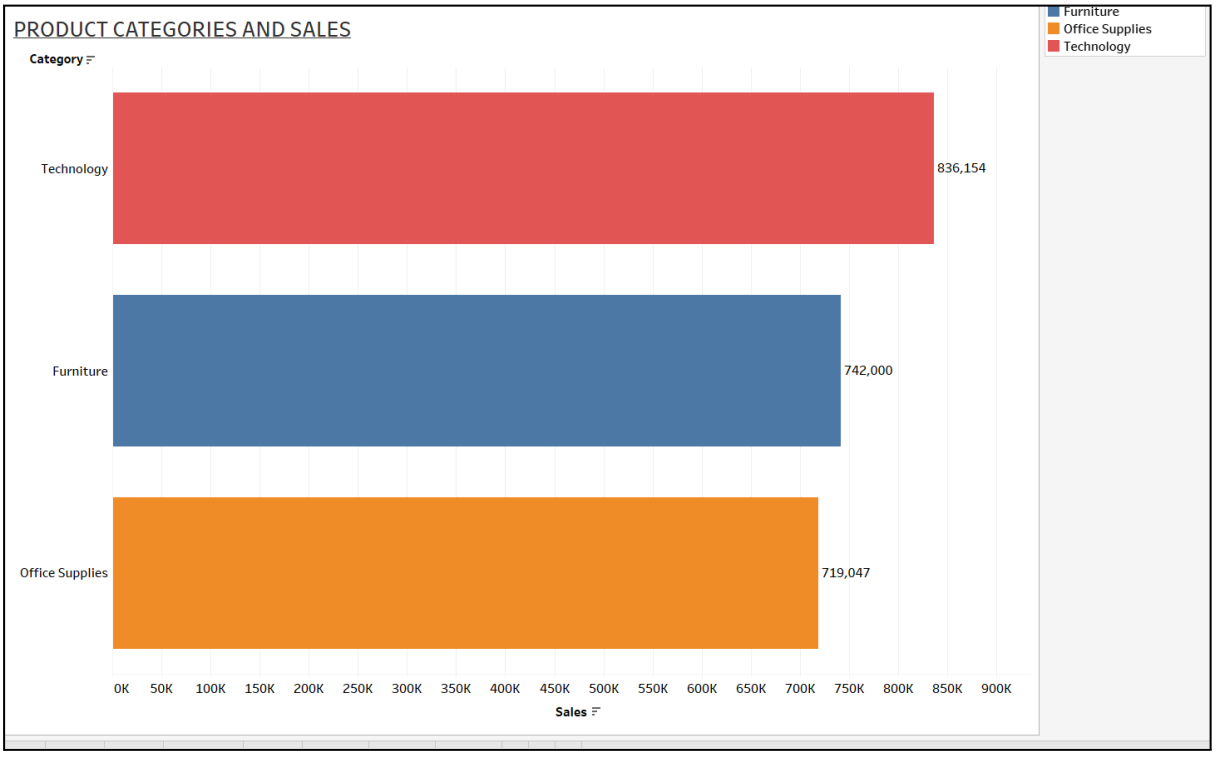
DATA VISUALISATION AND ANALYSIS PROJECT

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

1. Which product categories have the highest total sales in the "Superstore" dataset?
   Since we are dealing with categorical data (product categories) and numerical data (sales), a bar chart is a natural choice. It visually emphasises the relationship between the category and its corresponding sales.

   Bar charts are excellent for comparing different categories. Each bar represents a category, and the length or height of the bar corresponds to the total sales, making it easy to compare the sales of different categories at a glance.

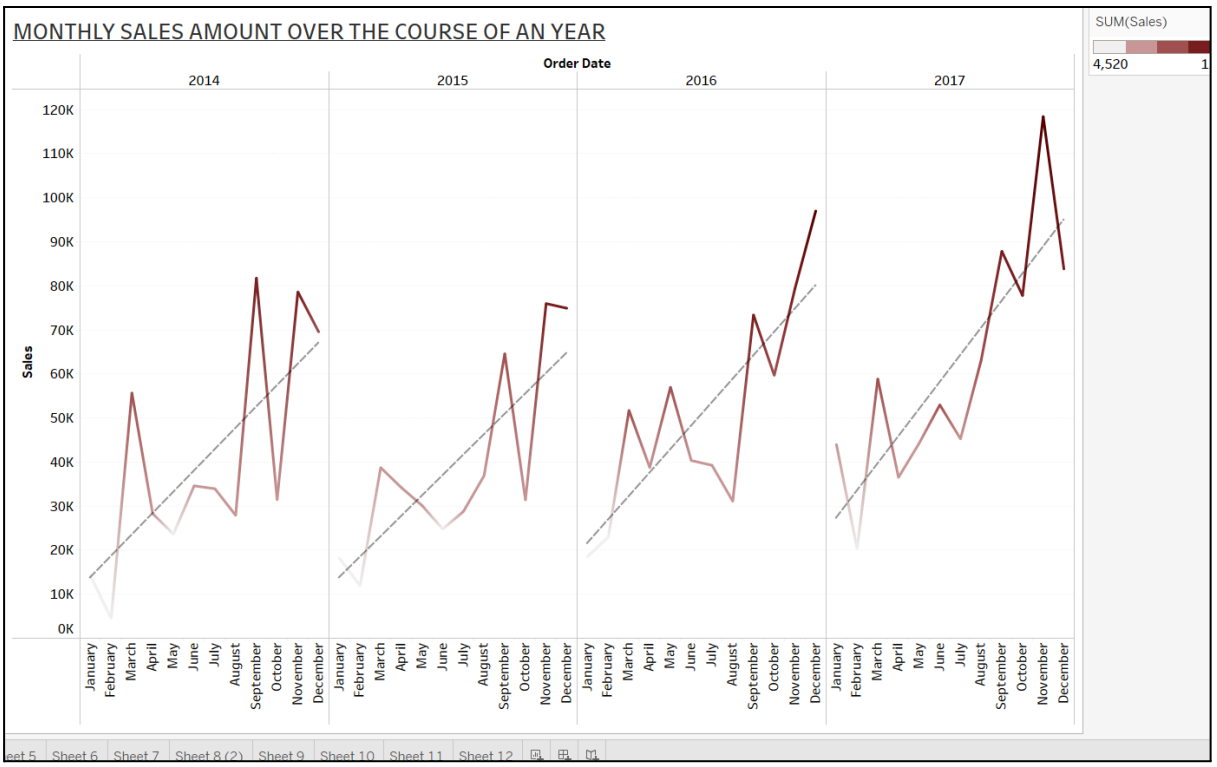


   Products from the technology category exhibit the highest sales according to the chart. This indicates a strong market demand and consumer interest in

technological products. Investing in and expanding this category could further boost overall sales and revenue.

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

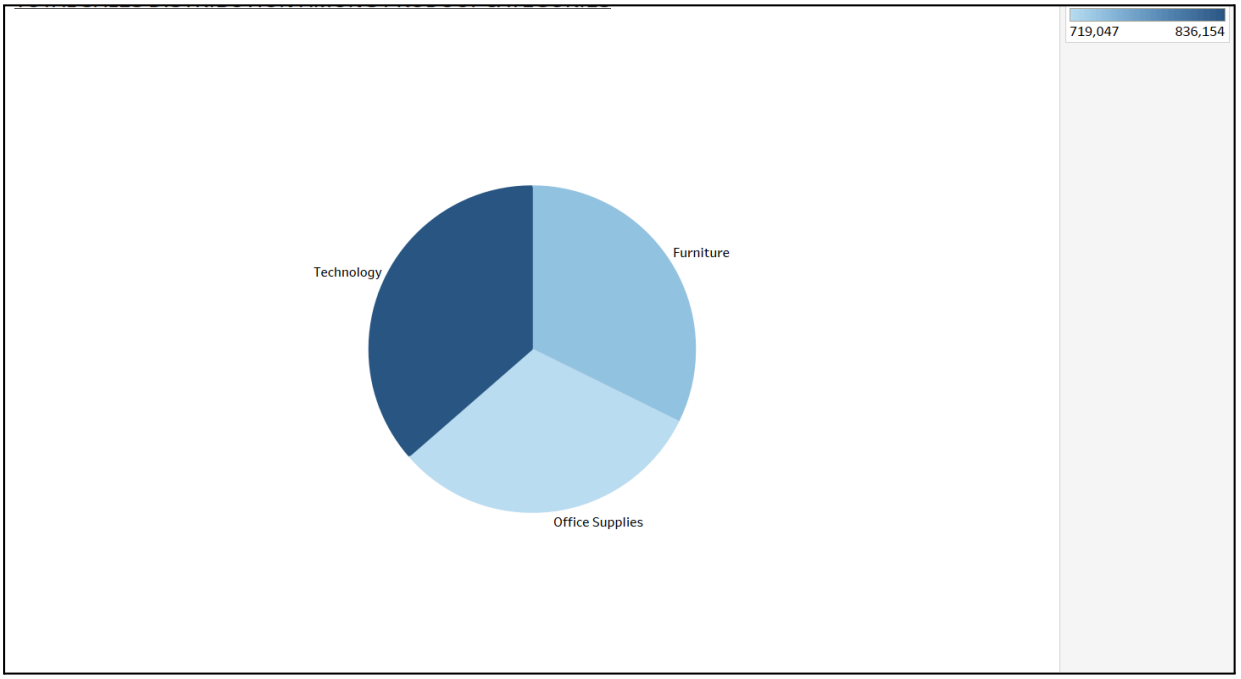
   A line chart is an excellent choice to visualise the monthly sales amounts over the course of a year. This type of chart clearly displays trends and fluctuations over time, making it easy to identify patterns such as seasonal peaks, dips, and overall growth or decline. By connecting data points with lines, it provides a continuous view of the sales performance, highlighting any changes in momentum and allowing for a straightforward comparison between different months.



   The line chart reveals that sales start slowly in January but pick up momentum afterward, exhibiting some fluctuations throughout the year. Notably, there is a significant dip in sales every October, a recurring pattern observed annually.

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

A pie chart is an ideal choice to illustrate how the total sales amount is distributed among different product categories. This type of chart provides a clear and immediate visual representation of the proportional contributions of each category to the overall sales. By displaying each category as a slice of the pie, it allows for easy comparison of the relative sizes and highlights the dominant categories as well as those with smaller shares, making the distribution of sales across categories intuitively understandable.

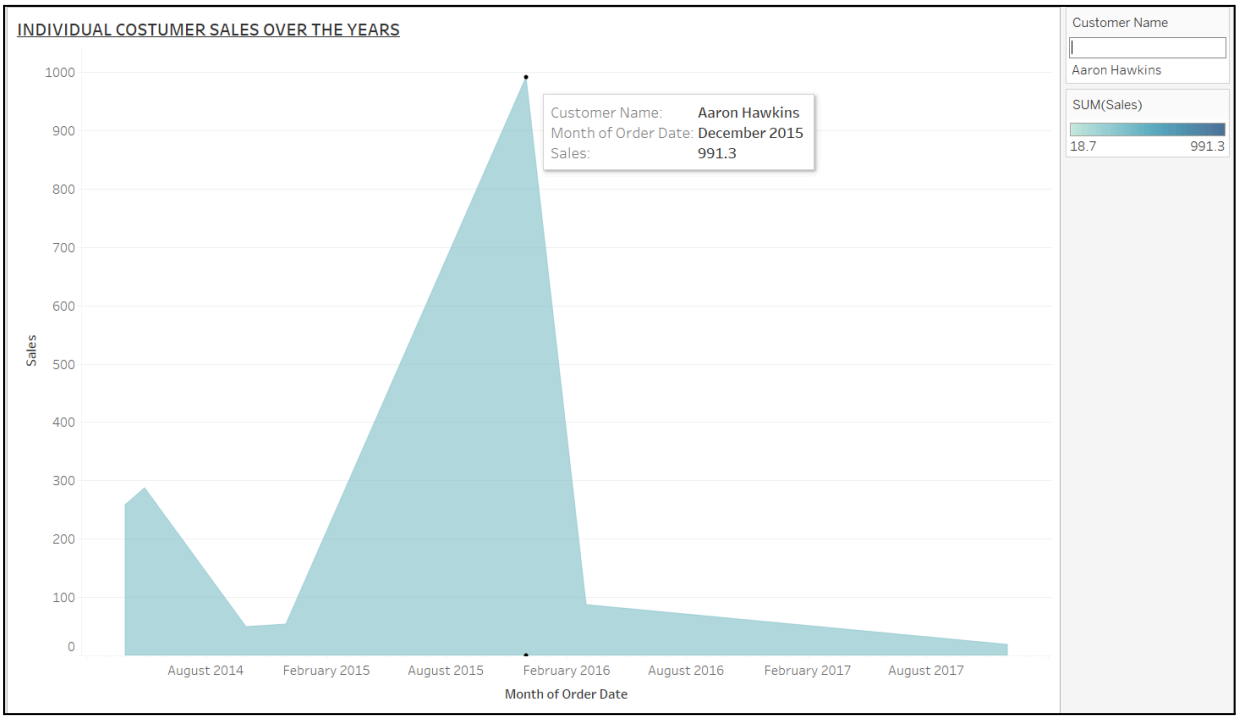


We have already observed that the technology category holds the largest contribution among all the categories,followed by furniture then office supplies.

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

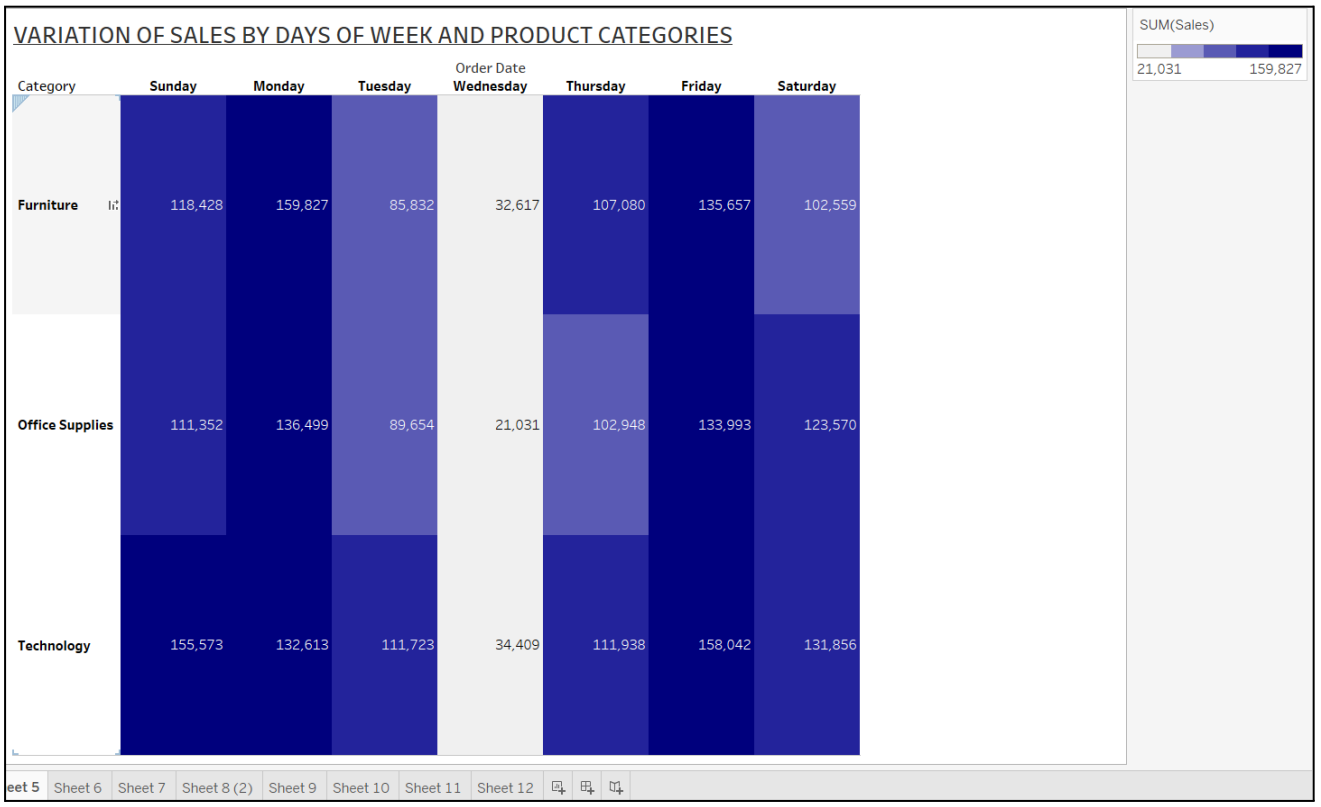
   Using a line chart with the capability to search for individual names and view their performance over time is a practical solution for analysing large datasets. This interactive approach allows users to focus on specific customers, providing a detailed and visual representation of their performance. It enables easy identification of trends, patterns, and changes in individual behaviour over

time, making the data more manageable and actionable. This method combines the clarity of line charts with the flexibility of search functionality, facilitating a comprehensive analysis of customer performance.



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

   A heatmap is an excellent choice for analysing how sales vary based on different days of the week and product categories. This type of chart uses colour gradients to represent data values, allowing for a quick and intuitive comparison of sales patterns across multiple dimensions. With a heatmap, you can easily identify trends and anomalies, such as which days of the week are most profitable for certain product categories, and which combinations are underperforming. The visual intensity of the colours provides an immediate understanding of the data distribution and highlights significant variations effectively.

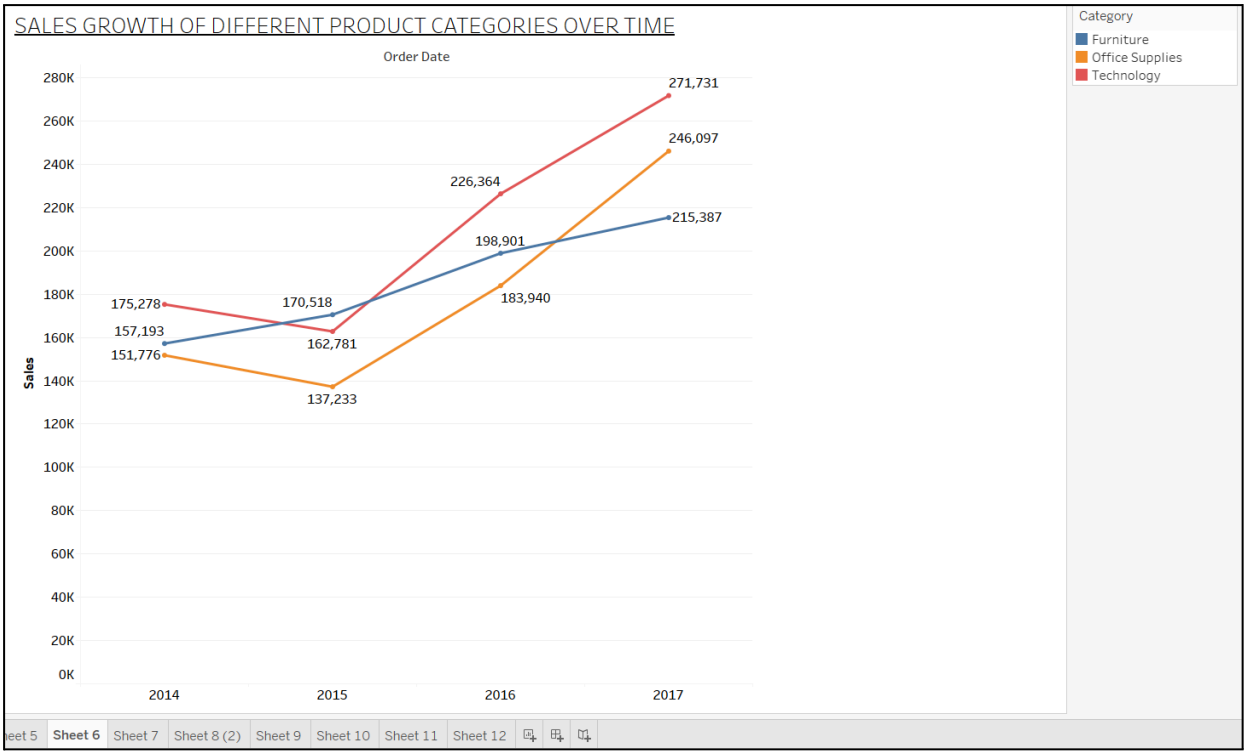


The heatmap clearly indicates that Wednesday is consistently a slow day for sales across various product categories. In contrast, Monday and Friday stand out as the days with the highest sales volumes. This pattern suggests potential opportunities for targeted promotions or strategic planning to boost mid-week sales, while also leveraging the high performance at the beginning and end of the workweek to maximise revenue.

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

   We can visualise the sales growth of different product categories over time using a line chart. A line chart is particularly effective for this purpose because it clearly shows trends and changes over a continuous time period. By plotting each product category as a separate line, it allows for easy comparison of growth rates and performance between categories. The continuous nature of the lines helps to highlight patterns such as consistent growth, seasonal

fluctuations, or periods of decline, providing valuable insights into the dynamics of each product category's sales over time.

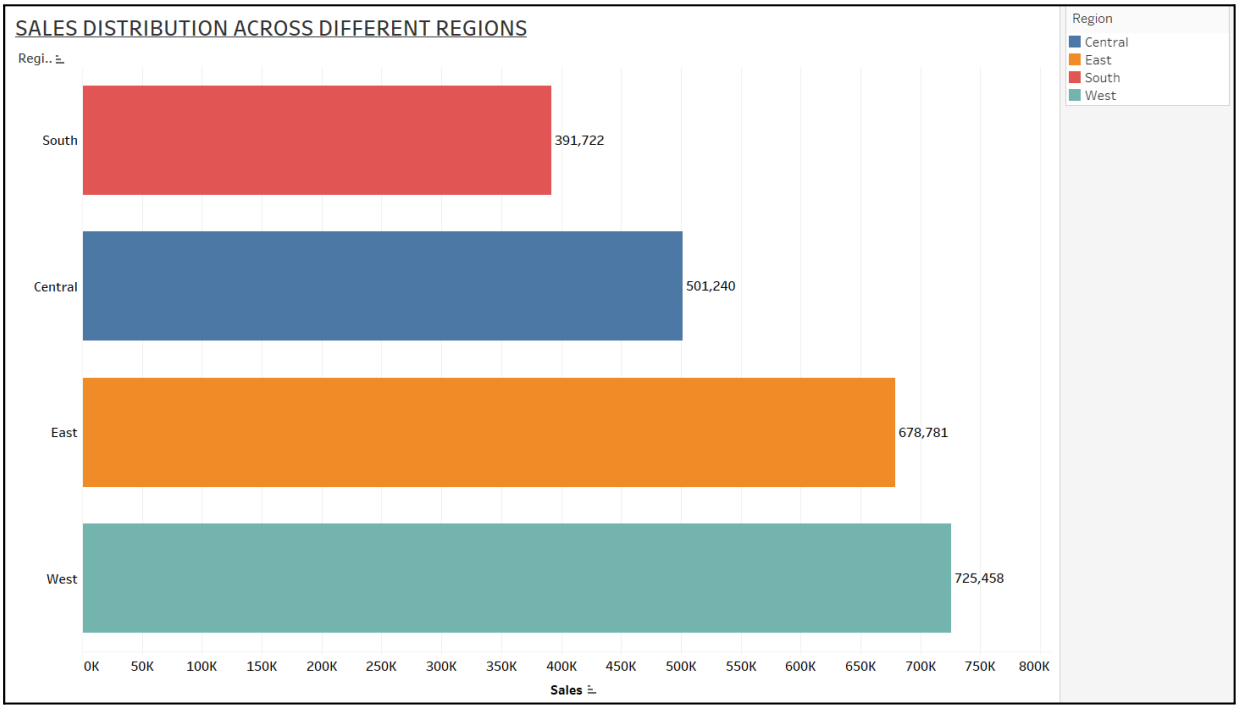


Technology products began with a stronger performance compared to other categories and have maintained their lead over time. Initially, furniture and office supplies exhibited similar sales figures, but office supplies, after experiencing a decline, have now surpassed furniture in terms of sales. This evolution suggests dynamic shifts in consumer preferences or market conditions, emphasising the importance of ongoing analysis and adaptation to capitalise on emerging trends and opportunities within each product category.

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

   Bar charts are effective for comparing values across categories, making them ideal for showcasing the sales performance of various regions. Each region can be represented by a separate bar, with the height of the bar indicating the sales amount. This allows for a clear comparison of sales figures between

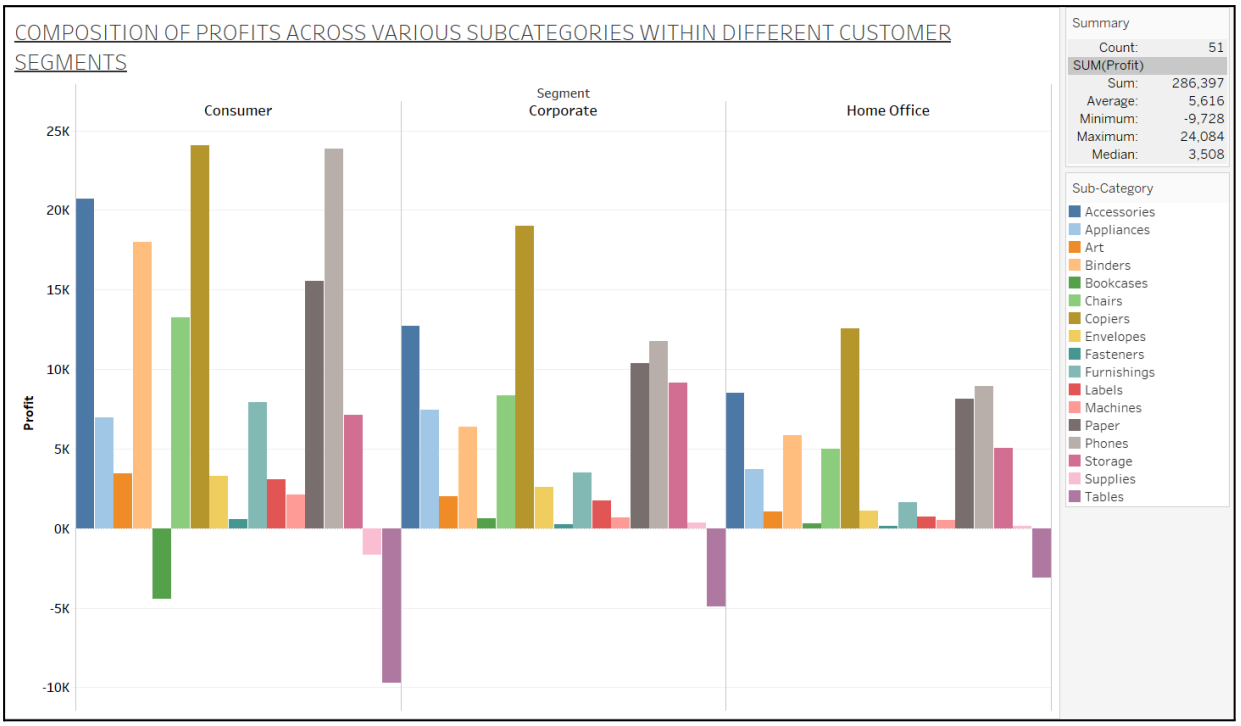
regions, making it easy to identify which regions are top performers and which may require further attention or analysis. Additionally, bar charts can accommodate large datasets with multiple categories, making them suitable for displaying the sales distribution across different regions comprehensively.



The bar chart reveals that the West region leads in sales, followed closely by the East, which performs comparably well. Meanwhile, the Central region outperforms the South, which lags behind significantly in terms of sales. This distribution underscores regional variations in sales performance, highlighting potential areas for strategic focus or investment to enhance sales growth across all regions.

8. Can we visualise the composition of profits across various subcategories within different customer segments?
   A side-by-side bar chart is a very appropriate method to visualise the composition of profits across various subcategories within different customer segments. This type of chart allows for a clear and direct comparison of profit compositions between different customer segments, facilitating insights into how profits are distributed across subcategories within each segment. By placing the bars side by side, it becomes easy to identify any disparities or

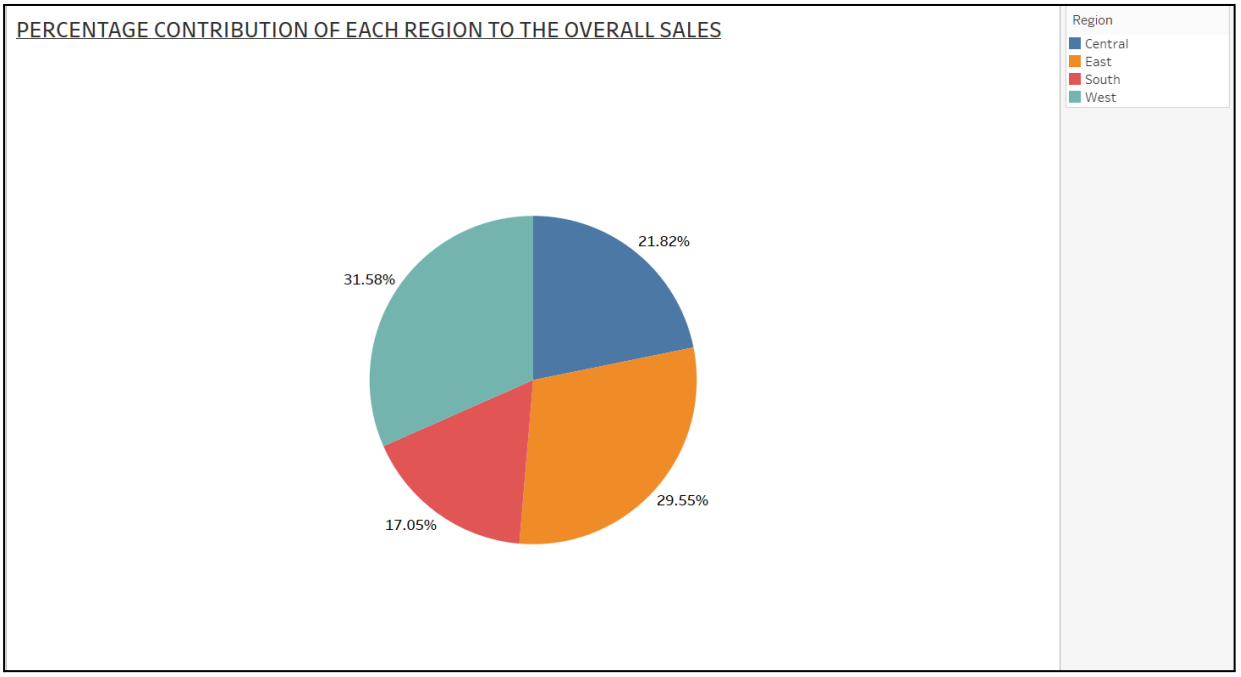
similarities in profit distribution across the subcategories within each customer segment, enabling informed decision-making and targeted strategies to maximise profitability. Additionally, this visualisation method is effective for presenting complex data in a visually appealing and easy-to-understand format, making it accessible to a wide audience.



The side-by-side bar chart shows that copiers generate the highest profit across all three customer segments, with the consumer segment leading the way. Phones also perform well in all segments, particularly excelling in the consumer segment. Accessories follow a similar trend, showing strong profits across the board, especially in the consumer segment. In stark contrast, tables are underperforming in every segment, with the consumer segment experiencing the highest negative profit. Bookcases have a moderate profit performance in the home office and corporate segments but struggle significantly in the consumer segment, which is an interesting deviation worth further investigation. These insights highlight key areas of profitability and concern, offering opportunities for strategic adjustments in product focus and marketing efforts.

9. What is the percentage contribution of each region to the overall sales?
   A pie chart is an ideal choice for visualising the percentage contribution of each region to the overall sales. This type of chart effectively illustrates the

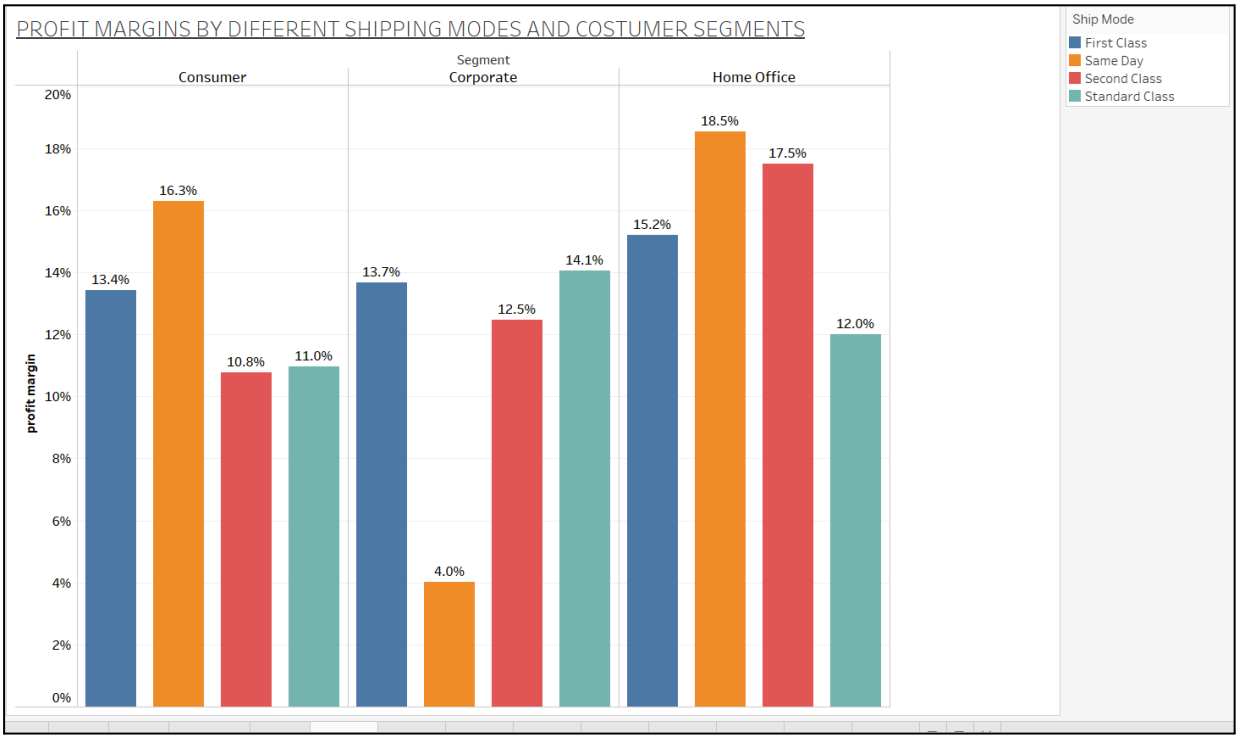
proportion of total sales that each region contributes, providing a clear and immediate visual representation of the distribution. The pie chart segments can be labelled with percentages, making it easy to compare the relative size of each region's contribution and quickly identify which regions are the major contributors and which are less significant. This visual approach simplifies the understanding of regional sales distribution and aids in strategic decision-making.



This is the same as the answer to question number 7 but in percentage contribution.

10. Can we visualise the profit margins associated with different shipping modes and customer segments?
    A bar chart is yet again an appropriate choice for visualising the profit margins associated with different shipping modes and customer segments. This type of chart provides a clear and straightforward comparison of profit margins across multiple categories, making it easy to see how different shipping modes perform within each customer segment. By using separate bars for each shipping mode within each segment, the bar chart allows for direct comparison of their profitability. This method highlights variations and trends in profit

margins, aiding in the analysis and decision-making process regarding shipping strategies and customer segment targeting. Additionally, bar charts are easy to read and interpret, making them a practical choice for presenting complex data in a simplified manner.



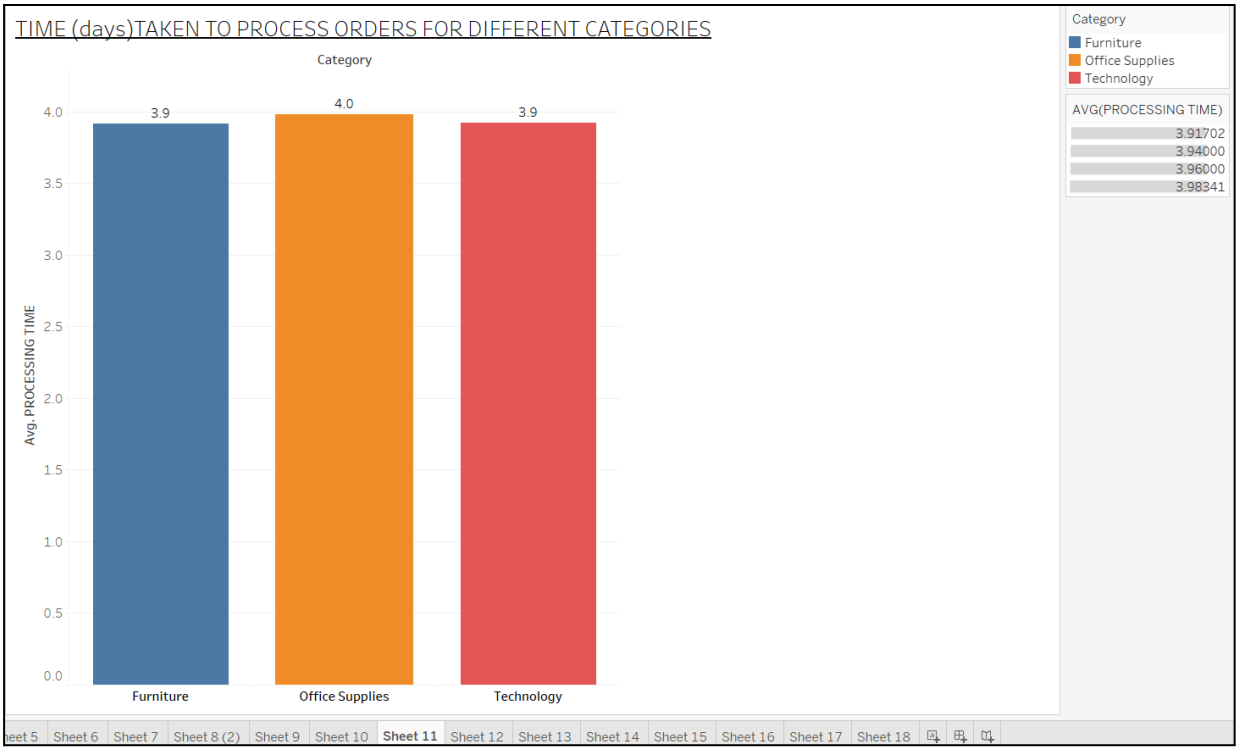
Promote Fast Shipping: For consumer and home office segments, same day shipping is highly profitable. Promotional strategies highlighting the convenience and speed of same day delivery can drive higher margins.

Value-Driven Corporate Shipping: For corporate customers, emphasising the reliability and cost benefits of standard class shipping will likely be most effective, given their preference for standard over faster, more expensive options.

Balanced Approach for Home Office: Offering a mix of fast and economical shipping options can cater to the varied preferences within the home office segment, ensuring higher profit margins by meeting diverse needs.

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

A bar chart is an ideal and straightforward choice for visualising the processing times for different product categories. This type of chart effectively compares the duration it takes to process orders across various categories, making it easy to identify which categories have shorter or longer processing times. By displaying each product category as a separate bar, the chart provides a clear visual representation that facilitates quick understanding and analysis. This allows for identifying any bottlenecks or areas for improvement in the order processing workflow.



The bar chart reveals that the time taken to process orders is fairly consistent across all product categories. This uniformity suggests that the order processing system is standardised and efficient, regardless of the product type. However, it also indicates that there might be limited flexibility to expedite processing for specific high-demand categories. This consistency can be advantageous for maintaining predictable delivery schedules and managing customer expectations effectively. Additionally, the uniform processing times

highlight the potential for optimising specific areas to further enhance overall efficiency.

12. How do discounts affect overall profit?

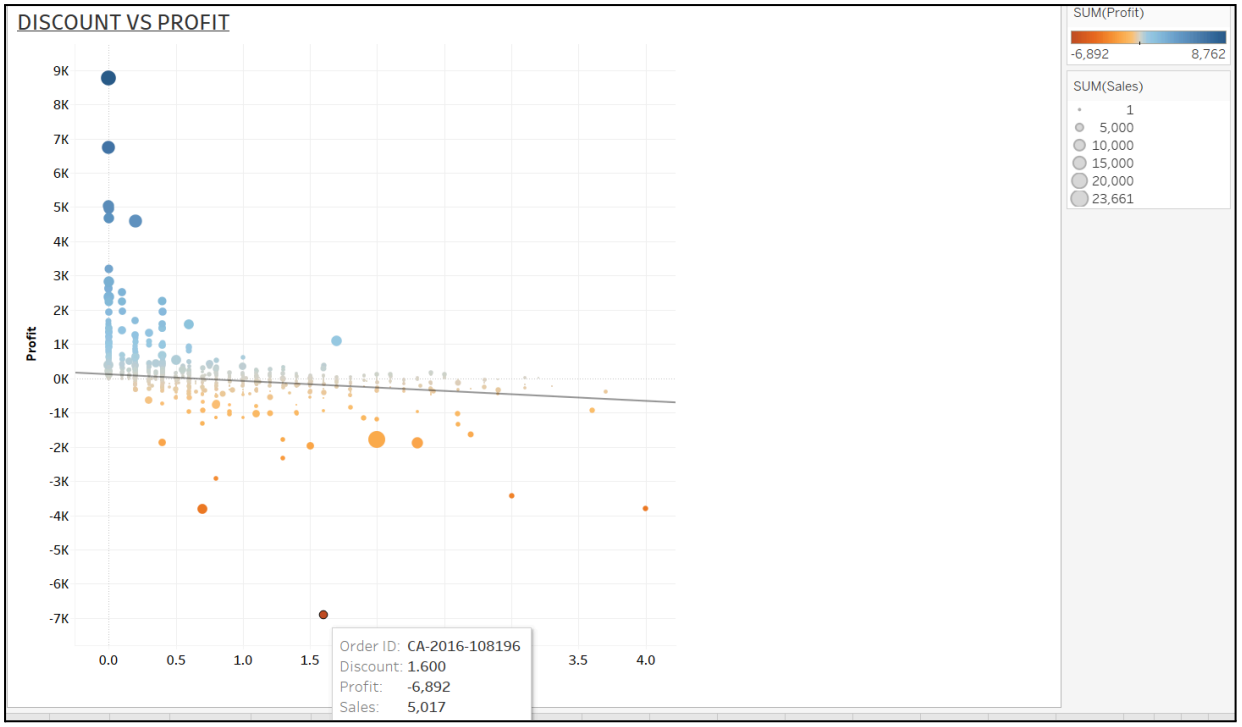
    Why scatter plot?

    Relationship Visualisation: Scatter plots are excellent for showing the relationship between two continuous variables, in this case, discount and profit. Each point represents an individual order, allowing you to see how profit changes with different levels of discount.

    Trend Identification: By adding a trend line, you can quickly identify whether there is a positive, negative, or no significant relationship between discount and profit.

    Outlier Detection: Scatter plots help in identifying outliers—orders where the discount had an unusually high or low effect on profit. This can be useful for further investigation.

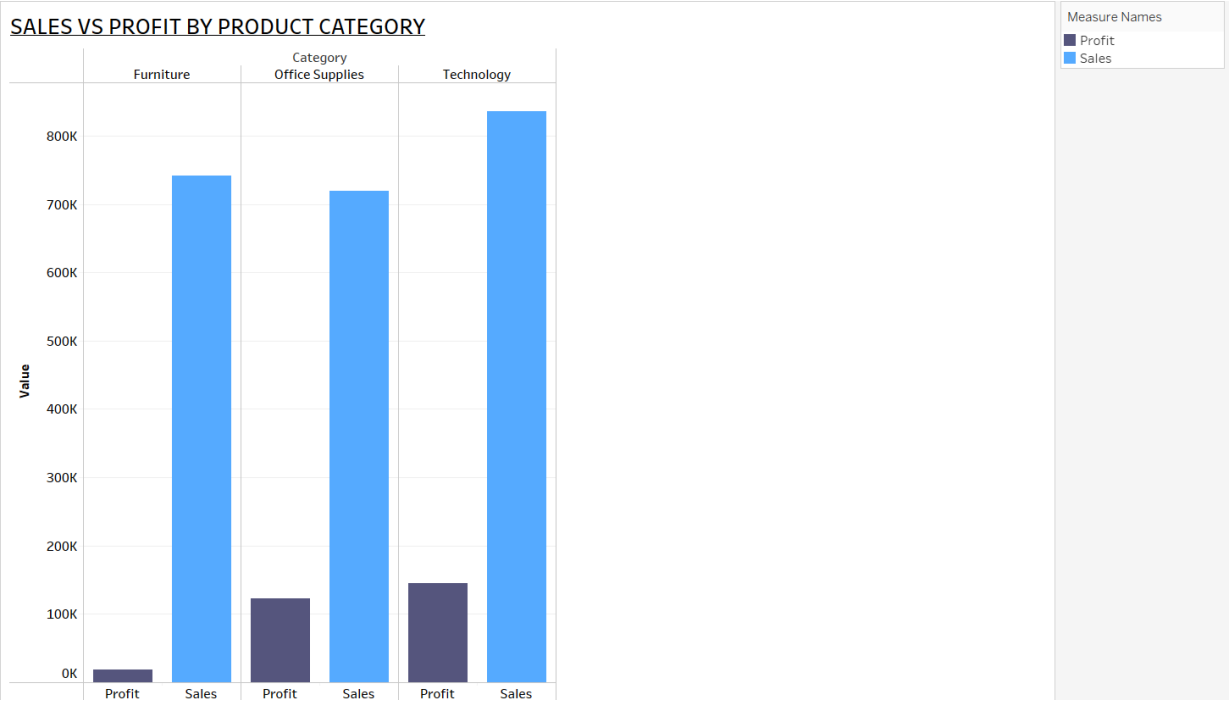
    Distribution and Variability: It shows the distribution of data points and how widely spread they are, indicating the variability in how discounts affect profits.



The analysis shows that orders generating the highest profits were typically associated with minimal discounts. However, it's important to note that some orders with relatively low discounts still failed to turn a profit. Despite these exceptions, the overall trend indicates that lower discounts tend to correlate with higher profitability. This suggests that while discounting can attract customers, maintaining a balance is crucial to ensure profitability. Businesses should consider strategically managing discount levels to maximise profit margins while still offering competitive pricing to attract and retain customers.

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

    Yes we can,and bar charts are here to the rescue again.using a bar chart is simple and straightforward and that is exactly what we want.By displaying separate bars for sales and profits within each category, the chart makes it easy to see how well sales translate into profitability. This visual representation can help identify categories where high sales do not necessarily lead to high profits, and vice versa, providing valuable insights for strategic decision-making and resource allocation.



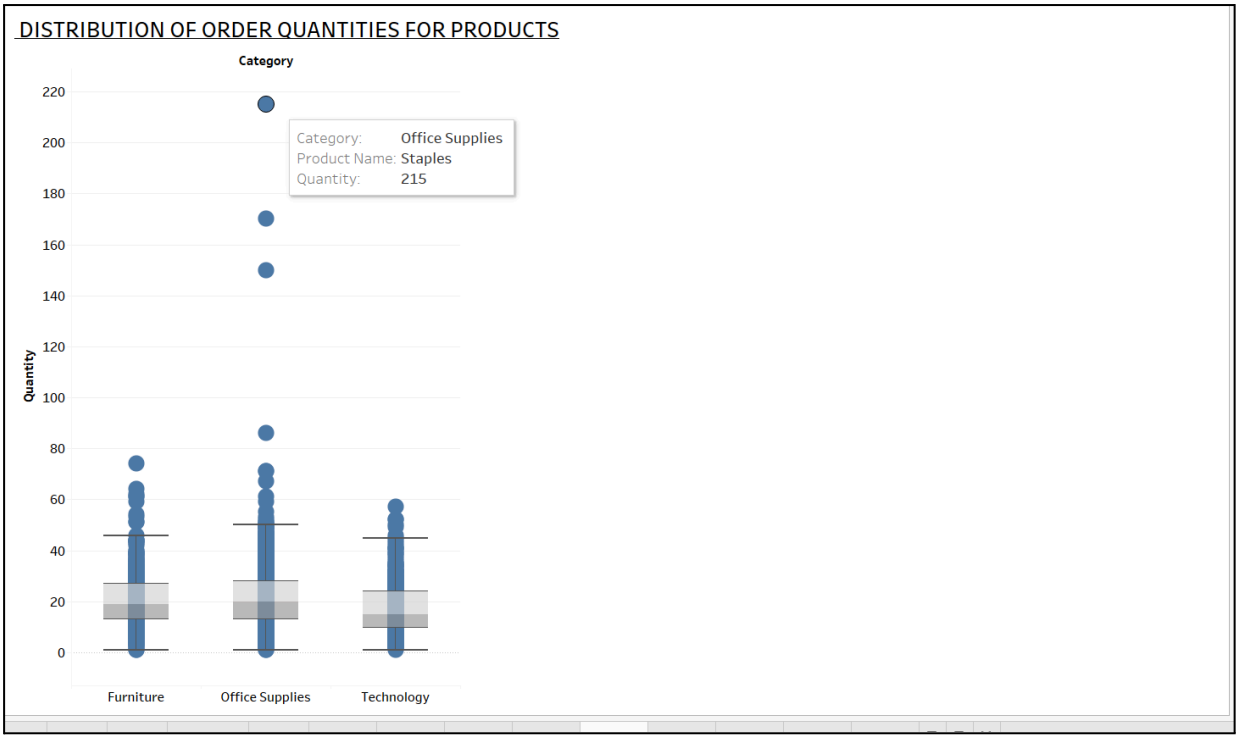
Technology has the highest sales and the highest profit ,closely followed by office supplies in terms of sales and profit.furniture,even though has sales slightly higher than office supplies but lacks very much on profit.

14. What is the distribution of order quantities for products in the dataset? Why box plot for this visualisation?

    Handles Large Data: Box plots summarise data distribution and handle large datasets effectively, avoiding clutter from too many data points.

    Shows Distribution: It displays the median, quartiles, and outliers, providing a clear summary of the distribution of order quantities.

    Comparison: Box plots allow for easy comparison between different products or categories.

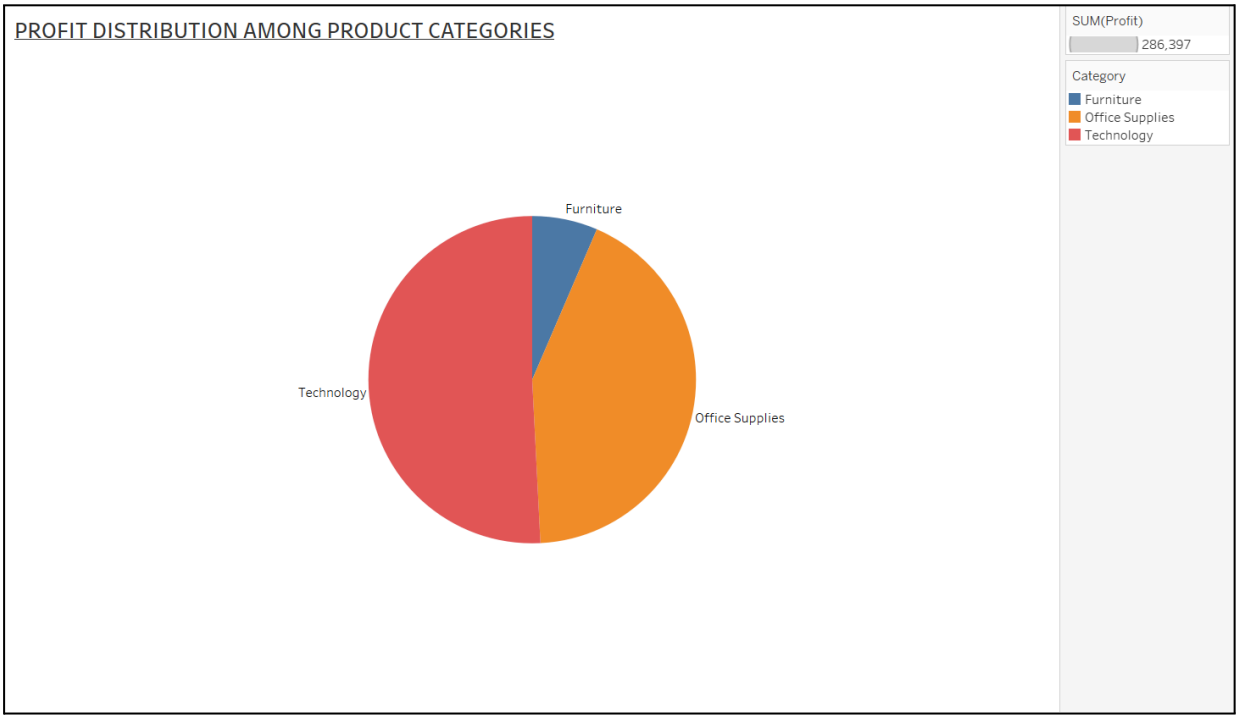


The tooltip feature greatly enhances the ability to quickly access detailed information from the chart. From this chart, we can conclude that staples in the office supplies category have the highest order quantities, followed by staple envelopes. In fact, the top four items in order quantity all belong to the office supplies category. Furniture and technology categories closely follow each other in terms of order quantity. However, it's important to note that technology products lead in both sales and profitability, underscoring their dominant position in the market. This insight highlights the significant contribution of office supplies to order volume, while technology drives overall revenue and profit.

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

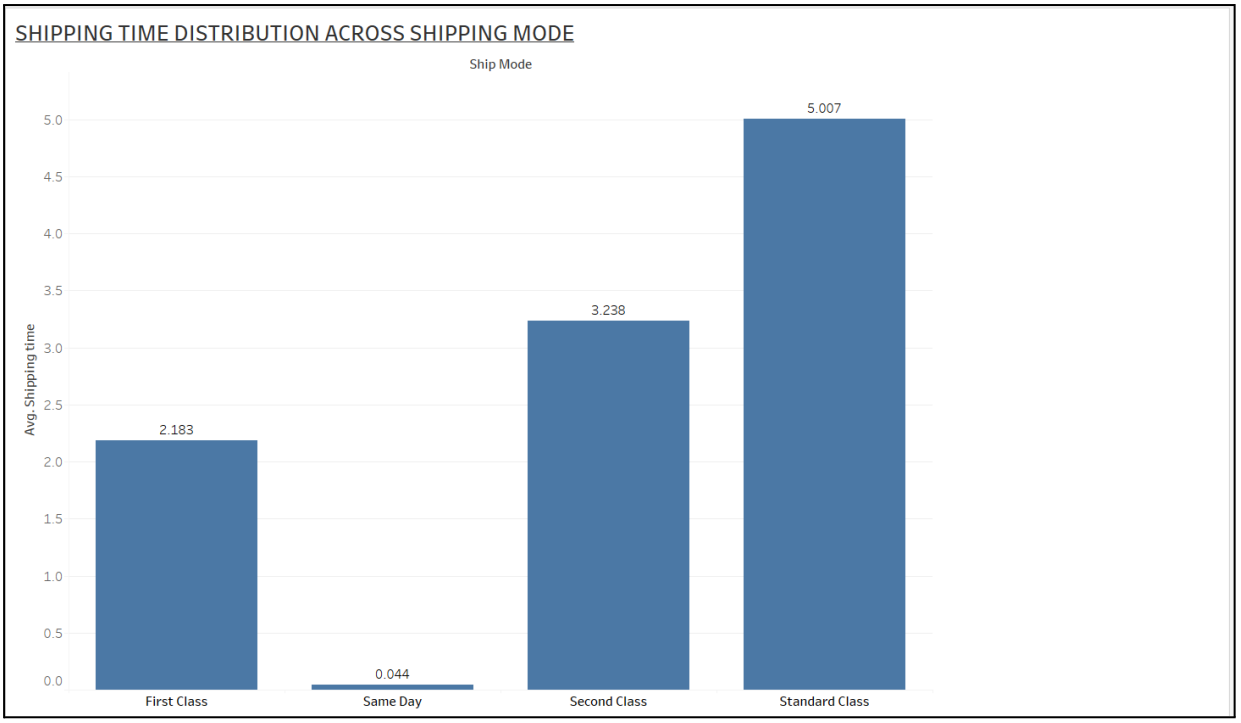
    A pie chart would be an ideal choice for visualising the profit distributions across different product categories. This type of chart effectively displays the proportion of total profits contributed by each category, making it easy to compare their relative profitability. By representing each category as a slice of

the pie, the chart provides a clear and immediate visual impression of which categories are the most and least profitable. This visual format simplifies the understanding of profit distribution and highlights key areas contributing to overall profitability.



The pie chart clearly shows that technology leads in profitability, followed by office supplies, with furniture trailing behind. While office supplies contribute significantly to overall profits, they still lag considerably behind technology. Furniture, although contributing to the total profit, has a substantial gap to close to match the profitability of technology and office supplies. This distribution underscores the dominant financial performance of technology products and highlights potential growth opportunities within the furniture category to enhance its profitability.

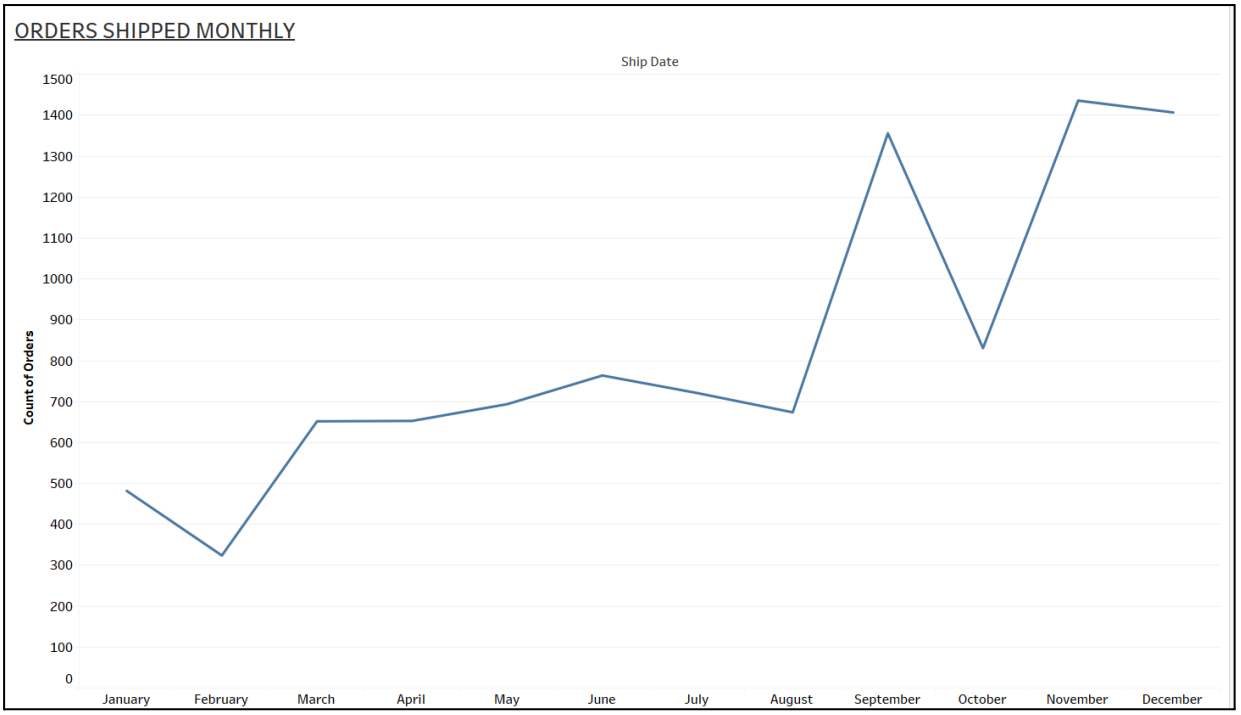
16. Can we compare the shipping time distributions for different shipping modes? Bar chart would be the simplest and efficient way to go.separate bars for each shipping mode and the length of the bar to show the shipping time.



The shipping time analysis reveals that same day shipping is the fastest, as expected. First class shipping typically takes 2 days, followed by second class, which takes slightly longer. Standard shipping has the longest delivery time, averaging 5 days. This hierarchy in shipping speeds underscores the varying levels of service and the trade-offs between cost and delivery time.

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

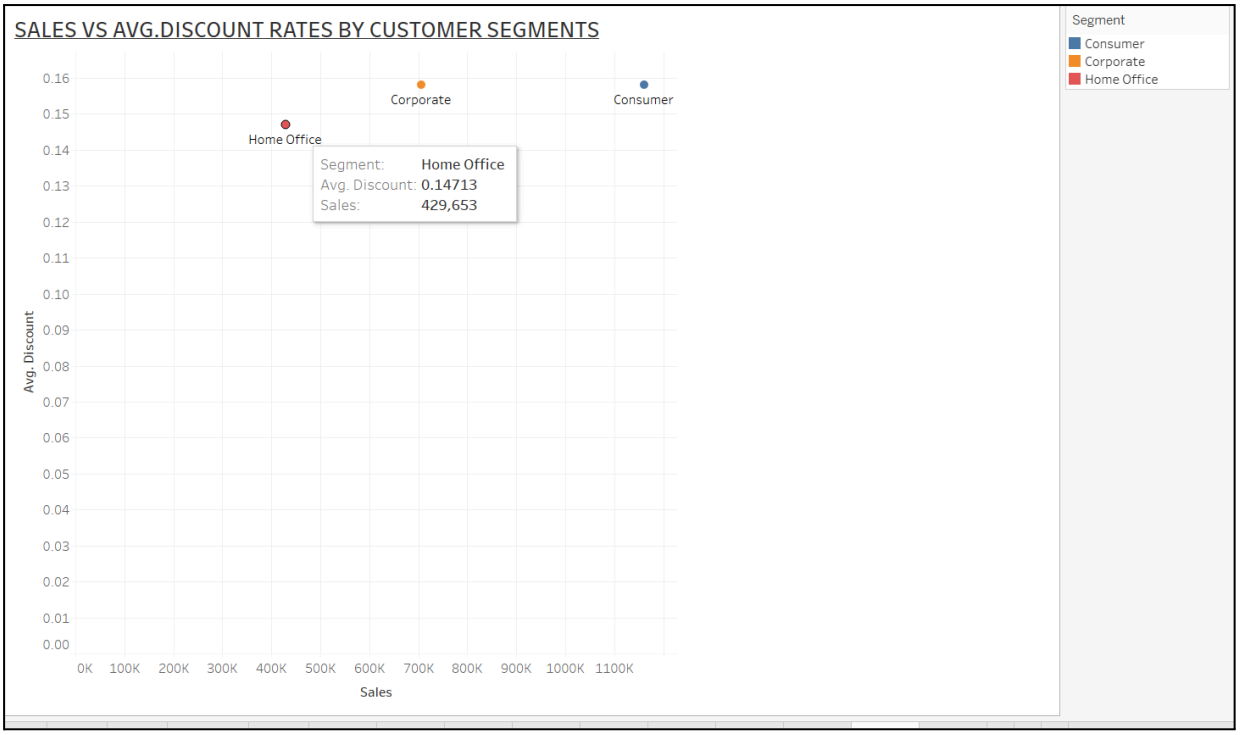
    A line chart is the most suitable choice for visualising the monthly trend in the number of orders shipped. This type of chart effectively displays changes over time, making it easy to identify trends, patterns, and fluctuations on a month-to-month basis. By connecting data points with lines, the chart provides a clear visual representation of how the number of orders varies throughout the year, allowing for straightforward comparison between different months and the identification of seasonal trends or anomalies.



The aggregate data from 2014 to 2017 shows a clear monthly trend in the number of orders shipped. The year starts with a slow count of orders, which begins to pick up pace around March. There is a notable spike in orders in September, followed by a dip in October. The order count then gradually recovers in November, ending the year with over 1300 orders in December. This trend highlights seasonal fluctuations, with significant increases in late summer and early fall, and a strong finish towards the end of the year.

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

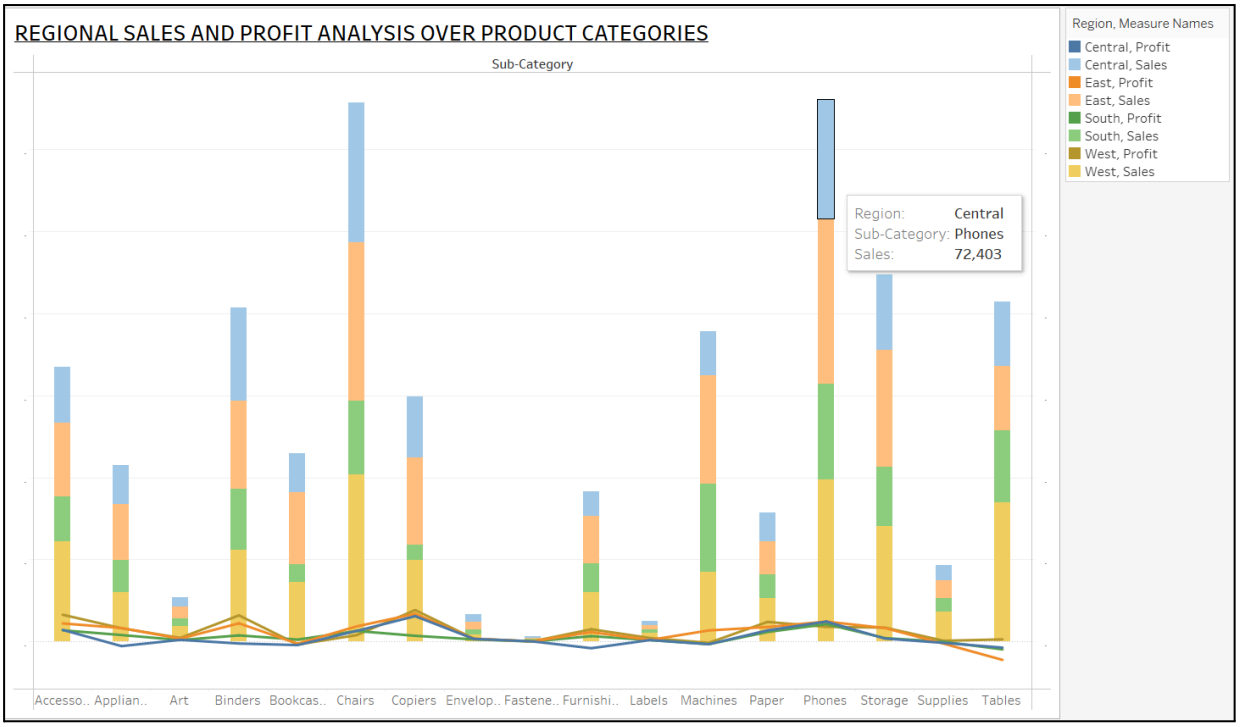
    A scatter plot is an ideal choice for visualising the relationship between different customer segments in terms of sales and discount rates. By plotting sales against discount rates for each customer segment, the scatter plot provides a comprehensive view of how each segment performs in terms of sales volume and the extent of discounts offered. This visual representation can reveal insights into which segments are more price-sensitive, which segments generate higher sales despite lower discounts, and how discount rates impact overall sales performance across different customer segments.

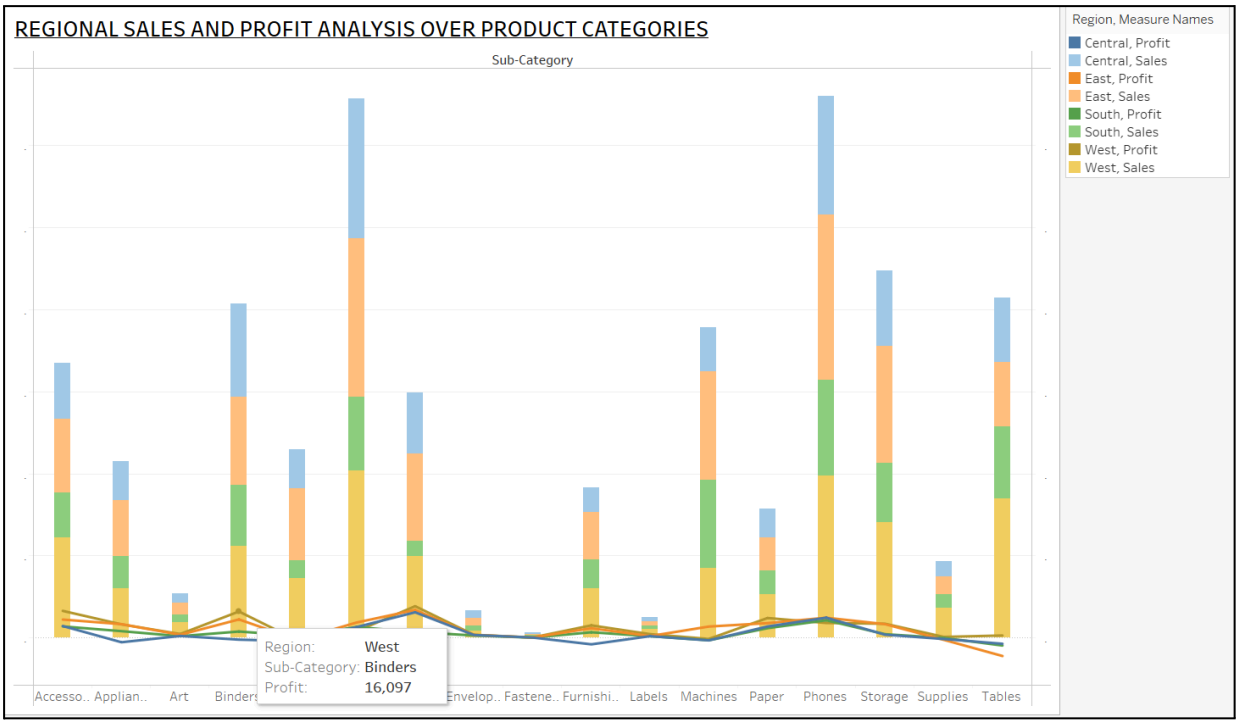


In terms of sales performance, the consumer segment stands out with the highest sales volume, despite offering comparatively higher discount rates. Following closely behind is the corporate segment, which also maintains a similar level of discount rates. Surprisingly, the home office segment achieves considerable sales figures with the lowest discount rates among the three segments. This observation suggests nuanced consumer behaviours across different segments, with the consumer segment possibly being more price-sensitive, while the home office segment prioritises value over discounts.

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

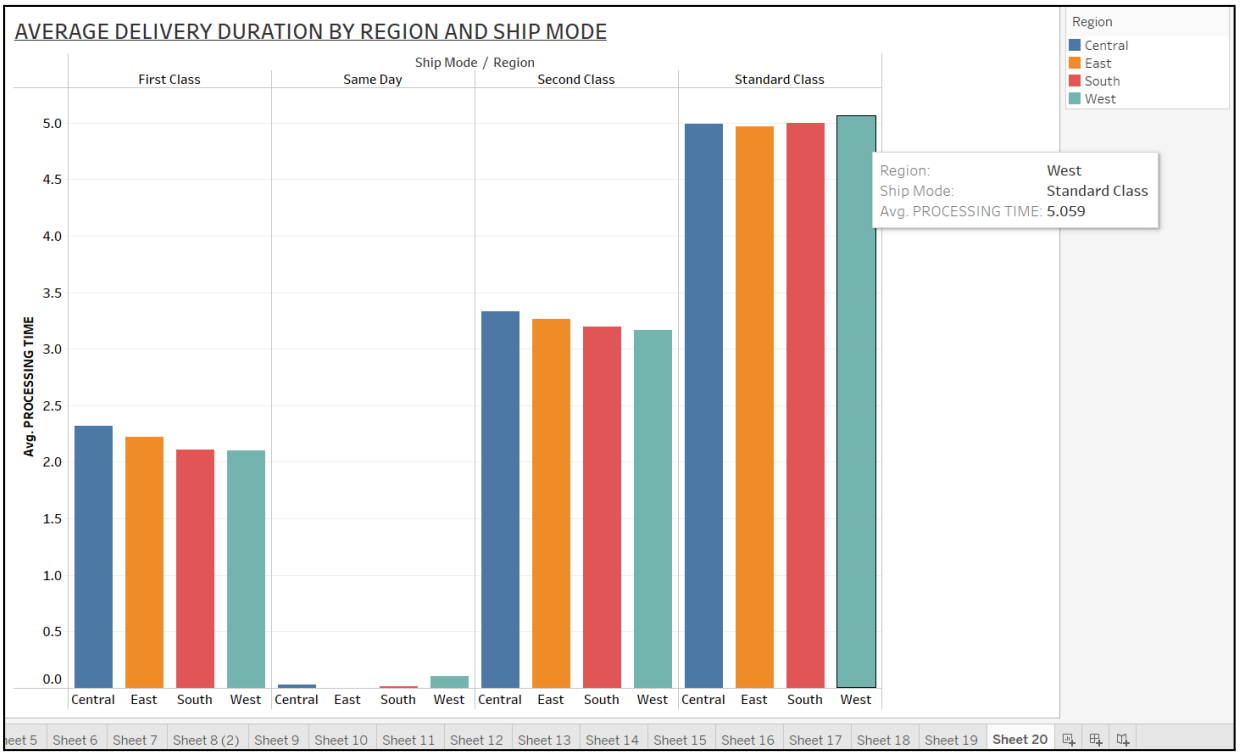
    A dual-axis chart with a stacked bar for sales and a line chart for profit margin across product subcategories and regions is justified due to its capacity to visualise intricate data relationships, compare sales and profitability trends simultaneously, highlight regional disparities, and identify correlations effectively.





The tooltip feature will swiftly provide all essential information. A bar chart, colour-coded by region, will depict sales data, while a line chart will represent profit, with each line representing a distinct region. This comprehensive visualisation method enables quick insights into sales performance across regions and highlights profitability trends by region.

20. What is the average delivery duration for different regions and ship modes?
    A bar chart would be perfect for this scenario.each bar representing the region and its length representing the avg.delivery duration divided by different ship mode.

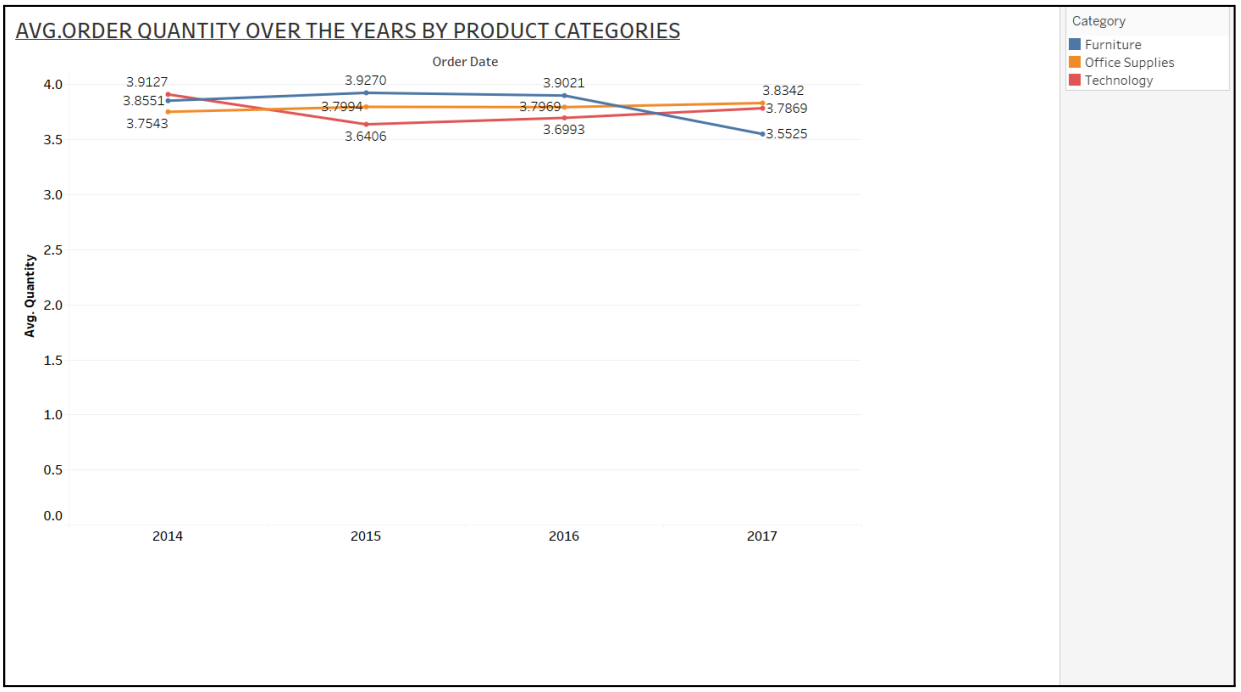


    It's evident that same day shipping has the shortest delivery duration, followed by first class, second class, and standard, which has the longest delivery time. Interestingly, the region doesn't seem to influence delivery duration significantly, as each shipping mode's delivery time remains consistent across all regions.

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

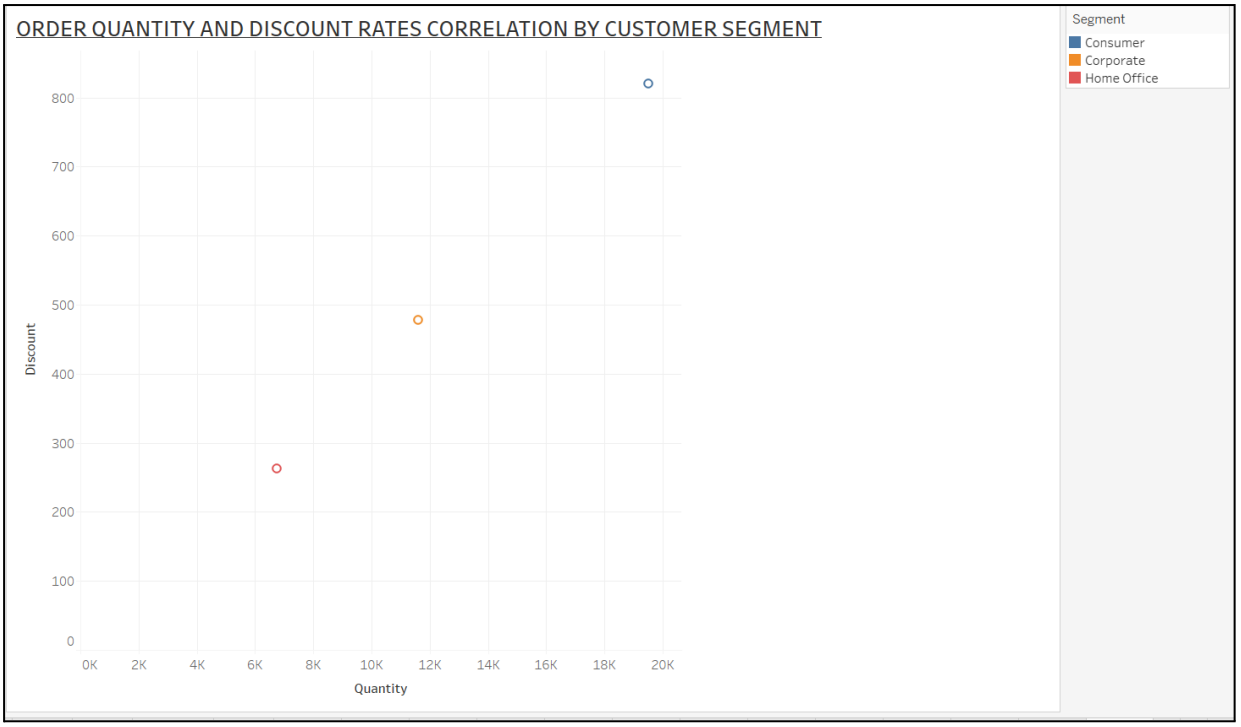
    A line chart is the optimal choice for visualising changes in the average order quantity over the years across various product categories. This type of chart effectively displays trends and fluctuations over time, making it easy to identify any shifts or patterns in ordering behaviour for each product category. By plotting average order quantity on the y-axis and years on the x-axis, the line chart provides a clear and comprehensive view of how ordering patterns have

evolved over time, facilitating insights into changes in customer demand, market trends, and product popularity across different years and categories.



All three categories had almost the same start .technology fell off after after 2014 but then had a consistent growth and in 2017 has the second most order quantity.office supplies remained consistent from the start and in 2017 had the the most number of orders.furniture too had a consistent growth but faced a decline after 2016.

22. Can we visualise the correlation between discount rates and order quantities for different customer segments?
    A scatter plot is the perfect choice for visualising the correlation between discount rates and order quantities for different customer segments. This type of chart effectively displays individual data points, allowing for a clear examination of the relationship between two variables. By plotting discount rates against order quantities for each customer segment, the scatter plot enables the identification of any patterns, trends, or correlations between the two variables. This visual representation is essential for understanding how discount rates impact order quantities across different customer segments and can provide valuable insights for pricing strategies and promotional planning.



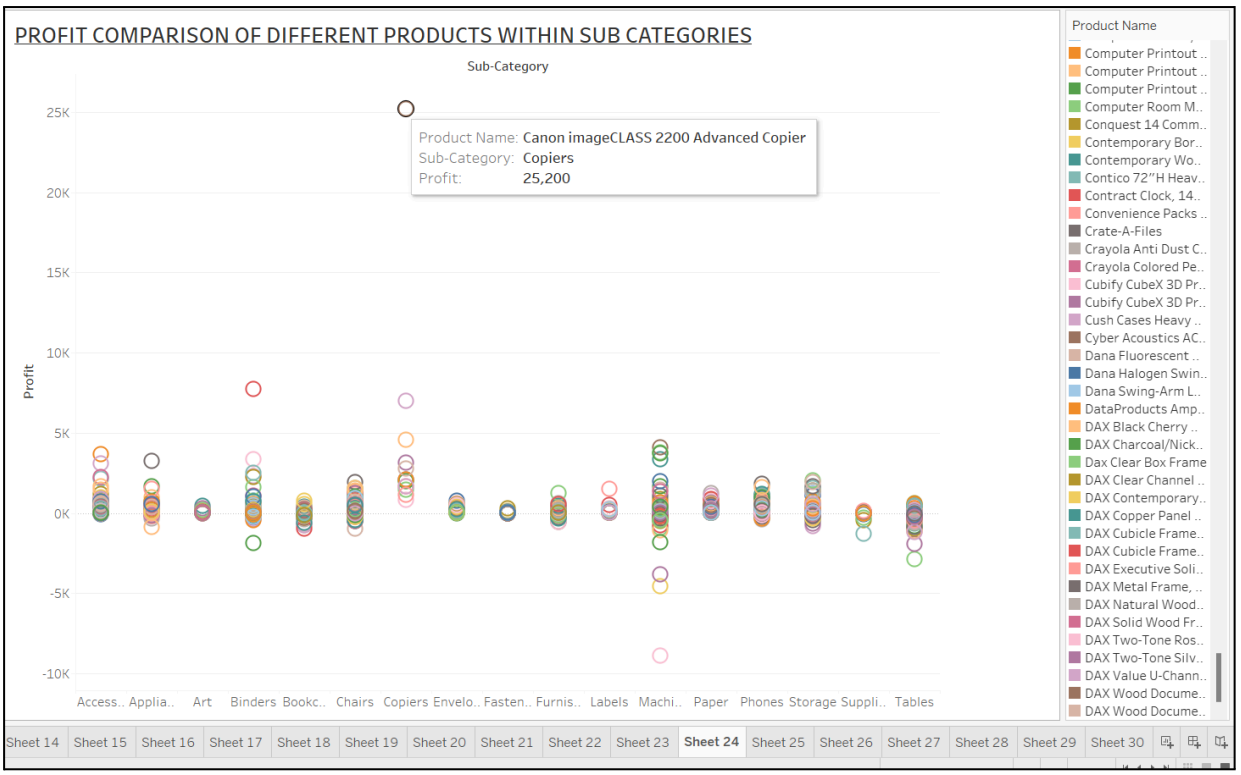
The analysis reveals that the home office segment exhibits the lowest order quantity, accompanied by the least amount of discount rates. Following this, the corporate segment falls in the middle range, while the consumer segment displays both the highest order quantity and the highest discount rates. This observation underscores distinct purchasing behaviours across the different customer segments, with consumers showing a tendency toward larger order quantities, possibly influenced by higher discount incentives. Meanwhile, the home office segment appears to prioritise purchases with fewer discounts, indicating potential preferences for product quality or specific features over price discounts.

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

    Bar chart with each bar representing a region and its length representing the number of returns would be the most effective and simpler visualisation.

The chart clearly shows that the West region has the highest number of returns, significantly outpacing the East, which has the second most returns. The remaining regions have return numbers closely following those of the East. This highlights a notable disparity in return rates between the West and other regions.

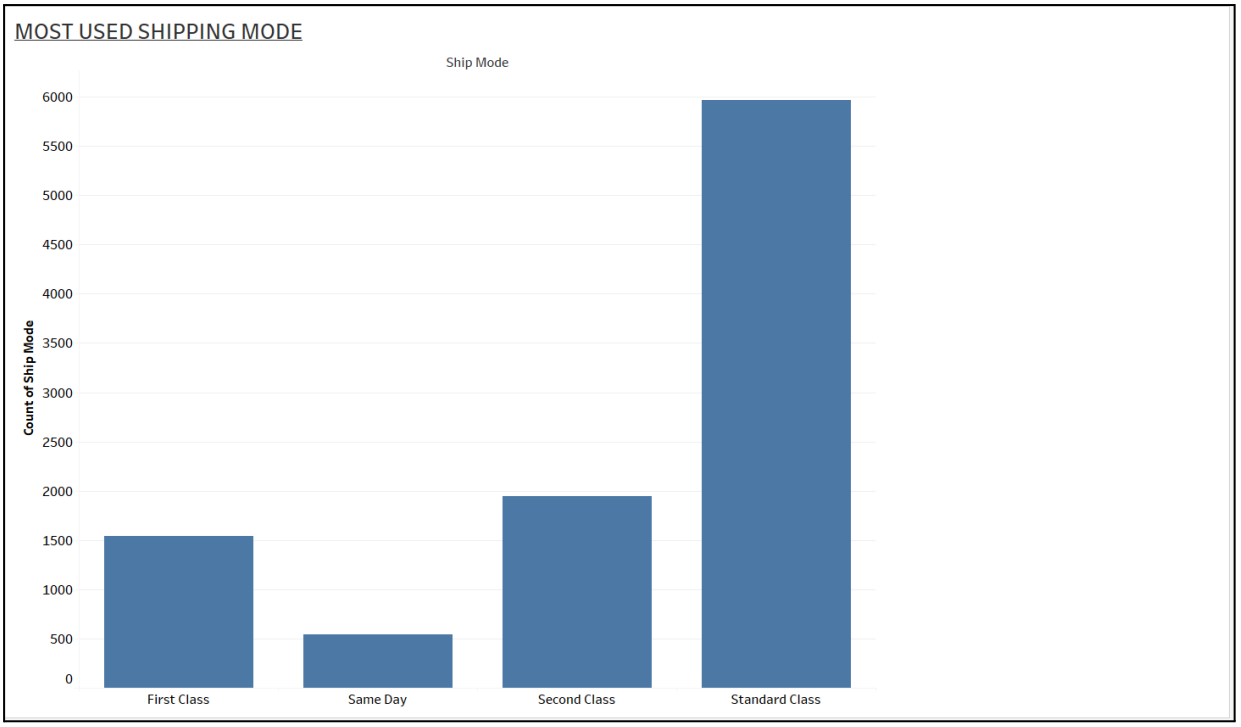
24. Can you compare the profit of different products for different subcategories?
    A scatter plot is an ideal choice for comparing the profit of different products across various subcategories, especially given the large number of products involved. This type of chart allows each product to be represented as an individual data point, making it possible to effectively visualise and analyse the distribution and variations in profit across different subcategories. By plotting profit on one axis and subcategories on the other, the scatter plot provides a clear view of how individual products perform within each subcategory, highlighting trends, outliers, and potential areas for strategic focus.



The scatter plot reveals that copiers generate the highest profit within the subcategory, while machines yield the least profit. The subcategory of binders has the second most profitable product. Additionally, most products in the other subcategories appear to have similar profit margins, indicating a relatively uniform performance across these categories.

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

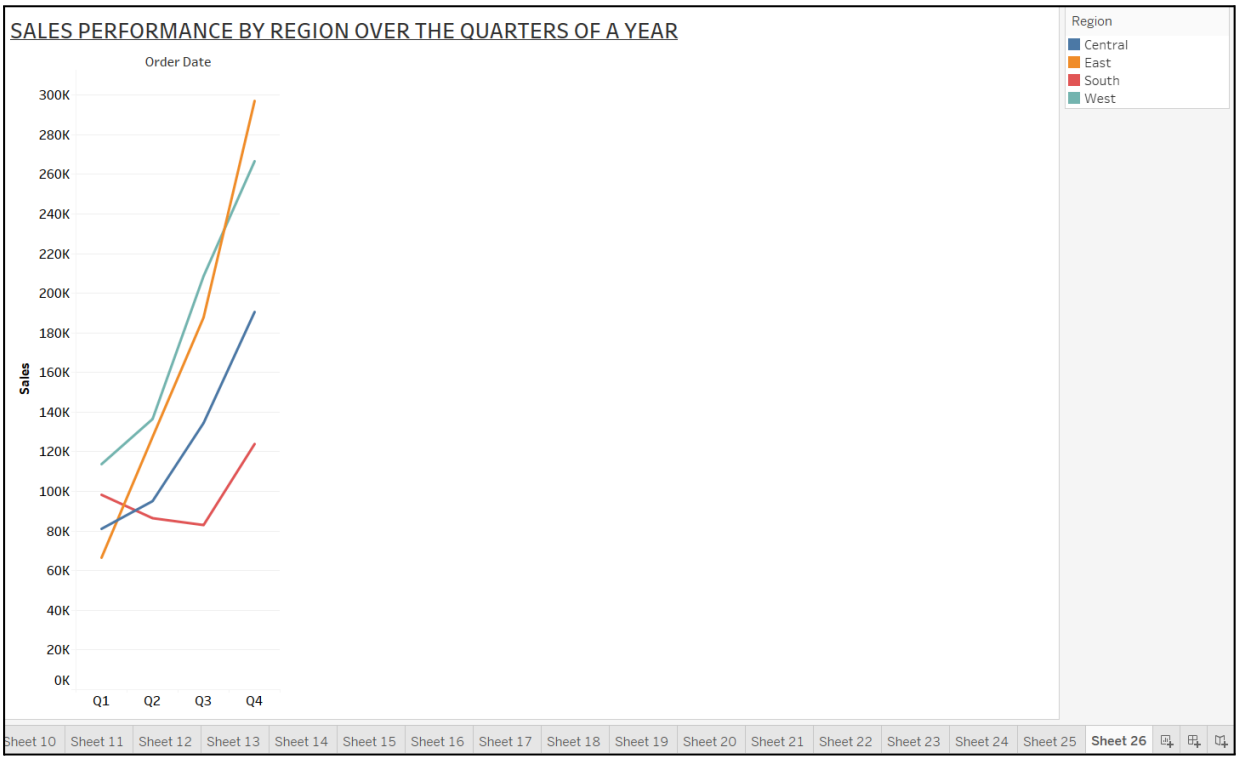
    This scenario definitely calls for a bar chart.each bar represents a ship mode and the length of the bar indicates the count of the ship mode which inturn tells us how many times they were used



The most utilised shipping mode is standard, despite it having the longest shipping duration. This is followed by second class, though not closely, then first class, and finally, same day shipping. Interestingly, the order of preference for shipping modes is the exact opposite of their respective shipping durations. This suggests that factors other than speed, such as cost or reliability, play a significant role in customers' shipping mode choices. It indicates that while faster shipping options are available, customers might prioritise affordability or other factors over delivery speed.

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

    A line chart is the optimal choice for visualising the sales performance of different regions throughout the quarters of a year. This type of chart effectively illustrates changes over time, making it easy to track and compare sales trends across different regions from quarter to quarter. By plotting the sales data for each region on the same chart, with time on the x-axis and sales on the y-axis, a line chart allows for a clear visual representation of the evolution of sales performance. This helps in identifying patterns, seasonal variations, and any significant fluctuations in sales across different regions, providing valuable insights for strategic planning and decision-making.

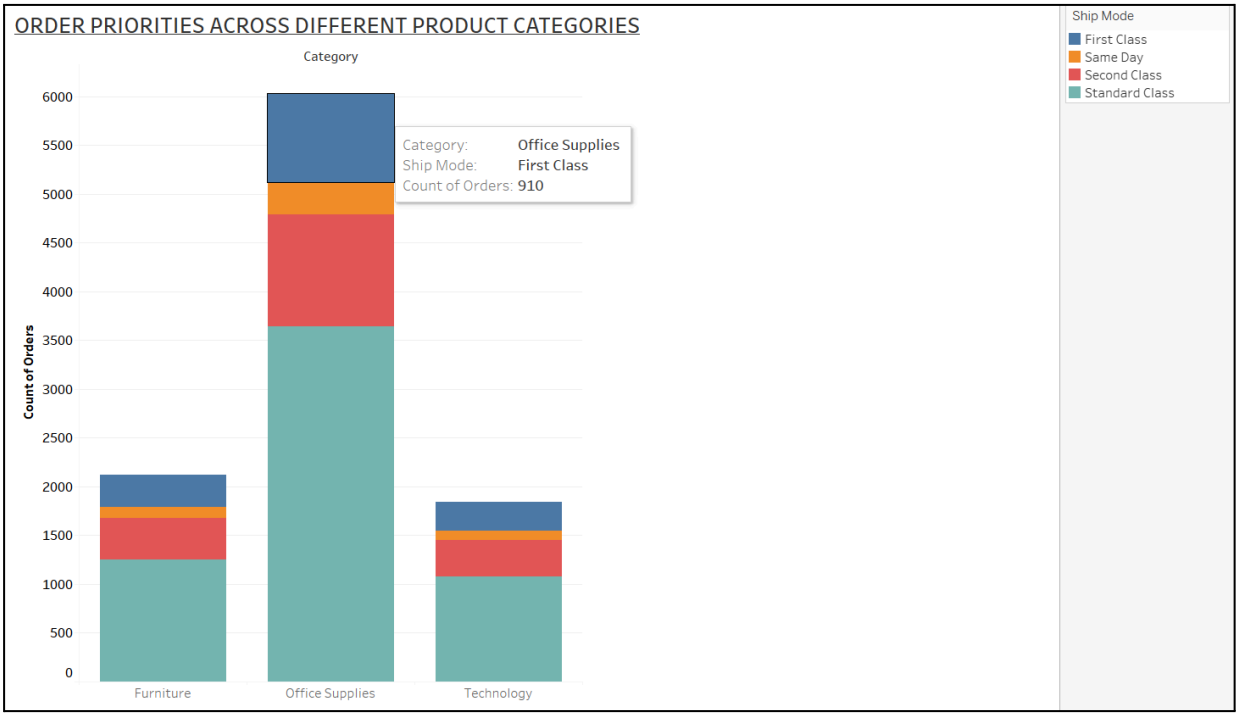


The sales in all regions show an upward trend, although at different rates. The East region started slowly but then accelerated, finishing the year with the highest sales. The West followed a consistent upward trajectory, ending with strong sales. The Central region showed moderate growth, performing adequately but not as impressively as the East or West. Meanwhile, the South had the least amount of sales, indicating slower growth compared to the other regions. This analysis highlights the East's significant late-year surge, the steady performance of the West, the Central region's steady but less remarkable growth, and the South's lagging sales, suggesting potential areas for targeted marketing and sales strategies.

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

A stacked bar chart is an excellent choice for visualising the distribution of order priorities across different product categories. This type of chart allows for the comparison of multiple categories simultaneously while also displaying the composition of each category in terms of order priorities.For instance, a stacked bar chart can show which product categories have a higher proportion of high-priority orders and which categories are more commonly associated with

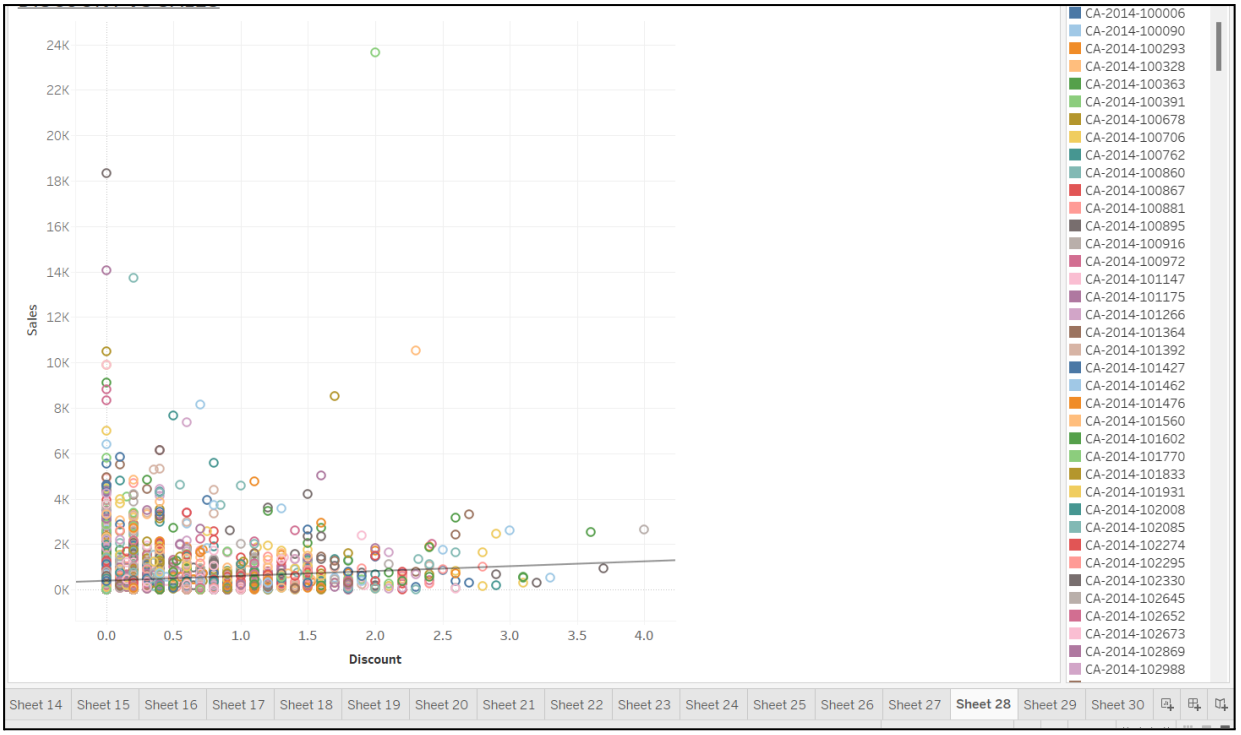
standard or low-priority orders. This insight is valuable for managing inventory, optimising supply chain processes, and tailoring customer service strategies to meet the demand dynamics of each category.



The preference for standard shipping across all categories indicates that, generally, customers are willing to wait longer for their orders if it means saving on shipping costs. However, the notable use of faster shipping modes for office supplies highlights that delivery speed is still a priority for certain categories, reflecting their urgent or frequent usage. This insight can help businesses optimise their shipping options and inventory management to better meet customer expectations and demands.

28. What is the relationship between discounts and sales?

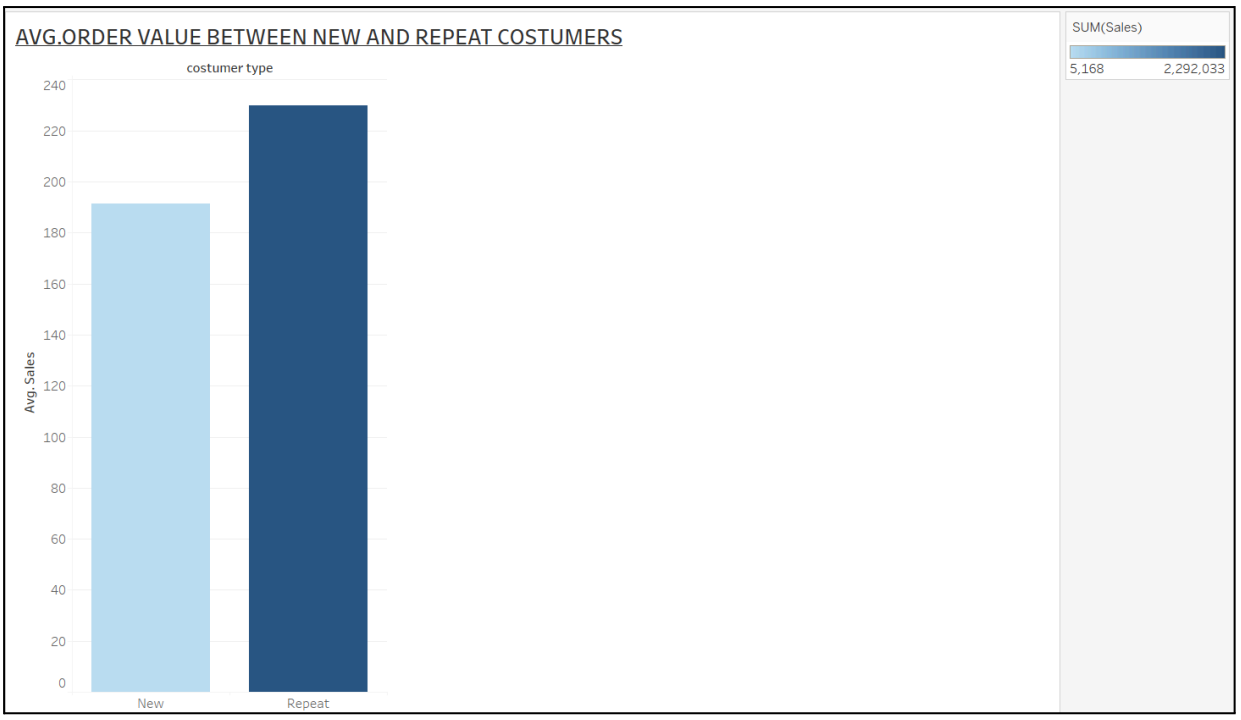
To answer this question a scatter plot with a trend line will be a powerful tool to understand the relationship between discounts and sales, providing both a visual and quantitative assessment that can inform pricing and marketing strategies.



An upward-sloping trend line suggests that higher discounts lead to higher sales. This indicates that customers are more likely to purchase products when offered larger discounts, reflecting a high level of price sensitivity among consumers.This suggests that discounts have a significant impact on sales performance and can be an effective tool for driving revenue growth.

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

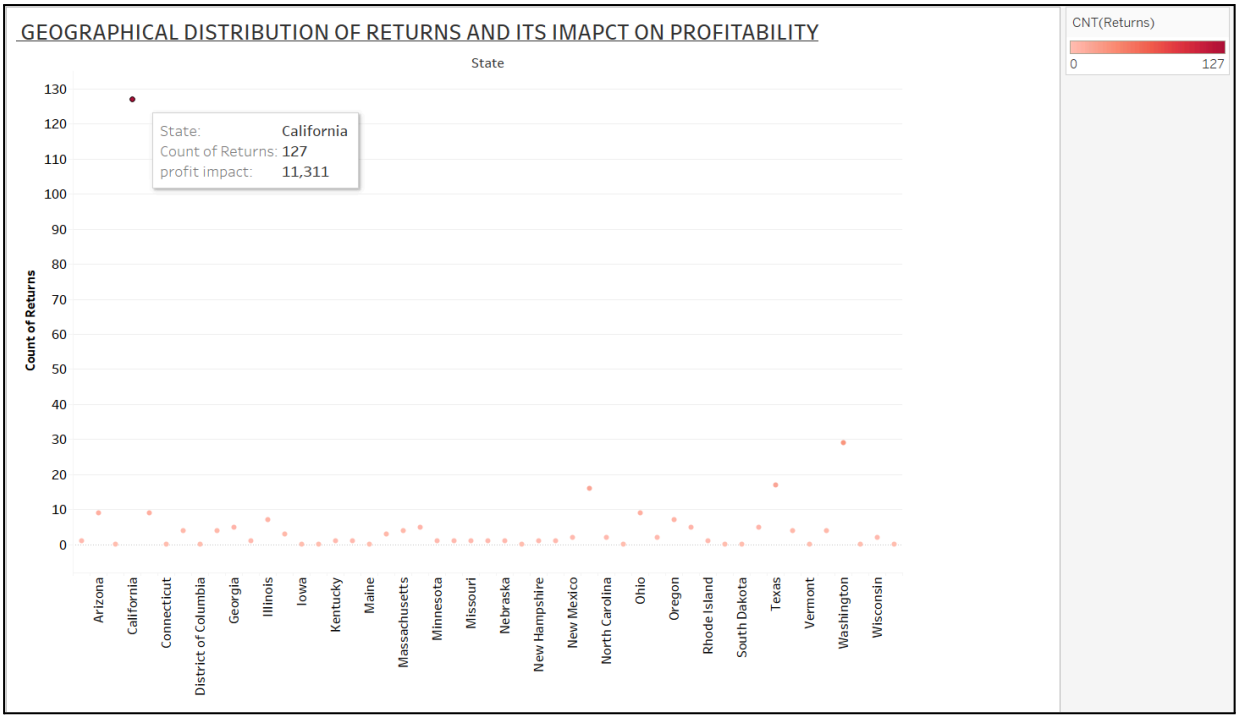
    All we need is a simple bar chart to properly visualise this scenario.a bar each for new and repeat customers where the length shows how much sales they bring in.



It's evident that repeat customers contribute more to sales compared to new customers, which aligns with expectations. This observation underscores the importance of customer retention and loyalty in driving overall sales revenue. By nurturing relationships with existing customers and encouraging repeat purchases, businesses can capitalise on the higher lifetime value of loyal customers. Additionally, this finding emphasises the significance of implementing strategies to cultivate customer loyalty, such as personalised marketing campaigns, loyalty programs, and excellent customer service, to sustain long-term business growth and profitability.

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

    While a map chart may not be suitable due to the limited geographic scope of the data, a scatter plot offers a versatile and informative alternative for visualising the geographical distribution of returns and its impact on overall profitability.



The chart illustrates a significant disparity between California and other states concerning the number of returns and their respective impact on profitability. California stands out with a notably higher number of returns, coupled with a substantial impact on profitability. In contrast, other states exhibit relatively consistent numbers of returns, suggesting a more uniform distribution across those regions. This observation underscores the unique market dynamics and customer behaviours in California, indicating potential opportunities for targeted strategies aimed at addressing return rates and maximising profitability in this region.

Insights from Sales, Profitability, and Customer Behavior Analysis

Monthly Sales: Sales start slow in January, increase through the year, dip significantly in October.

Sales by Product Category: Technology leads in sales.

Customer Performance: Line charts with search functionality help track individual customer performance.

Sales by Day and Category: Heatmaps show Wednesday as the slowest day; Monday and Friday have the highest sales.

Product Category Growth: Technology leads in growth; office supplies recovered to surpass furniture.

Regional Sales: West leads in sales, followed by East, Central, and South.

Profit by Shipping Mode and Segment: Same day shipping is most profitable for consumers and home office; standard class leads for corporate.

Order Processing Time: Consistent across product categories.

Discounts and Profit: Higher profits are generally associated with lower discounts.

Quarterly Sales by Region: East shows the most growth, followed by West, Central, and South.

Order Priorities: Standard shipping is most used; first and second class preferred for office supplies.

Discounts and Sales: Higher discounts lead to increased sales.

Customer Segment Performance: Consumer segment has highest sales and discounts, followed by corporate and home office.

Sales and Profit Trends: Technology leads in sales and profit; dual-axis charts show trends by subcategory and region.

Repeat vs. New Customers: Repeat customers generate more sales.

Geographical Returns: California has the highest returns and profit impact; other states have lower, similar return rates.