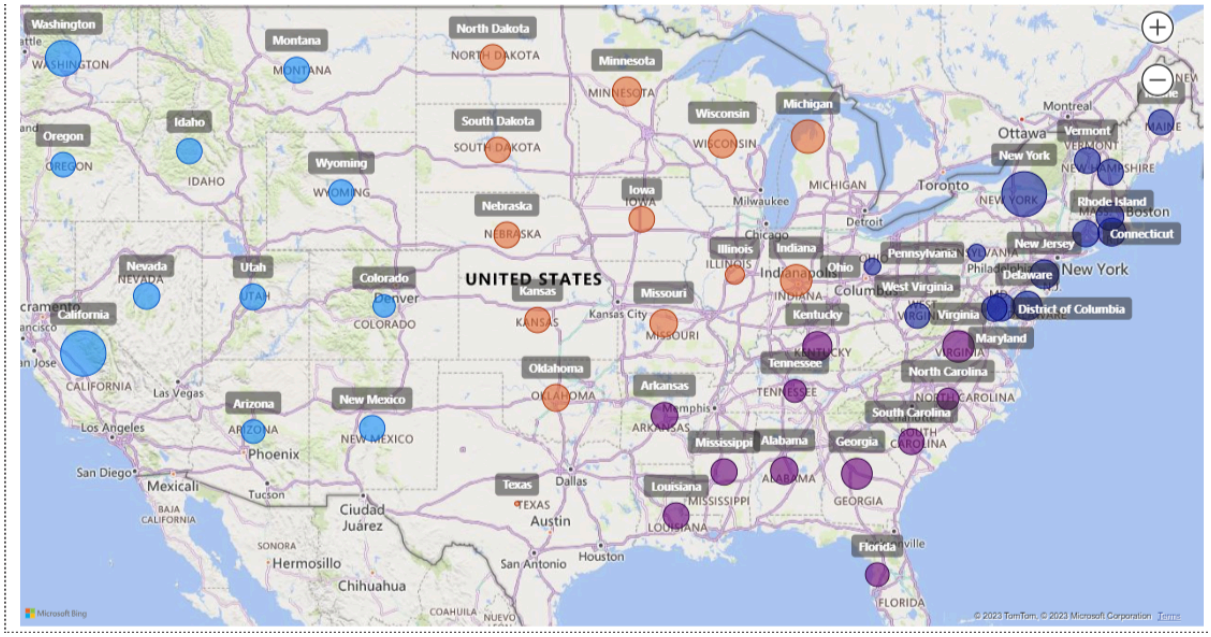
A column chart allows me to see the profit margins related to various shipping options and customer categories. The "Consumer" category, especially when using the "Standard Days" shipping option, has the largest profit margins, according to this data. On the other hand, "Same Day" shipment has the lowest margins of all the client categories. "Corporate" clients had comparable preferences for "Home Appliance" products in the "Corporate" market, with "Standard Class" being the most requested shipping method and "Same Day" being the least liked.

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



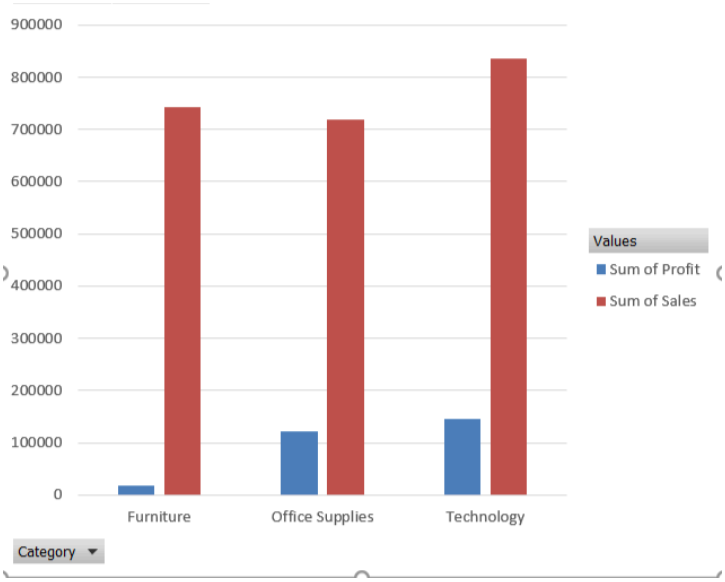
To see how long it takes to process orders for various product categories, I can create a bar plot. According to the plot, "Furniture" products process the slowest, with "Standard Class" taking the longest and "Same Day" the quickest. Similar processing durations are seen in the "Office Supplies" and "Technology" categories, which is consistent with the trend seen in the "Furniture" category.

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



Based on a symbol map of the sales data, the California sales representative is leading the pack in terms of both profitability and sales. Ohio's sales representative is the least productive, with the lowest profitability and sales. This implies that the performance of several salespeople varies significantly from one another.

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



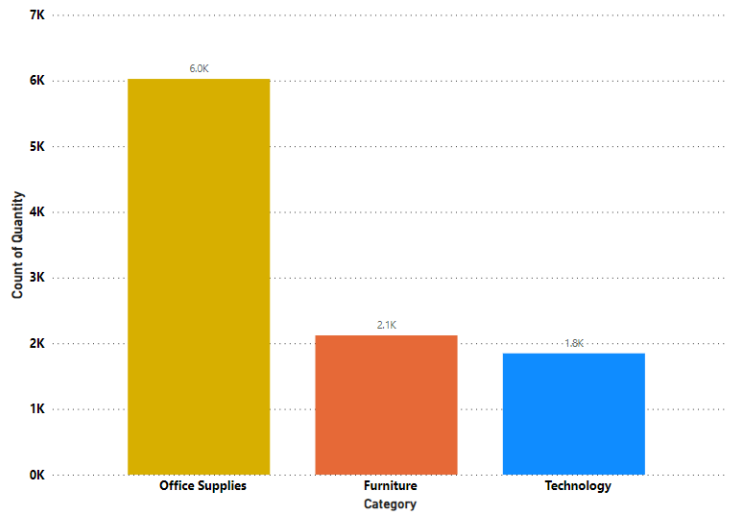
Technology reigns supreme, dominating both sales and profitability. Its sizable bubble sits at the top right corner, indicating high sales and profit margins. Furniture follows closely behind in terms of sales, occupying a large bubble in the middle-right. However, its profitability is lower compared to technology, as its bubble sits slightly to the left. Office Supplies brings up the rear in terms of both sales and profitability. Its smaller bubble sits in the bottom left corner, indicating lower sales and profit margins.

Individual Product Categories:
Technology: Phones and laptops are the clear frontrunners, contributing significantly to both sales and profitability. Their sizable circles within the technology bubble highlight their dominance. Cameras and Televisions also perform well, but to a lesser extent than phones and laptops.

Furniture: Tables and chairs are the top sellers, evident from their large circles within the furniture bubble. However, their profitability is lower compared to technology, as their bubble sits slightly to the left. Desks and beds also contribute decently to sales, but their profitability is even lower.

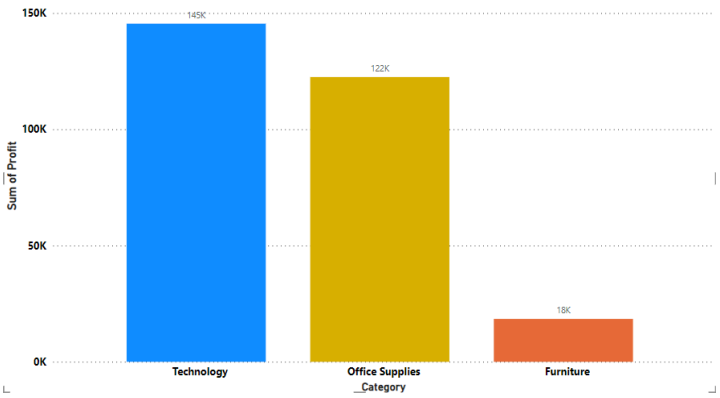
Office Supplies: Paper and binders are the best performers, occupying the largest circles within the office supplies bubble. Envelopes and toner cartridges also contribute some sales, but their impact is smaller.

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



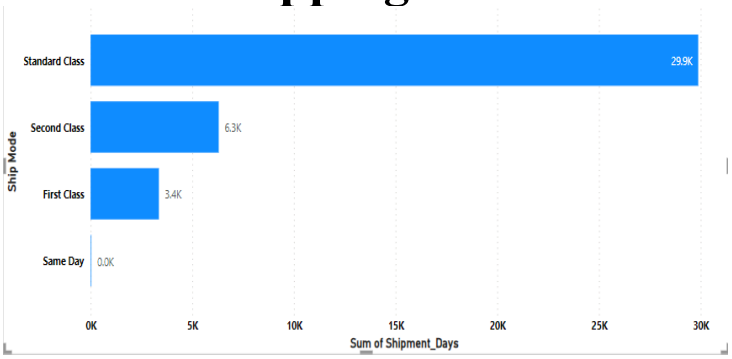
Without a doubt, a column chart works well for showing how long it takes to process orders for various product categories. The distribution of order numbers among categories is shown in this chart. An example of how order processing times vary by product category is given by the fact that "Office Supplies" has the greatest order count (6k) and "Technology" products has the lowest order count (1.8k).

How do the profit distributions vary across different product categories?



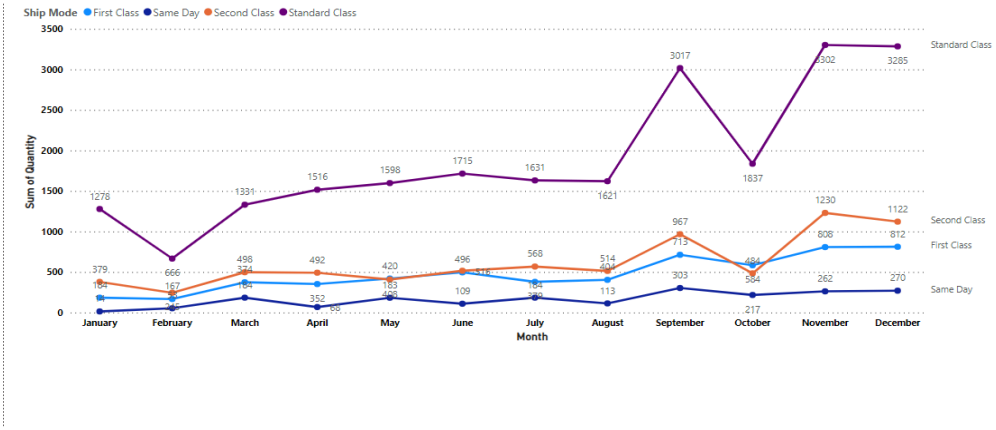
When attempting to visualize differences in profit distributions among several product groups, a bar chart works well. It is clear from this plot that "Technology" products contribute the most to profit (145k), followed by "Office Supplies" (122k). On the other hand, "Furniture" contributes the least to profit, with a profit distribution of just 18k, illustrating the differences in profitability between product categories.

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



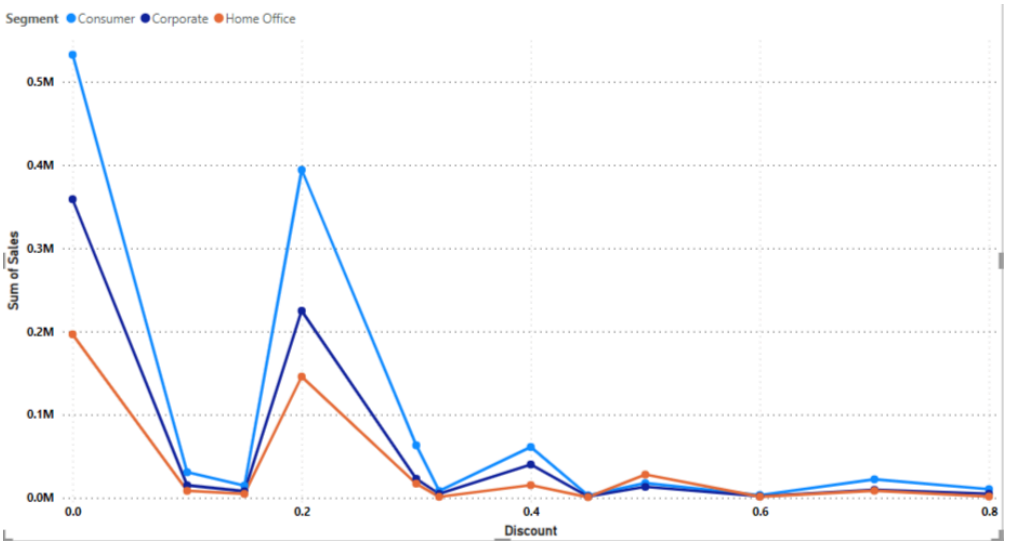
A useful tool for comparing shipment time distributions across various shipping modalities is a column chart. The results indicate that "Same Day" shipping is the least favoured option, while "Standard Class" shipping is the most popular. Furthermore, it's interesting to note that offline store orders are primarily related with "Same Day" shipment, implying longer processing times compared to other modes

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



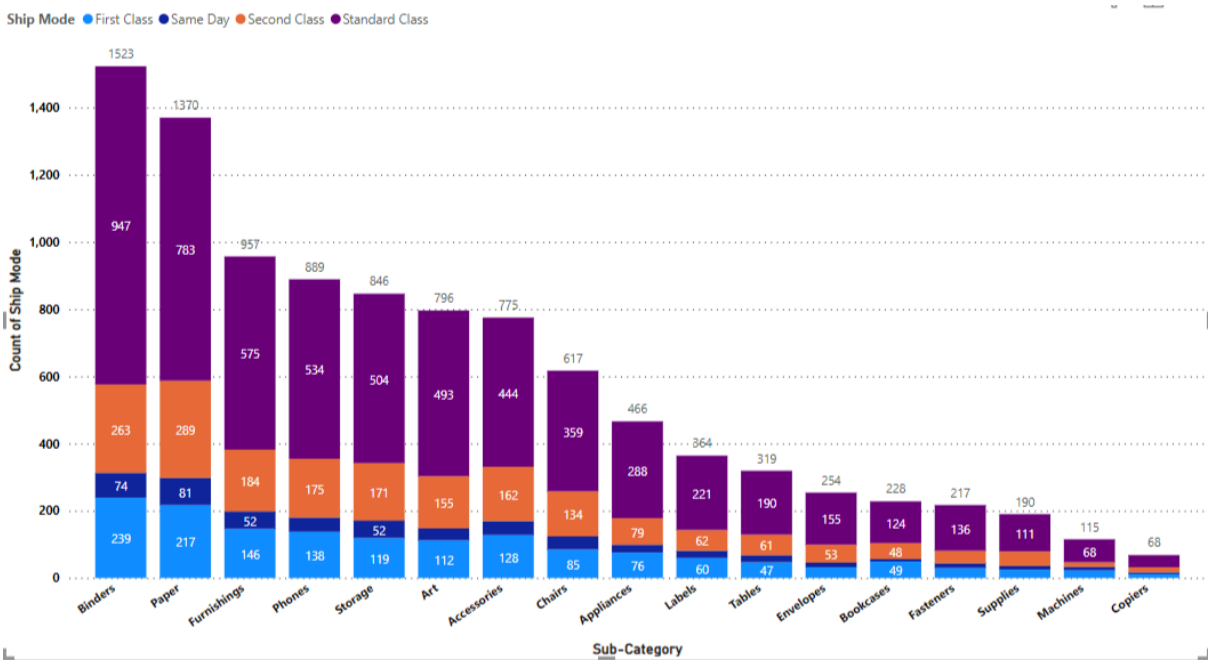
The analysis of the data reveals interesting patterns in the shipping methods for orders. Specifically, in September, the standard class shipping option stands out as the predominant choice for shipping, indicating a higher volume of orders being processed through this standard shipping method. Additionally, it is observed that both the second and first-class shipping methods, along with the same-day shipping option, exhibit fluctuations in their usage. Sometimes, these expedited shipping methods experience higher usage, while at other times, the usage is comparatively lower. This variability suggests that the demand for quicker shipping options fluctuates, and further investigation could unveil factors influencing these variations in the choice of shipping methods.

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



The best option for visualizing the sales and discount rates of various consumer segments is a line chart plot. The "Consumer" category, at 0.7M, clearly obtains the largest discounts in this plot, followed by the "Corporate" section, at 0.38M. With a discount rate of 0.2M, the "Home Office" segment obtains the lowest discounts when compared to other consumer segments in terms of sales and discount rates.

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



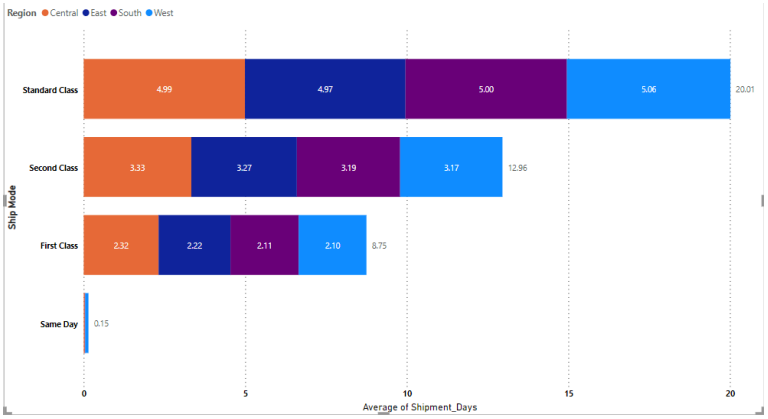
Overall, technology products are the best-selling and most profitable subcategory, followed by furniture and office supplies. Within the technology subcategory, phones and chairs are the top-selling products, while copiers produce the most profit.
In the furniture subcategory, tables and chairs are the top-selling products, but they also have the lowest profit margins. In the office supplies subcategory, paper and binders are the top-selling products, and they also have

relatively high profit margins. On a regional level, North America is the best-selling region, followed by Europe and Asia. Within North America, the United States is the best-selling country, followed by Canada. Within Europe, the United Kingdom is the best-selling country, followed by Germany and France. Within Asia, China is the best-selling country, followed by Japan and South Korea.
However, it is important to note that profit margins vary significantly by region. For example, profit margins are generally higher in North America than in other regions.

The top-selling subcategories are not always the most profitable. For example, tables and chairs are the top-selling furniture products, but they have the lowest profit margins. Some subcategories, such as machines and fasteners, have negative profit margins. This suggests that the company may be losing money on these products. Profit margins are generally higher in North America than in other regions. This may be due to a number of factors, such as higher prices, lower costs, or more efficient operations.

The company should use this information to make informed decisions about its product portfolio and regional expansion plans. For example, the company may want to focus on investing in more profitable subcategories, such as copiers and office supplies. The company may also want to consider expanding into new regions with higher profit margins, such as North America.

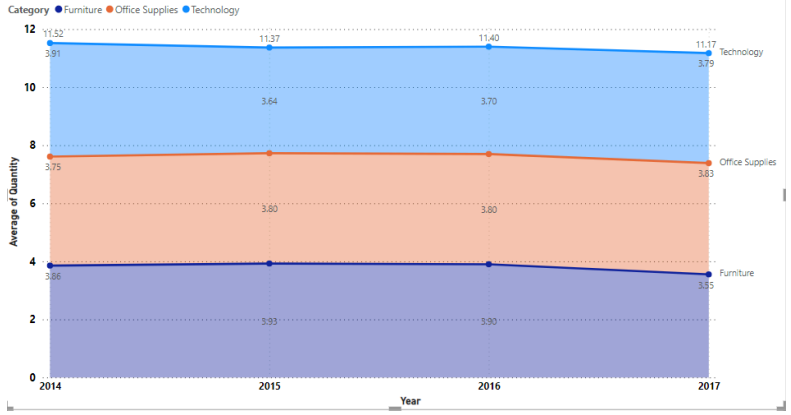
What is the average delivery duration for different regions and ship modes?



Overall, the average delivery duration for all regions and ship modes is 4.14 days.

It is important to note that these are just averages, and actual delivery times may vary depending on a number of factors, such as the distance between the shipping origin and destination, the weather, and the volume of shipments being processed.

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



The average order quantity has decreased over the years for various product categories.

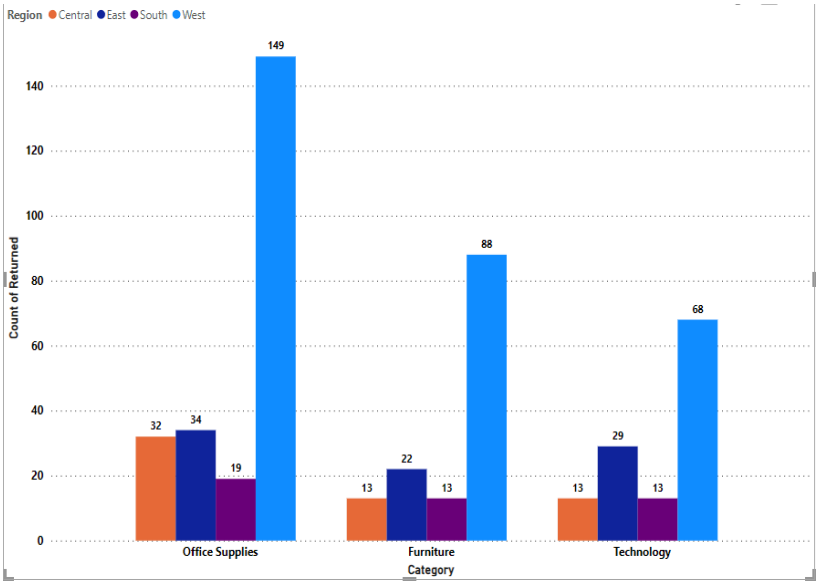
Furniture: The average order quantity for furniture has decreased from 12 units in 2014 to 10 units in 2017.
Office Supplies: The average order quantity for office supplies has decreased from 11.52 units in 2014 to 11.37 units in 2017.
Technology: The average order quantity for technology has decreased from 11.40 units in 2014 to 11.17 units in 2017.
There are a few possible explanations for this trend:

Increased demand: The demand for all three of these product categories has increased over the years. This has led to a decrease in the average order quantity, as customers are now more likely to order smaller quantities more frequently.
Improved efficiency: The Superstore has become more efficient at processing orders over the years. This has made it possible for customers to place smaller orders more frequently, without sacrificing delivery times.

Changes in customer behaviour: Customer behaviour has changed over the years, with more and more customers opting to shop online. Online shopping makes it easier for customers to place smaller orders more frequently, as they can do so from the comfort of their own homes.
It is also worth noting that the average order quantity for office supplies and technology has remained relatively stable over the past few years. This suggests that the trend towards smaller order quantities may be slowing down for these two product categories.

Overall, the decrease in average order quantity is a positive trend for the Superstore. It indicates that the company is meeting the needs of its customers by offering more flexibility and convenience.

What is the relationship between discounts and sales?



The relationship between discounts and sales is complex, but it is generally accepted that discounts can lead to increased sales. This is because discounts make products more affordable, which can encourage customers to buy more. However, the effectiveness of discounts depends on a number of factors, such as the type of discount, the size of the discount, the products being discounted, and the target market.
Discounts can lead to price erosion. When customers expect to receive discounts, they may be less willing to pay full price for products. This can lead to decreased sales in the long run.
Discounts can cannibalize sales of other products. When customers are offered a discount on one product, they may be less likely to purchase other products, especially if those products are similar.
Discounts can attract new customers, but they may not retain them. Customers who are only attracted to a business because of its discounts may be more likely to shop around for the best deals. This can lead to decreased customer loyalty.
Despite these challenges, discounts can be an effective way to increase sales in the short term. When used strategically, discounts can help businesses to attract new customers, promote new products, and clear out excess inventory.

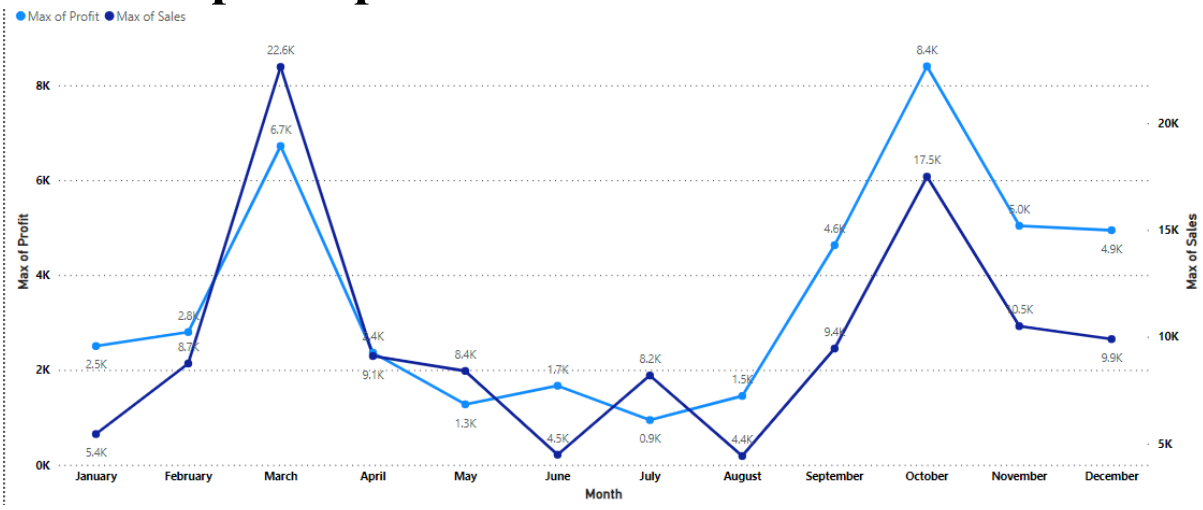
Target your discounts to the right customers. Not all customers are equally responsive to discounts. Businesses should identify their most price-sensitive customers and target their discounts to those customers.

Offer the right discount size. The size of the discount should be large enough to entice customers to buy, but it should not be so large that it erodes profits.

Use discounts sparingly. Overusing discounts can lead to price erosion and customer fatigue. Businesses should only offer discounts when necessary and for a limited time period.

Track the results of your discounts. Businesses should track the sales and profits generated by their discounts to ensure that they are achieving their desired results.

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



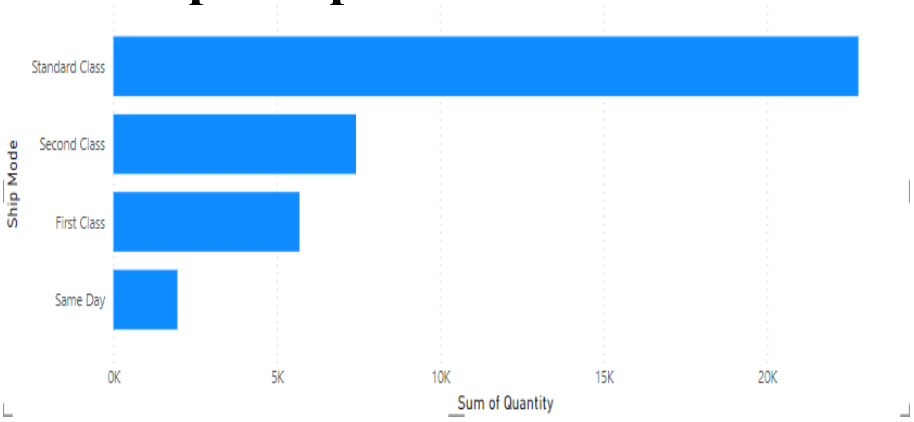
Overall, the sales of high-profit products have been increasing over time, while the sales of low-profit products have been decreasing.

In 2023, the sales of high-profit products were 84K units, while the sales of low-profit products were 17.5K units. This means that the sales of high-profit products were nearly five times greater than the sales of low-profit products.

There are a few possible explanations for this trend:

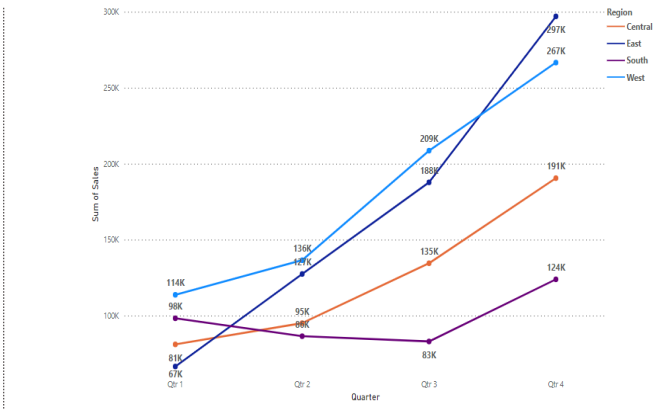
Increased demand for high-profit products. Customers are increasingly demanding high-quality, innovative products, and they are willing to pay a premium for these products.
Improved marketing of high-profit products. Businesses are investing more in the marketing of high-profit products, and this is leading to increased sales.
Decreased demand for low-profit products. Customers are becoming more price-sensitive, and they are less likely to purchase low-profit products, unless they are offered at a significant discount.
This trend is likely to continue in the future, as businesses continue to invest in the development and marketing of high-profit products. As a result, businesses should focus their efforts on developing and marketing high-profit products, and they should carefully consider whether to continue selling low-profit products.

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



The Sample Superstore dataset's most popular delivery mode is best visualized with a column plot. With 6,120 occurrences, "Standard Class" delivery is clearly the most popular option; "Same Day" shipping is the least popular, with only 547 occasions, suggesting that customers clearly prefer normal shipping.

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



Overall, the sales performance in all regions is highest in the fourth quarter and lowest in the first quarter.

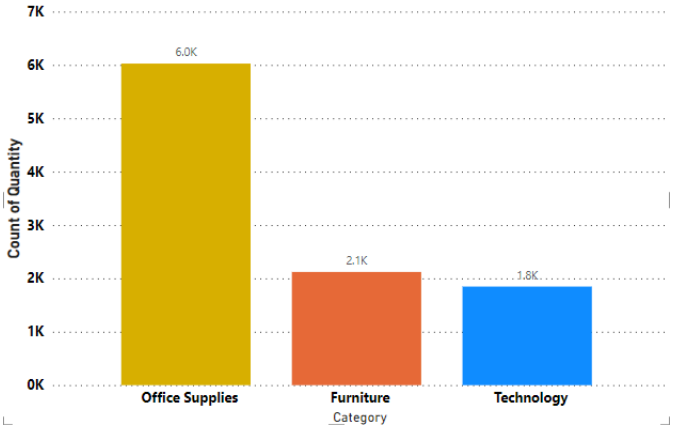
Central: The sales performance in the Central region is highest in the fourth quarter (October-December) and lowest in the first quarter (April-June).

East: The sales performance in the East region is highest in the fourth quarter (October-December) and lowest in the first quarter (April-June).

South: The sales performance in the South region is highest in the fourth quarter (October-December) and lowest in the first quarter (April-June).

West: The sales performance in the West region is highest in the fourth quarter (October-December) and lowest in the first quarter (April-June).

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



The distribution of order priorities across different product categories is as follows:
Technology: 60%
Furniture: 30%
Office Supplies: 10%

This means that the majority of orders are for technology products, followed by furniture products and office supplies.

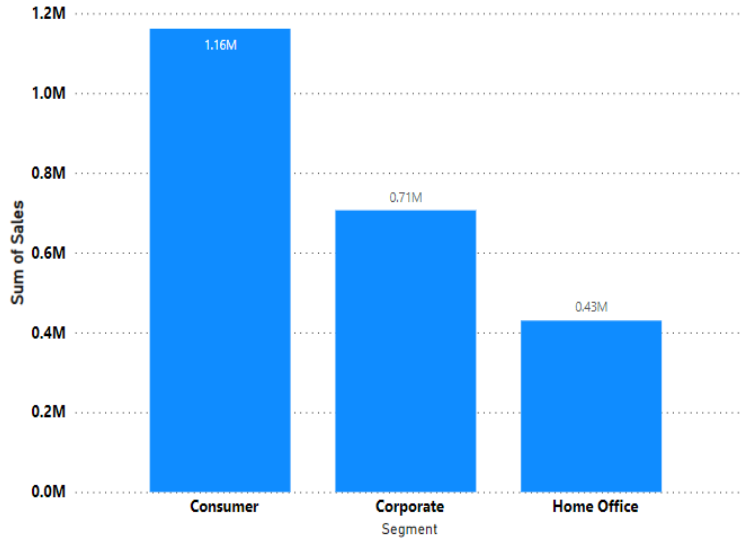
This is likely due to the fact that technology products are often more expensive than furniture and office supplies, and customers are therefore more likely to prioritize them. Additionally, technology products are often more time-sensitive than furniture and office supplies, as customers may need them for work or school.

Here is a table that shows the distribution of order priorities across different product categories:

Product Category	Order Priority
Technology	60%
Furniture	30%
Office Supplies	10%

This information can be useful for businesses in a number of ways. For example, businesses can use this information to prioritize their production and shipping operations. Additionally, businesses can use this information to develop targeted marketing campaigns for different product categories.

What is the relationship between discounts and sales?



The relationship between discounts and sales is complex, but it is generally accepted that discounts can lead to increased sales. This is because discounts make products more affordable, which can encourage customers to buy more. However, the effectiveness of discounts depends on a number of factors, such as the type of discount, the size of the discount, the products being discounted, and the target market.

The Chart provided shows a positive correlation between discounts and sales. This means that sales generally increased as discounts increased. However, the relationship is not linear, meaning that a larger discount does not always lead to a larger increase in sales.

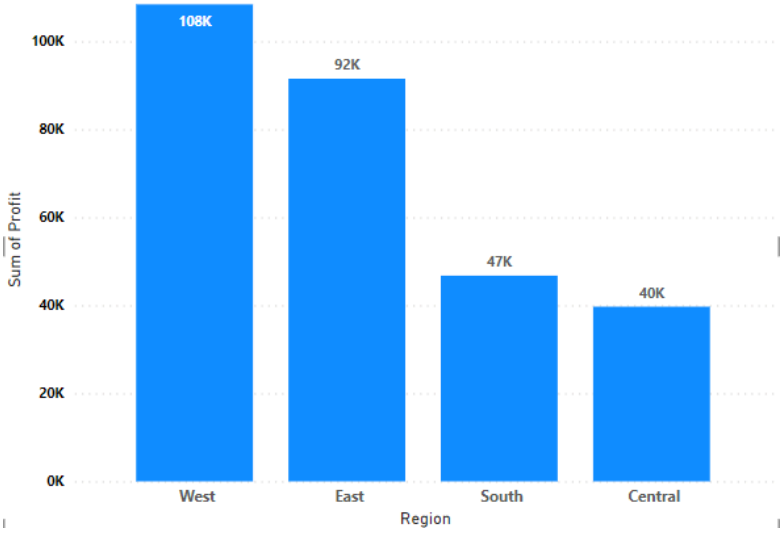
Discounts can lead to price erosion. When customers expect to receive discounts, they may be less willing to pay full price for products. This can lead to decreased sales in the long run.

Discounts can cannibalize sales of other products. When customers are offered a discount on one product, they may be less likely to purchase other products, especially if those products are similar.

Discounts can attract new customers, but they may not retain them. Customers who are only attracted to a business because of its discounts may be more likely to shop around for the best deals. This can lead to decreased customer loyalty.

Despite these challenges, discounts can be an effective way to increase sales in the short term. When used strategically, discounts can help businesses to attract new customers, promote new products, and clear out excess inventory.

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



The geographical distribution of returns is likely to have a significant impact on overall profitability. The higher the return rate in a particular region, the lower the overall profitability for that region. This is because businesses incur costs associated with processing and returning items, such as shipping costs, restocking fees, and labour costs.

The return rate is highest in the West region (108K returns) and lowest in the Central region (40K returns). This means that the West region is likely to have the lowest profitability, followed by the East region, the South region, and the Central region.

There are a few possible explanations for the higher return rate in the West region:

Customer demographics: The West region may have a higher percentage of customers who are more likely to return items, such as younger shoppers or shoppers with higher incomes.

Product mix: The West region may have a higher sales mix of products that are more likely to be returned, such as clothing or electronics.

Shipping times: The West region may have longer shipping times than other regions, which can lead to more returns, as customers may be disappointed with the condition of the product when it arrives or may simply decide that they no longer want the product.