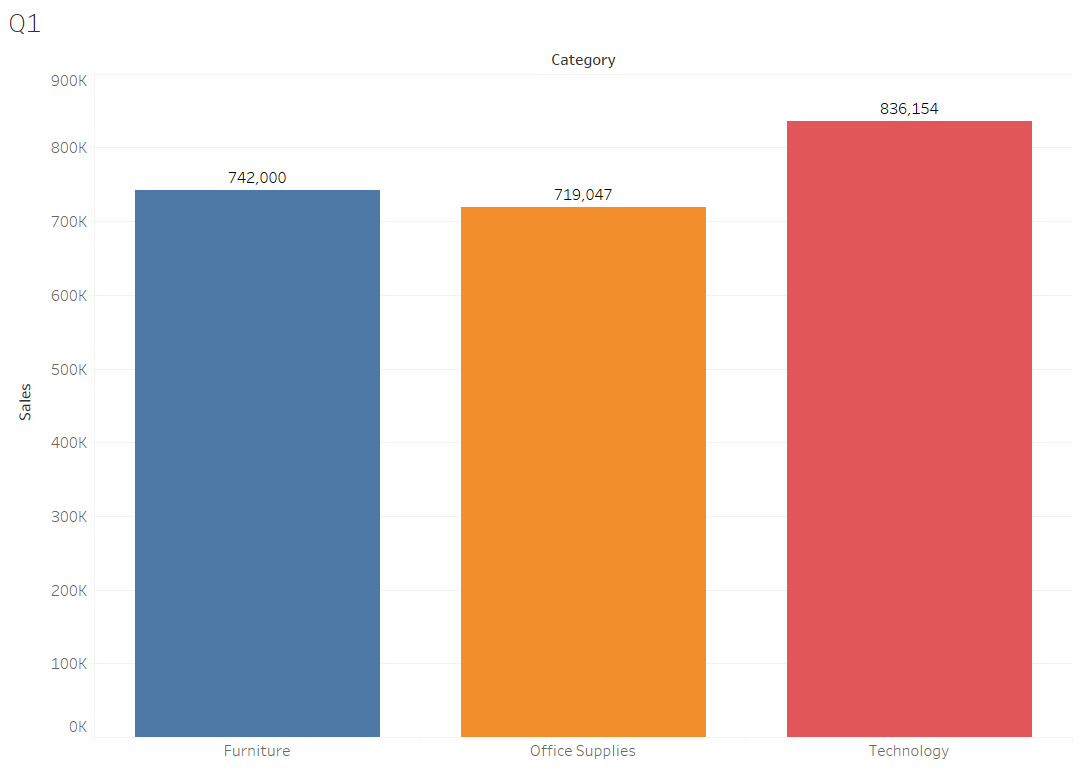
# Questions and Answers

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

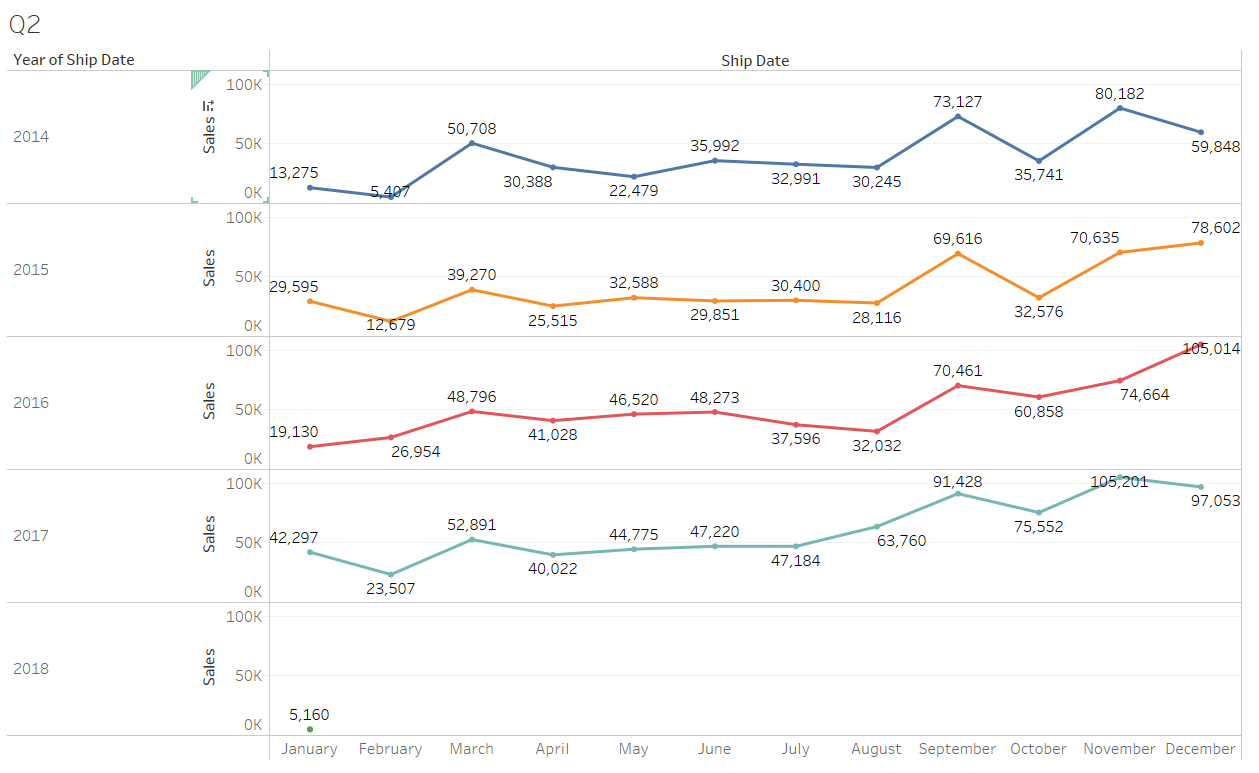
The question can be effectively answered using the following datasheet developed in the tableau environment. It shows the total number of sales as per the product categories. In this context, the choice of the graph is appropriate as per my understanding.



The choice of the graph is appropriate as it shows the required output from the described data and it provides values of sales for different categories. It shows that the Superstore dataset has 742000 sales in total for the Furniture category, a total of 719047 sales in the office supply category, and 836154 sales in total with the technology category. It also reveals the category with the highest number of sales which helps determine some key insights about the sales and products of the data.

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

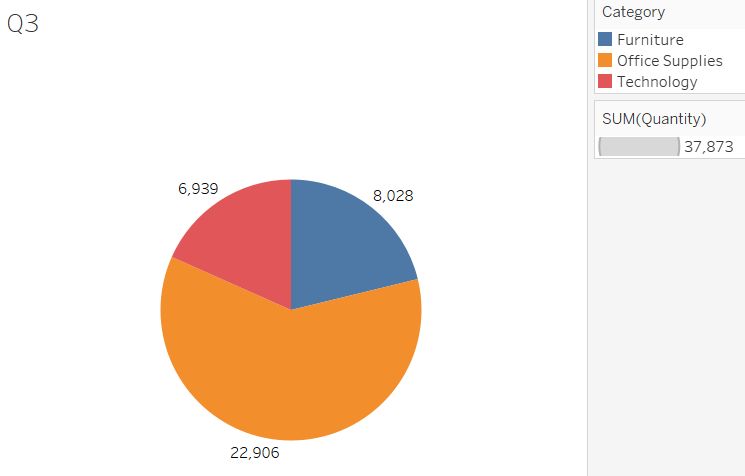
In this case, the sales amount change over a year is shown. This also shows the changes over months during different years. The choice of the graph is justified as it shows trends in sales over months in different years.



The graphical representation shows the ups and downs of the sales over different months. These show that February has high sales and September has the highest sales every year. This also emphasises the similarities of trends over different years. This helps the management of stores to take necessary actions in their operations.

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

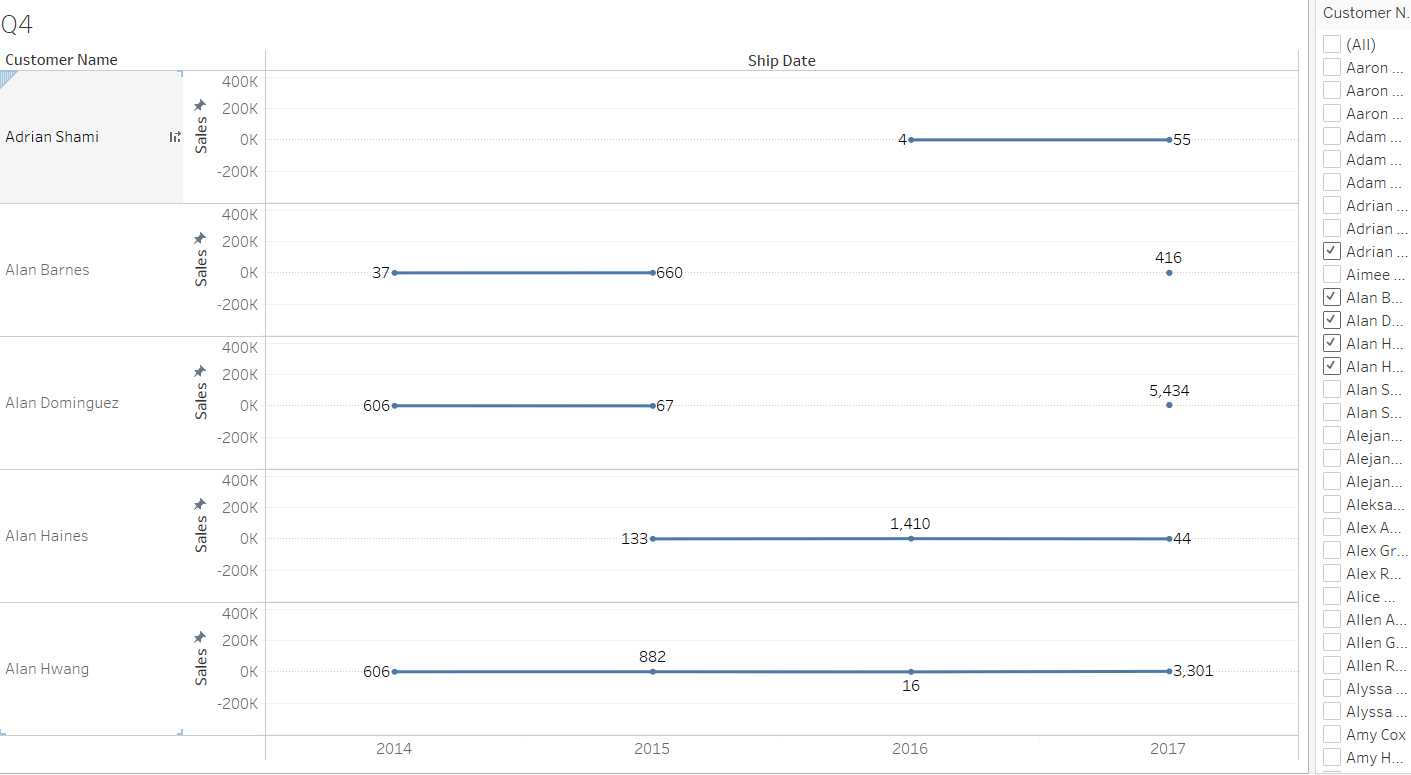
This pie chart shows the sales amount distribution of different product categories and it also shows that Office Supplies has the most sales. It can also be subject to deliver the idea about the amount of total sales.



The choice of the graph can be justified in assessing the volume of sales and the pie chart helps get a visual idea about the categories which have the majority of sales.

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

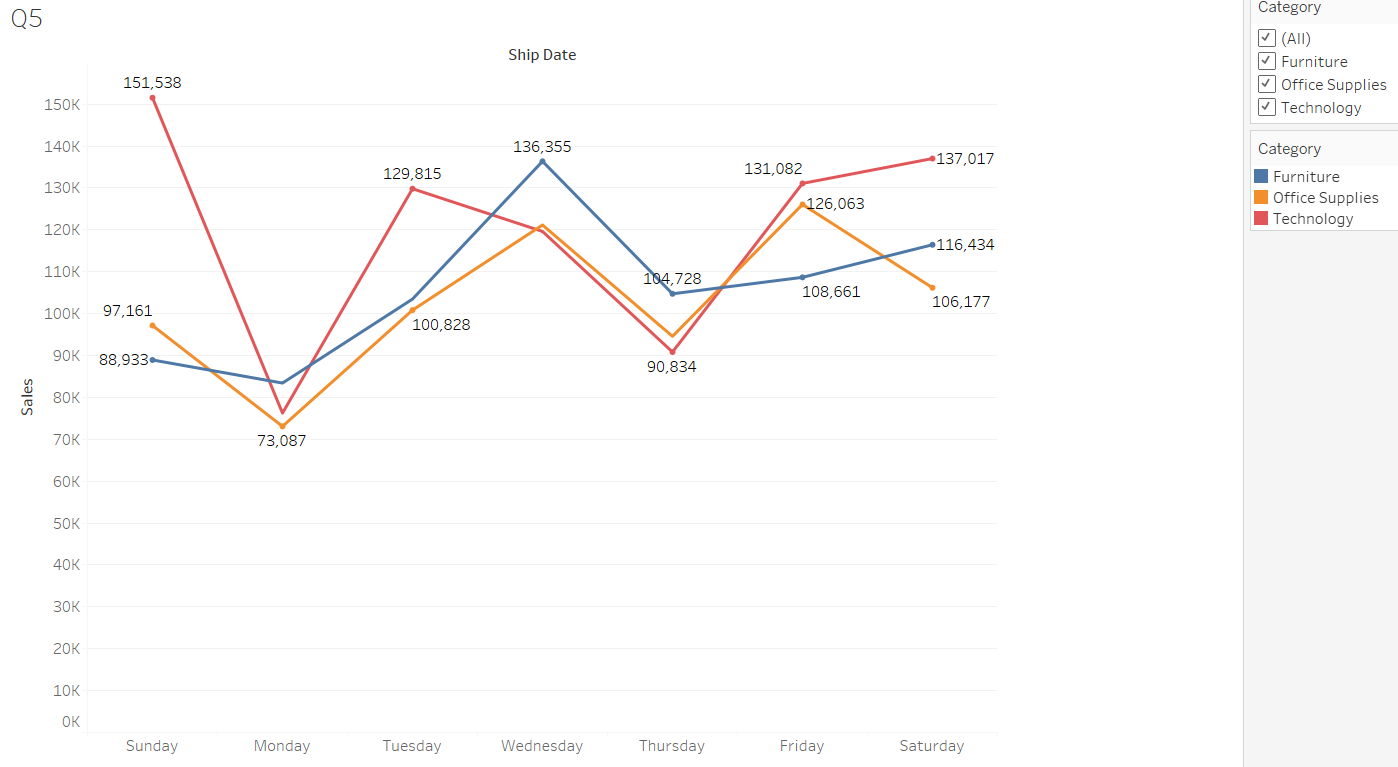
Yes, the chart shows the sales performance of individual customers over time and it implies different data from the choices of the customer. The choice of the chart and filter is justified as the sales performance over the years can be effectively observed.



The chart shows trends in sales over time and it enables the possible outcomes for the customers and their respective sales. The choice of the customers can also be explored through the filters and one or more customers’ respective sales can be observed.

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

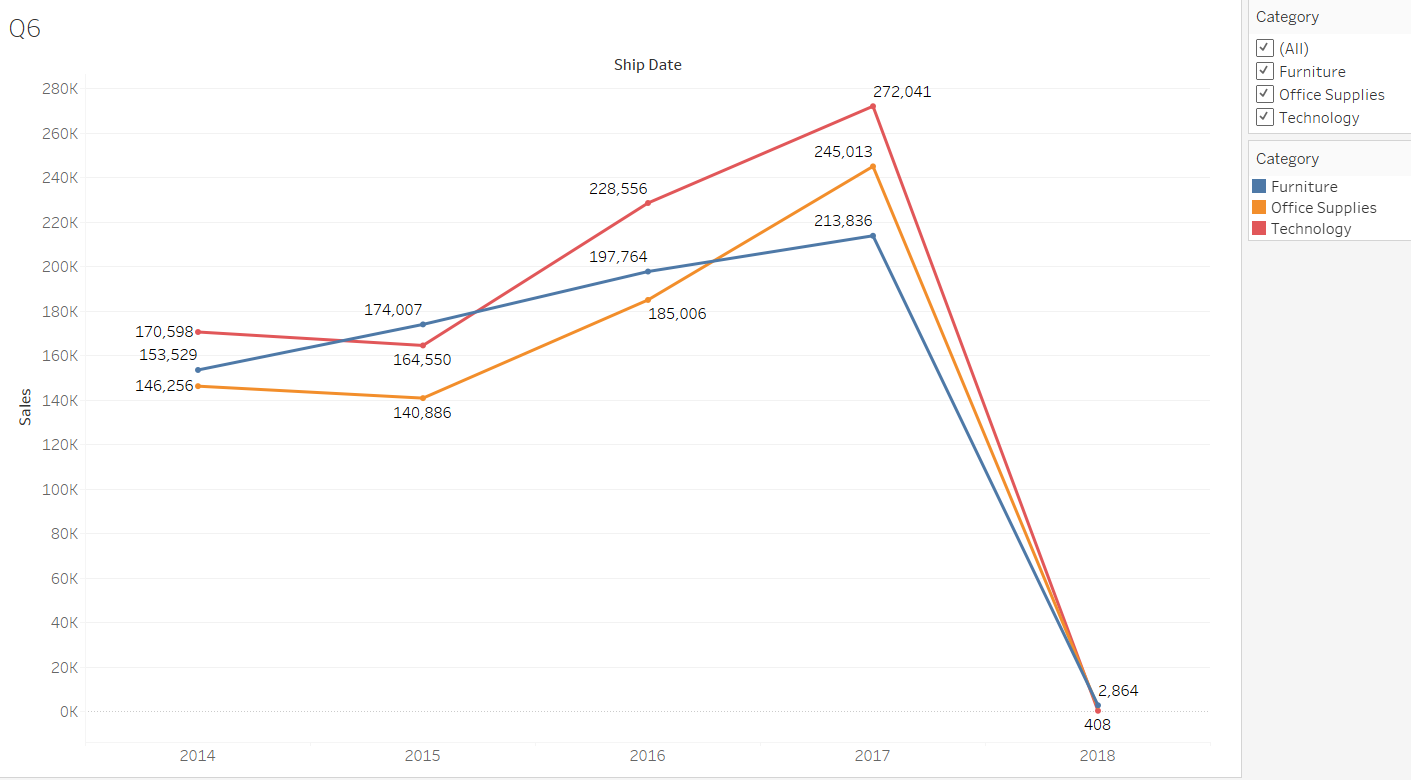
The different product categories are highlighted in different colours and the sales are represented graphically according to different days of the week. It can be subject to developing ideas on how the sales vary over the different days of the week.



The choice of the graph can be justified as the line plot effectively shows trends over a period of time, days, and weeks. This helps gain an idea of the different categories based on the different filters and colour codes.

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

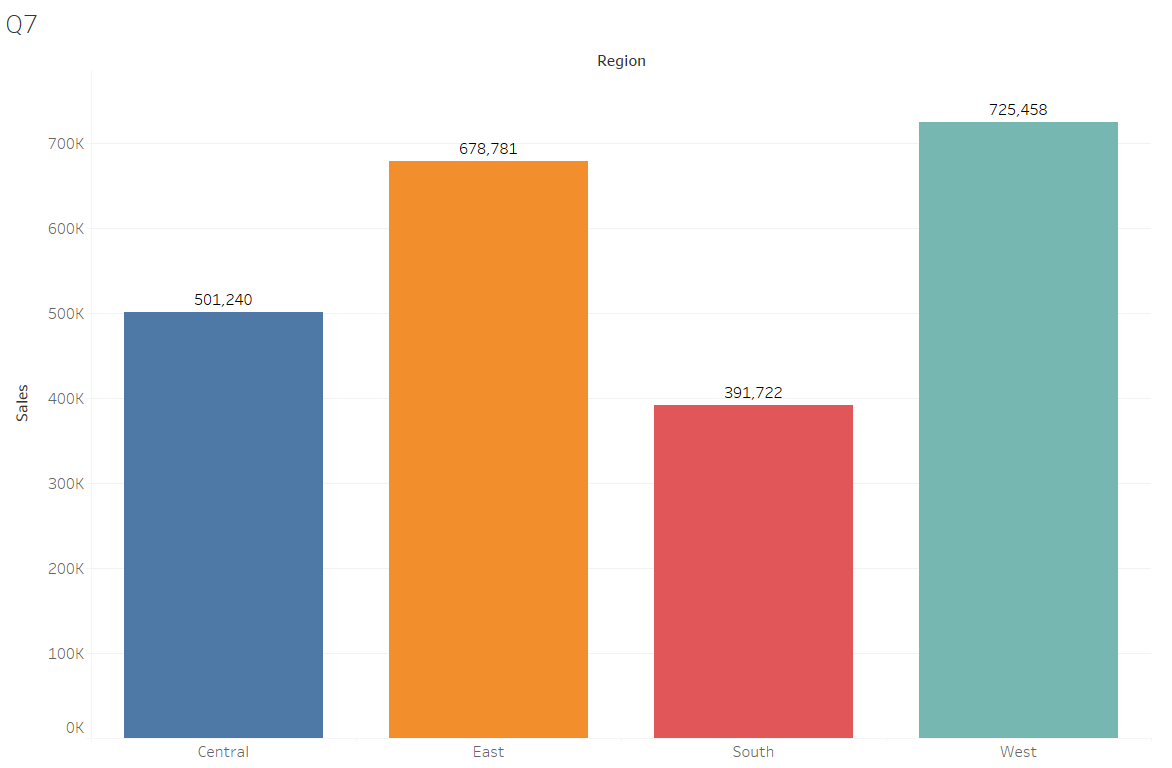
Sales growth of different product categories can be observed in this graph which shows that the sales have grown from 2014 to 2017. This also shows lowered sales growth from 2017 to 2018.



The sales growth of different products are highlighted in three categories and it is subject to show the required output. The graphical representation helps identify the sales growth in different years and the sales trends over the years have been effectively shown.

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

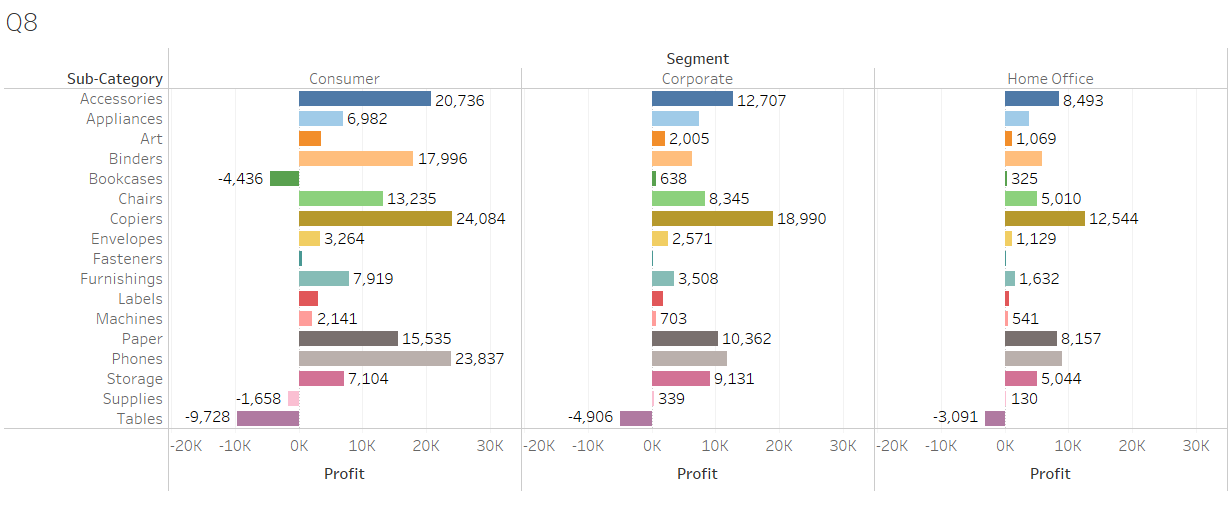
The sales distribution varies across different regions in the Superstore dataset and the following graphical representation shows the desired output. These four regions, central, east, west, and south are observed in this graph.



The graph shows how these sales distributions are in different regions and the choice of the graph is justified in terms of showing the distribution with levels and labels. It helps compare the different distributions of sales in the regions.

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

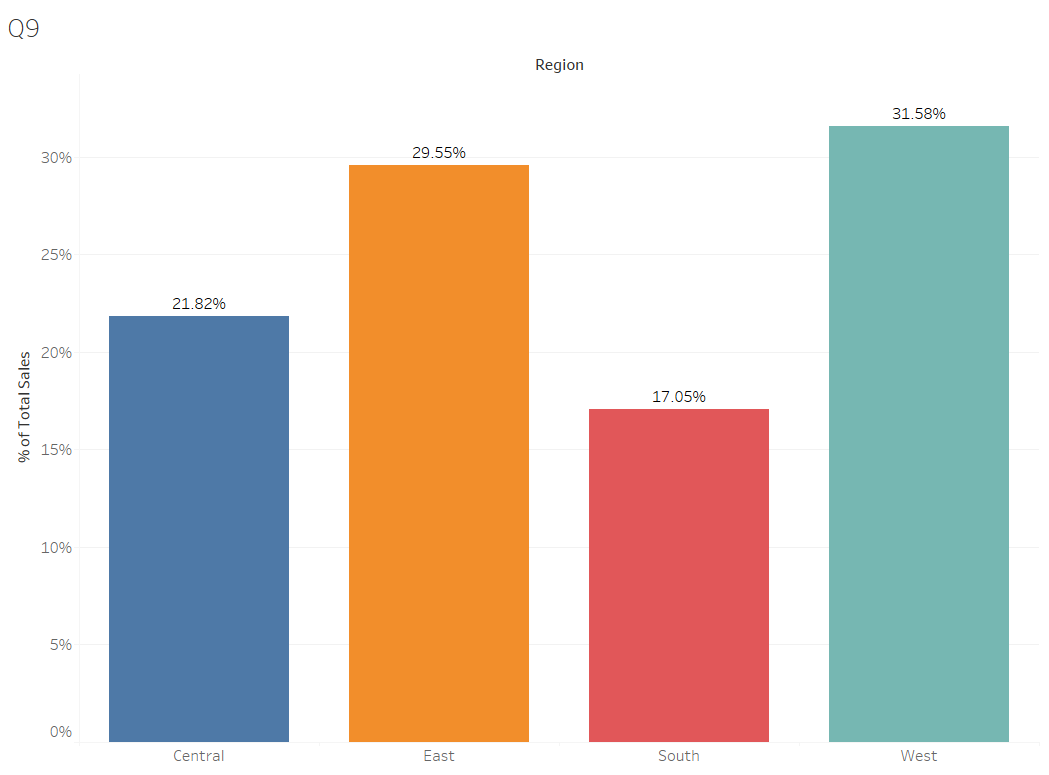
Yes, the profit and loss aspects of different subcategories of the products are observed in this graph and it also divides the output according to different customer segments. This is subject to delivering the idea of profit and loss statements of sales.



The choice of the graph is also justified as it provides the overall idea of profit and loss according to both subcategory and customer segmentation.

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

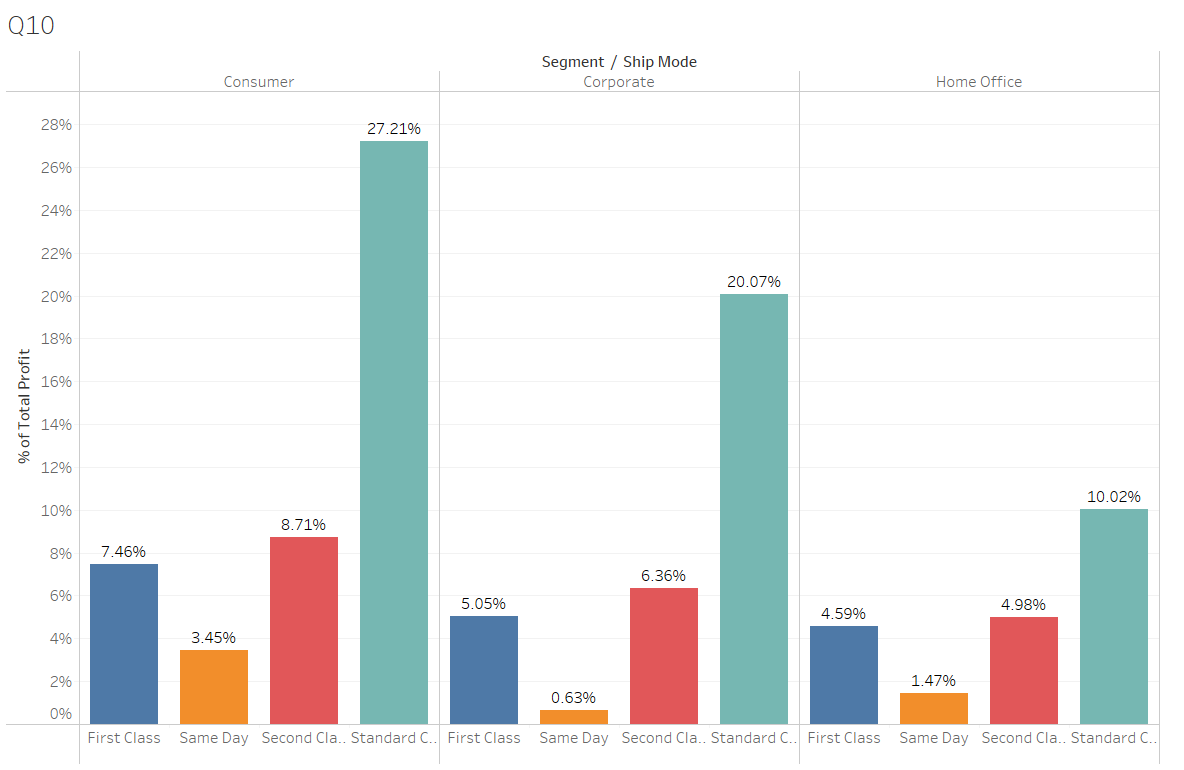
Region-based percentage contributions with overall sales have been represented in the graph and it is subject to deliver the differentiation of percentage of overall sales in different regions of the superstore. This helped assess the required output and the choice of graph is also justified.



The choice of the graph is appropriate as we can observe the differentiation of the sales percentage in different regions which is the required output. We can see that the West side has the highest sales percentage and the South side has the lowest sales percentage.

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

Yes, visualising the profit margins according to the shipping modes and customer segments is possible and it is an essential aspect of data analysis. The bar plot is essential, in this context, to observe the percentage of total profit for different customer segments and shipping modes.

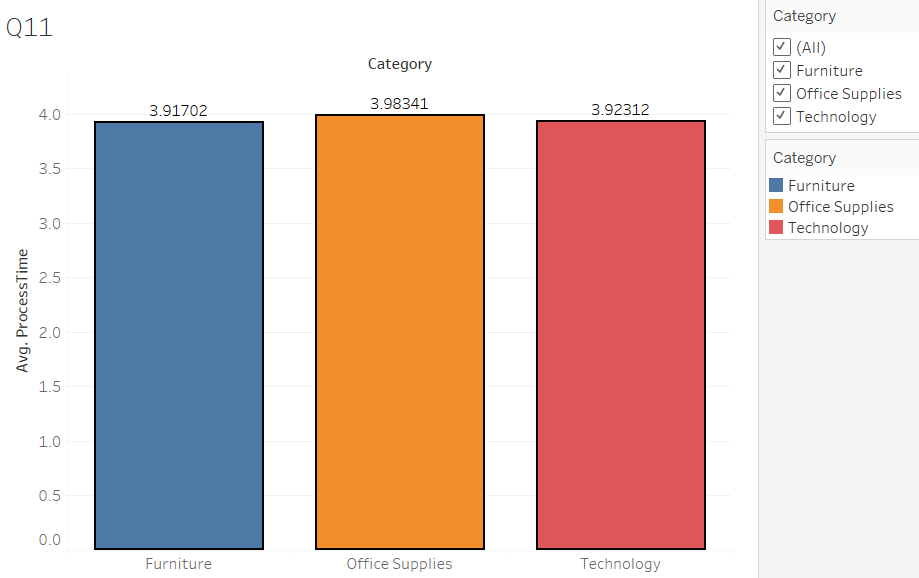


The choice of the graph is justified as the level of profit in different segments and shipping modes can be compared. This helps in data identification and analysis to a further extent and it delivers the idea that the standard class has the highest profit in every segment and shipping mode.

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

The processing time for supplying the orders can be derived from the difference between the Ship Date and the Order Date. This can be effectively done with the help of a calculated function in Tableau. The DATETRUNC function has been used and the hour value is calculated.

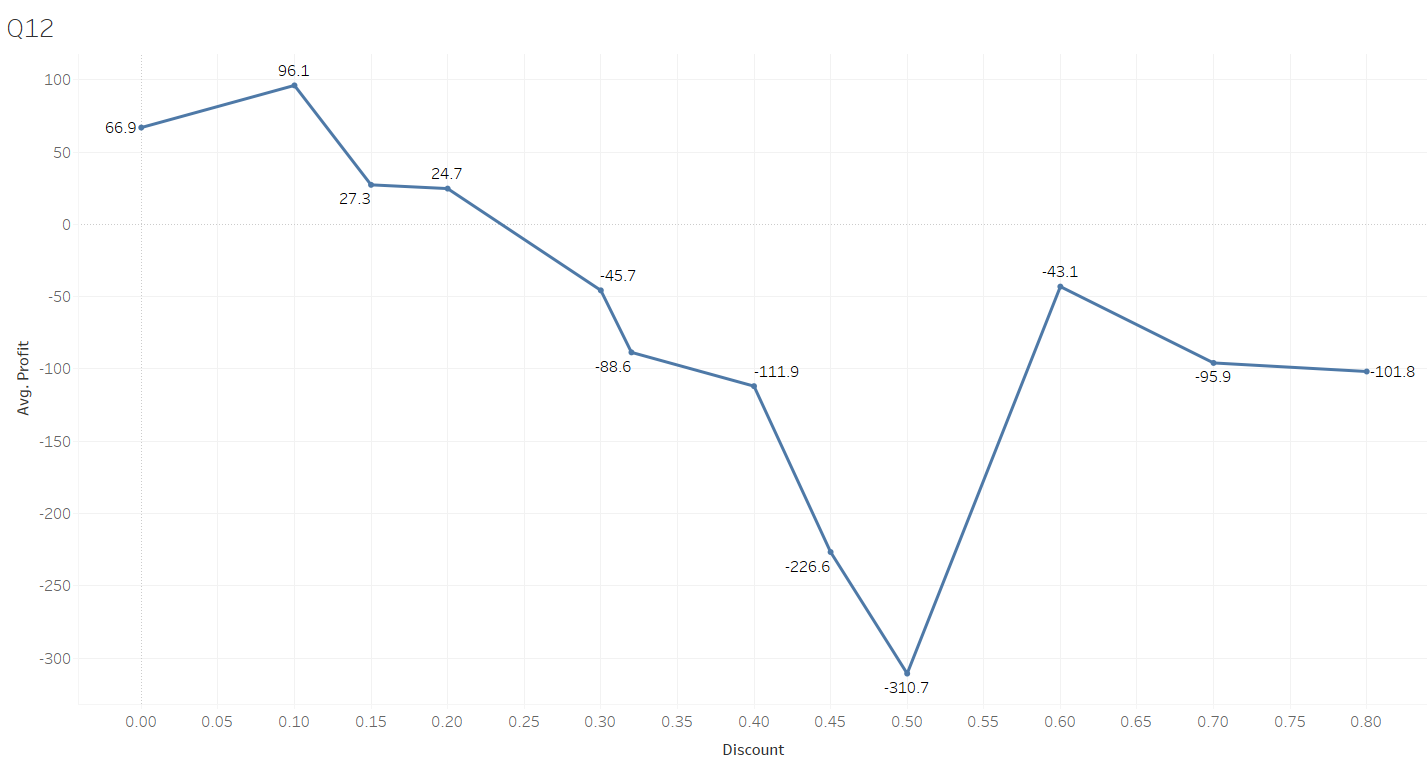
The formula is: ***“DATETRUNC('hour',[Ship Date])-DATETRUNC('hour',[Order Date])”***



The choice of the graph and the function is justified and it can help differentiate the process time of the orders. This provides an idea of how fast the different product categories are processed and this helps provide valuable insights from the graph.

## How do discounts affect overall profit?

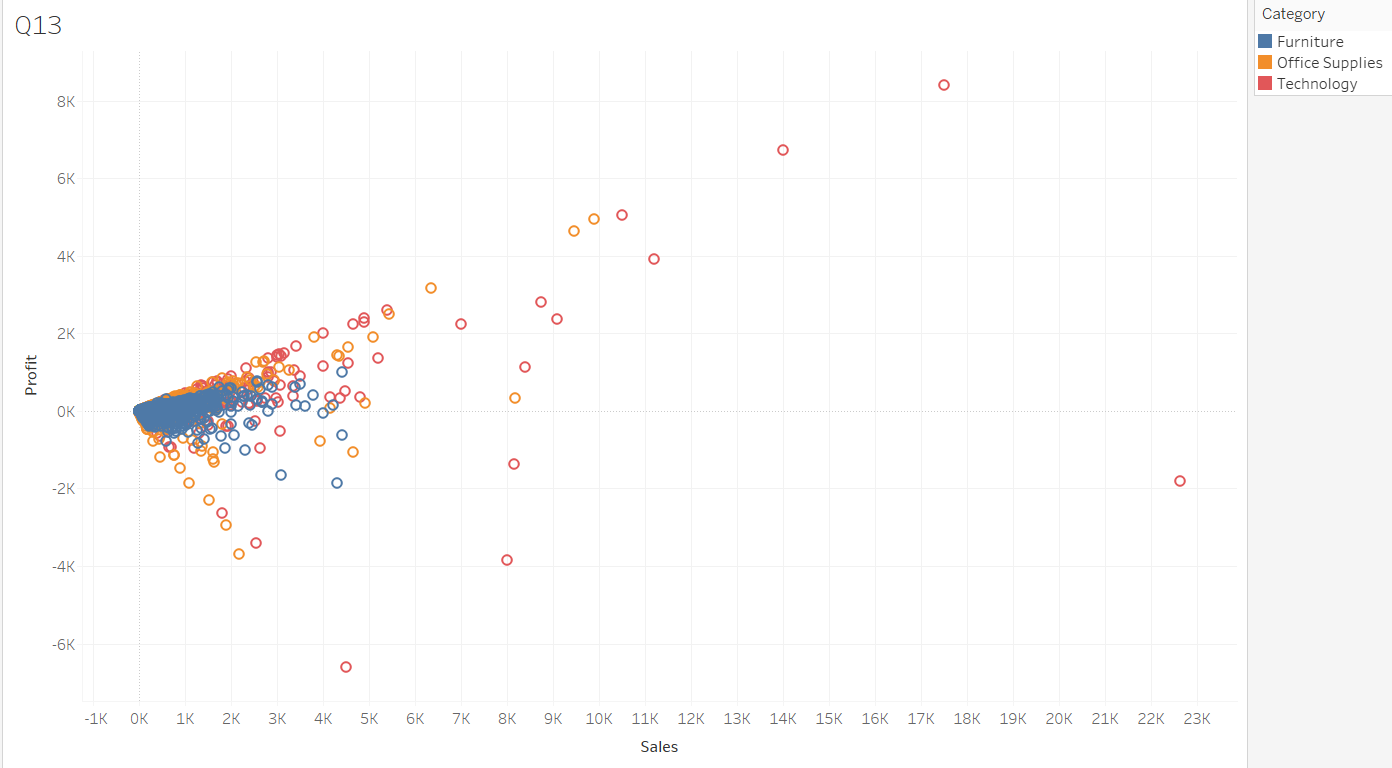
The following line chart provides an idea of how the discounts affect the overall profit from the sales. The sales data has the discount and profit calculated from the tableau and this helps determine how providing discounts can increase or decrease the overall profit.



The choice of the graph is justified as the representation shows that the change in discount helps develop profit. It can be observed that with a minimum value of discounts, the profit is high. The 40-60% discount shows the lowest discount and again the high discount provides a moderate range of profit. The line chart is effective in providing the trends over time or other variables which are effectively implied.

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

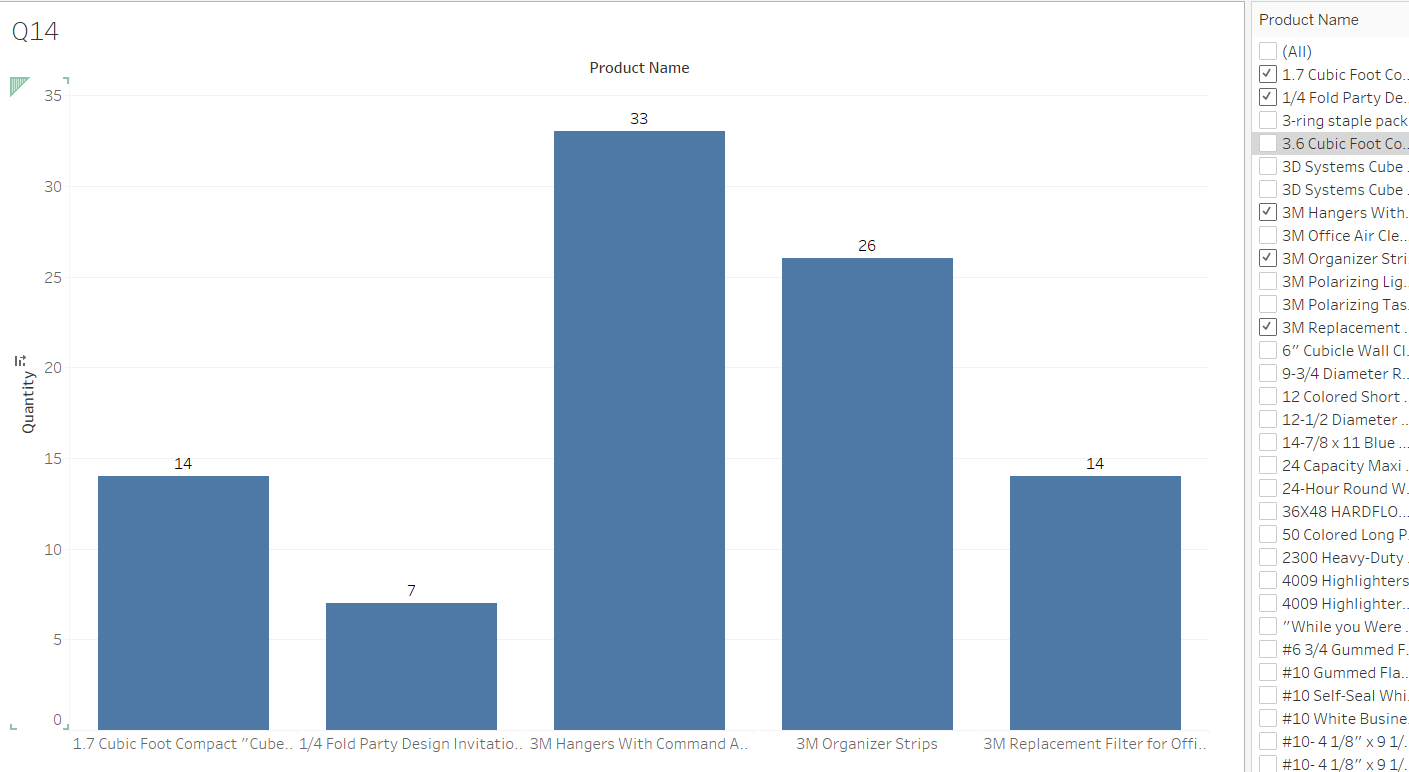
Yes, the relationship between product sales and profitability can be observed with the help of a line chart but the trends developed with the values do not show any fruitful results. Hence, the right choice is a scatter plot and this effectively shows the trends in different categories.



The observation shows that between the 0 to 4K value for sales, the profit is more in numbers as the sales are more. In this context, higher sales show a higher profit from fewer observations. Some of the high sales show low profits which can be outliers and should be ignored.

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

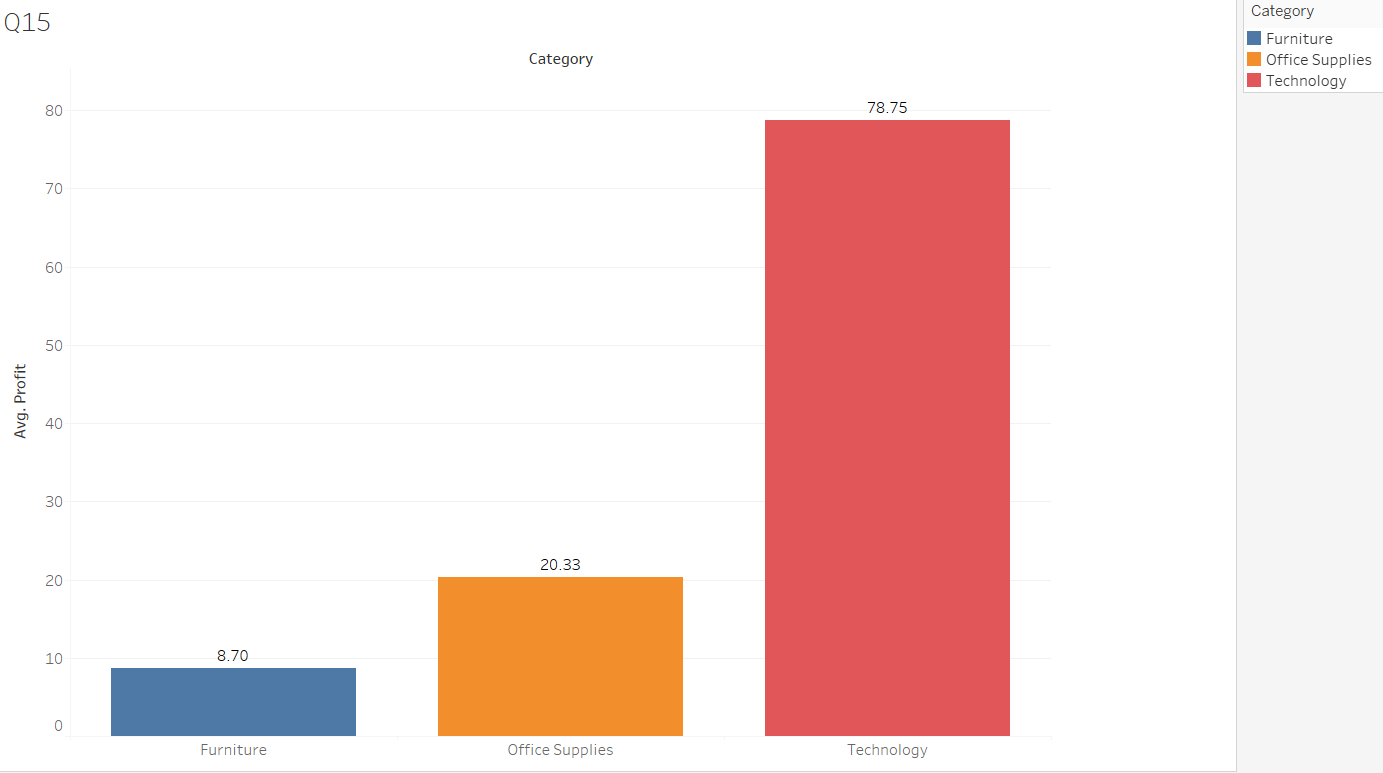
The order quantities for different products can be best visualised with the help of the bar plot and a filter is selected for the product list. This helps the users identify the specific products and their respective quantities in the dataset.



The choice of the graph is justified as it provides a comparative view of the data and it also provides the facility to choose specific products for comparison. This also helps the users to select the visualisations which are quite customised.

## How do the profit distributions vary across different product categories?

The average profit distributions are observed in the graphical representation and it shows with the bar plot that the Technology-based products acquire the maximum profit.



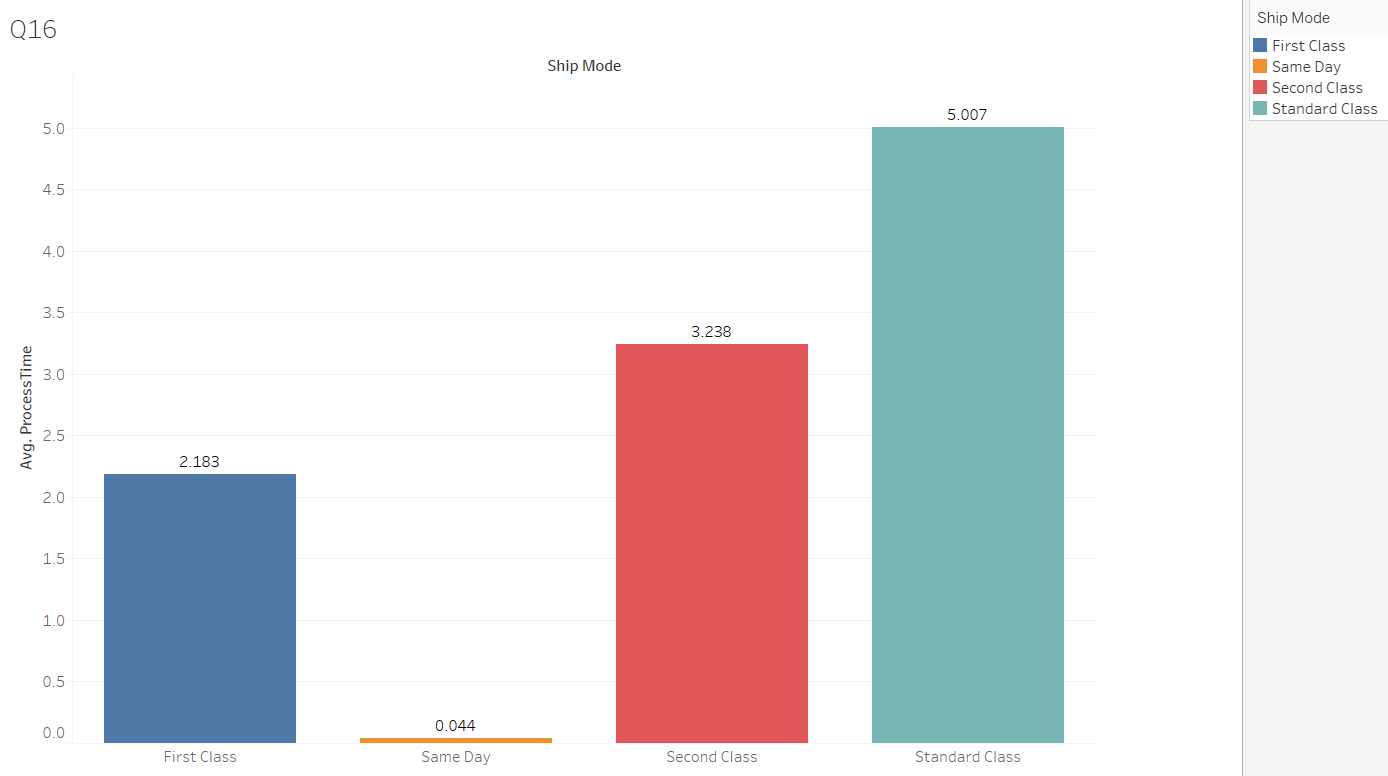
This helps identify how the average profit is distributed over the profit range. The choice of the graph is justified in terms of providing the required view. The bar plot helps assess the comparison of the categories.

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

Yes, it is possible to compare the shipping time or process time with the DATETRUNC function as calculated previously. The formula is:

***“DATETRUNC('day',[Ship Date])-DATETRUNC('day',[Order Date])”***

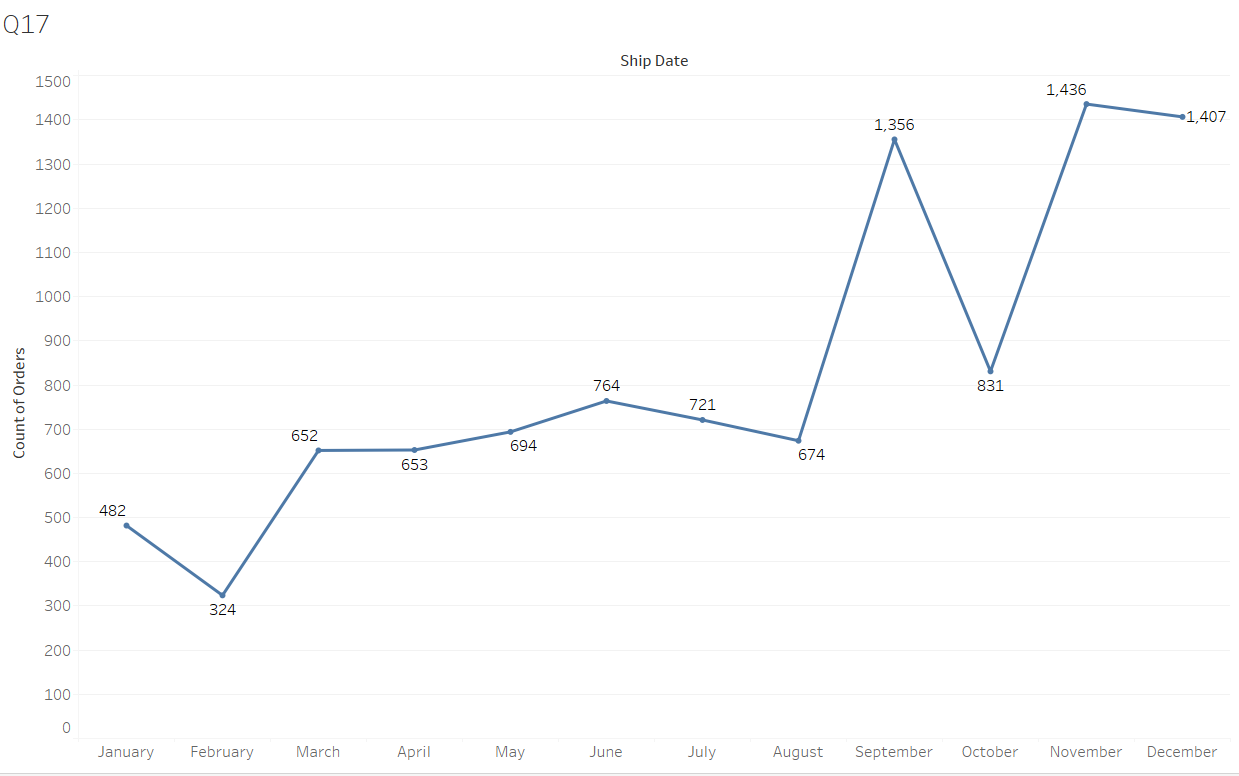
The following visualisation shows the processing time for different shipping modes and it can be observed that the standard class has the highest process time compared to the first and second classes and the same-day delivery has the lowest time.



The choice of the graph can be subjected to understanding the difference between the different shipping modes and the choice of the DATETRUNC function is also justified as it provides the time for processing. The comparative analysis can be effectively observed and it provides the required output.

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

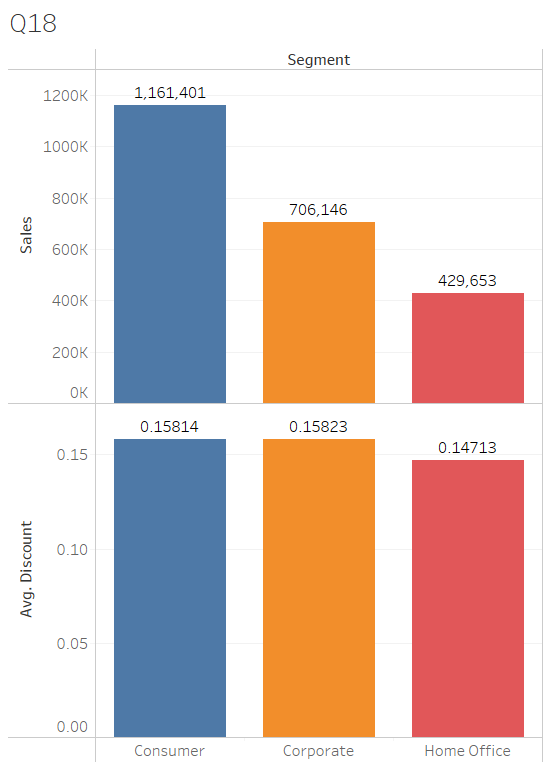
The month-wise trend of the orders shipped or the number of orders is visible clearly in the line plot and it provides an idea of how the orders are distributed over the years in a month-wise view.



This assesses some insight into the data where the trend shows that the number of orders is higher during September, November, and December. The count of orders is quite low in other months and the lowest in February. These can be effective for the business analyst and they can be able to predict the sales in following years.

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

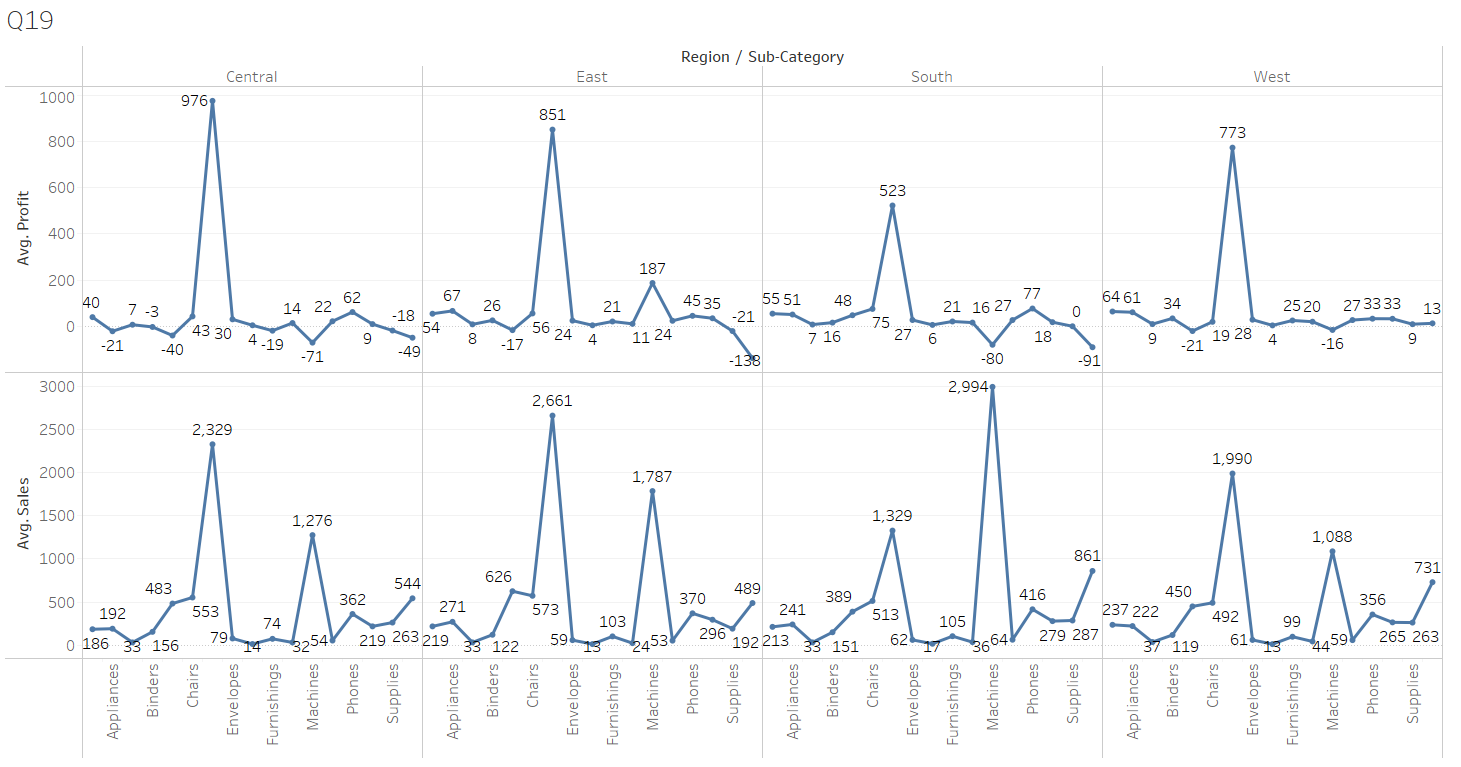
Sales and Discount rates can be observed with the Tableau visualisation and different customer segments. The different customer segments are consumers, corporate, and home office and the data shows that the sales and discount rates are more in the case of the customers.



The visualisation is applicable in the following context and it can be also observed that the different customer segments can be effectively visualised. The comparison of the data can be effectively observed and hence, the choice of the graph is justified.

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

The sales and profit trends across different product subcategories and regions can be effectively visualised with the line plot and subplots. The visualisation is subjected to four regions that are further divided as per the product subcategories.



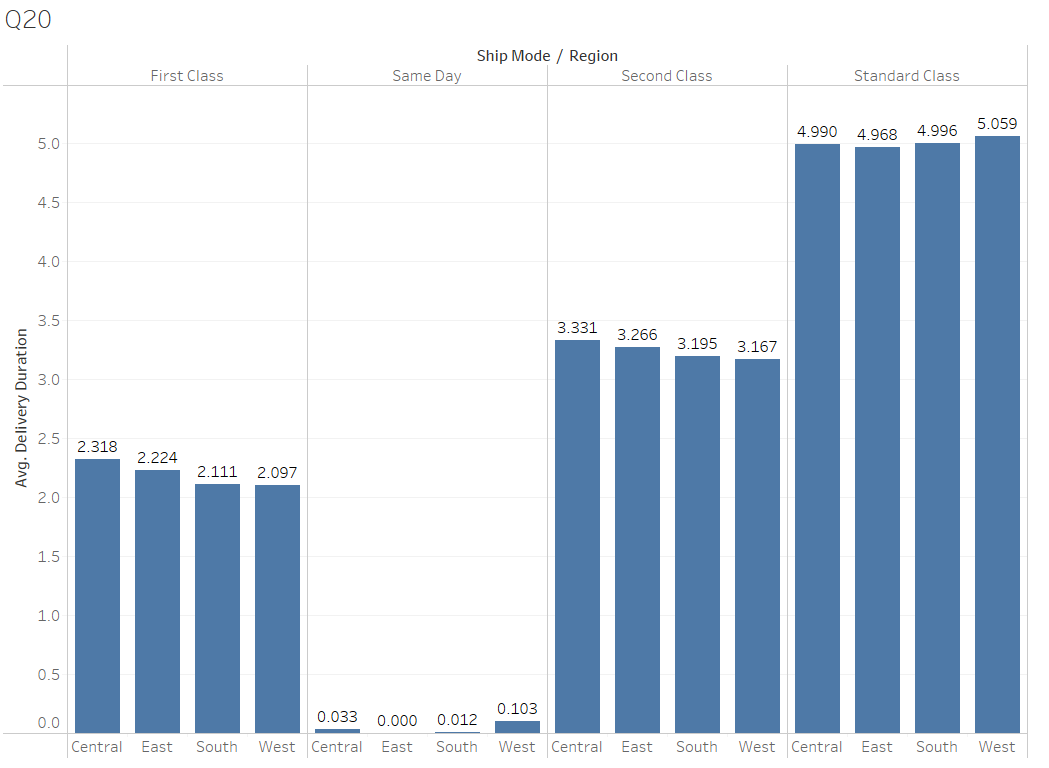
This is the most effective method to summarise the data and the required observations can be effectively done. Thus, the trend line shows how different regions and subcategories have sales and profit from the data.

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

The calculated delivery duration or the processing time can be observed with respect to the ship modes and regions. The formula is:

***“DATETRUNC('day',[Ship Date])-DATETRUNC('day',[Order Date])”***

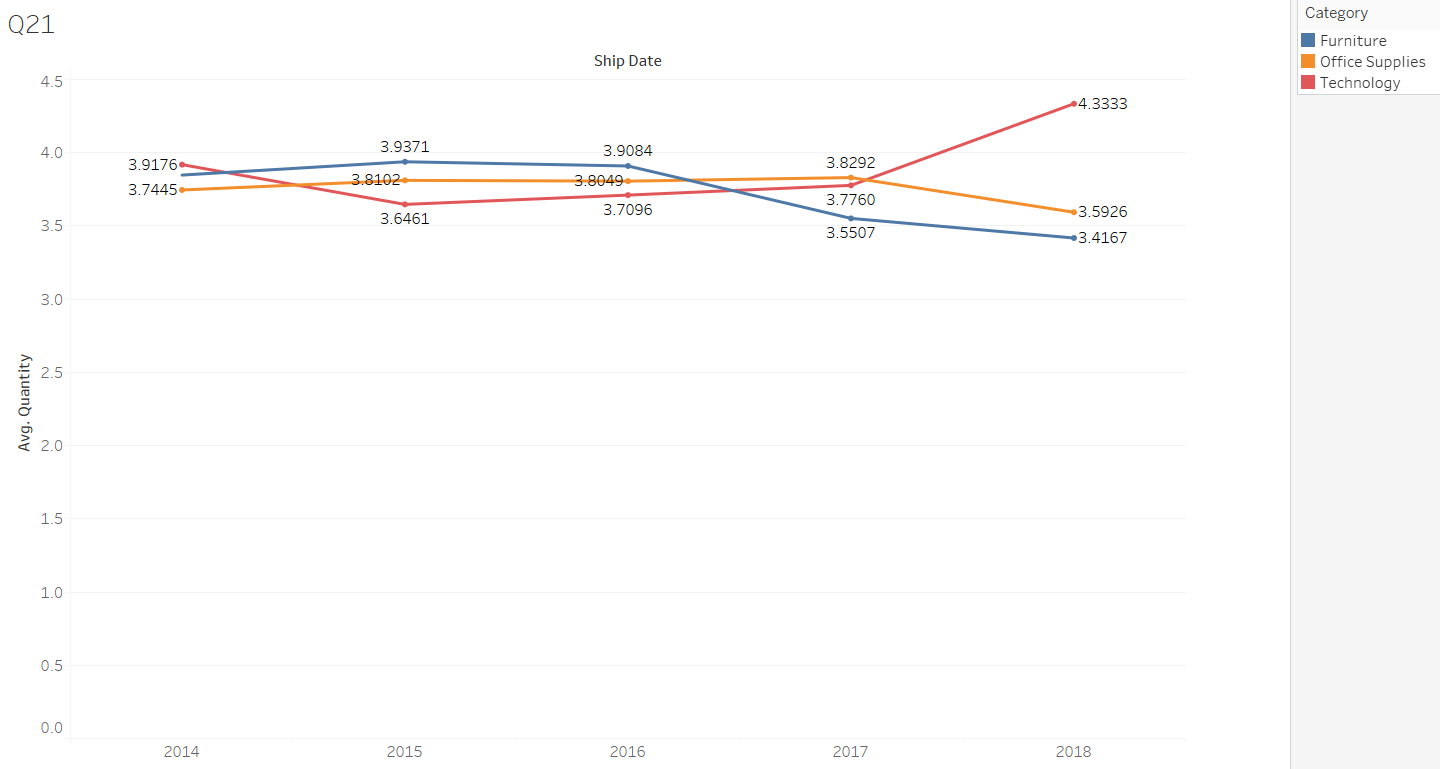
This helps identify the average delivery durations for several ship modes and regions through bar plots.



The bar plots help comparison of different factors through the same scale and it shows that the Standard class has the highest delivery time and same-day delivery is processed in the minimum time. The comparison in the region is also effectively observed and it is also effective to provide some insights on the region-based aspects.

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

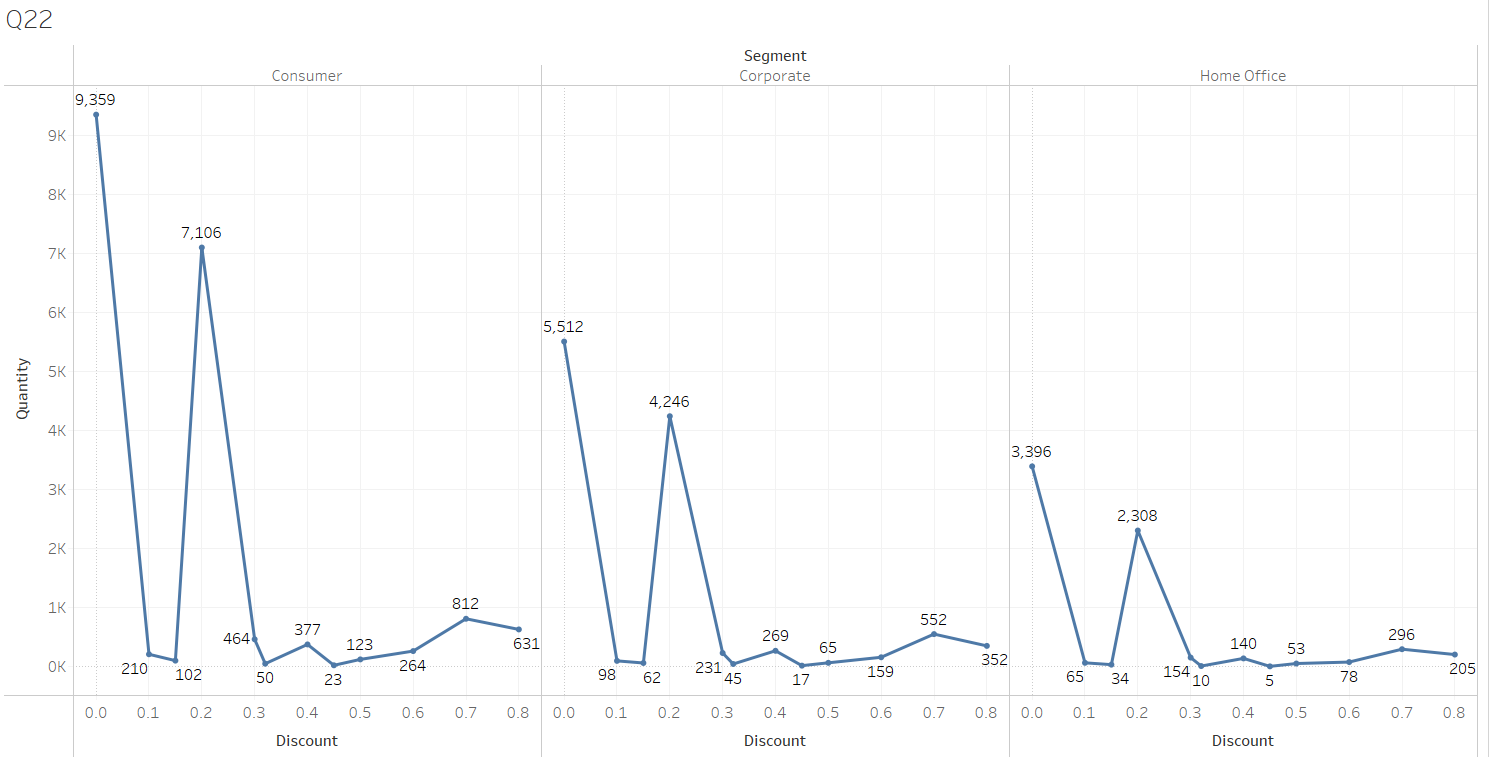
In this context, the yearly changes in the average quantity of the products have been summarised. This also denotes the different categories of the products and the quantity for each category is observed in the graph.



The choice of a line chart is effective in observing the trends of the product quantity for different categories. This is subject to provide an idea that the Quantity of Furniture is lowered and this is also true for office supplies. The technology category has gone up in terms of sales quantity.

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

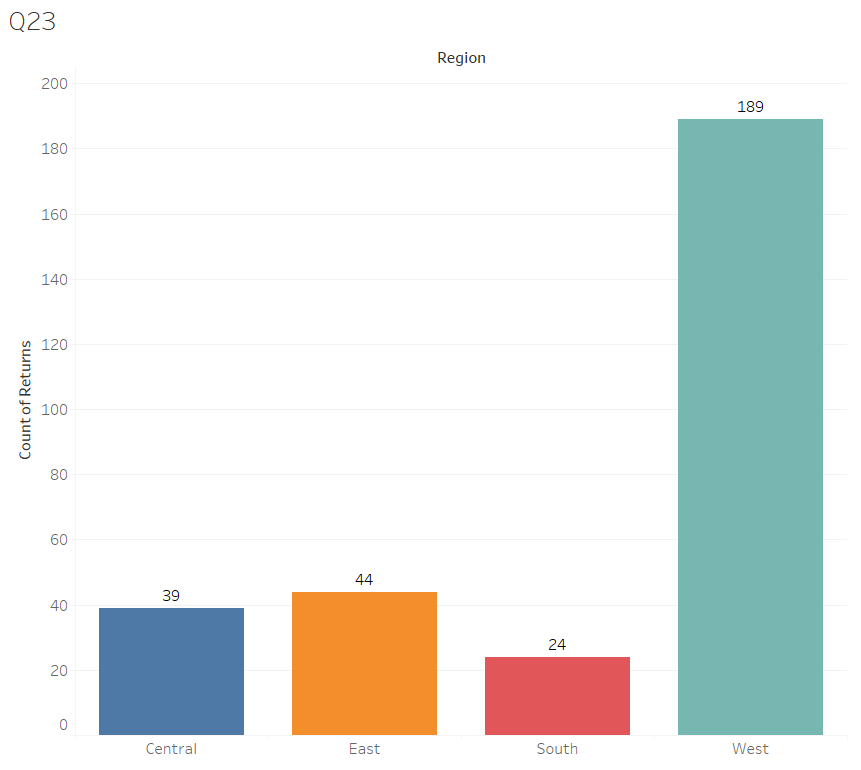
Yes, the visualisation can effectively correlate the discount rates and order quantities. The choice of the graph is appropriate as the relationship between the discount and the quantity can be effectively observed with the line chart. The choice of having different charts for different segments is essential for having a proper marking of the values as the values must be effectively observed. This is not possible with all the line plots being in a single graph.



The charts show effective results as the different trends of segments are suitably visible and the differences can be effectively established. It is also observable that the discount and quantity of products have similar trends which means the different segments of the consumers have similar buying behaviour.

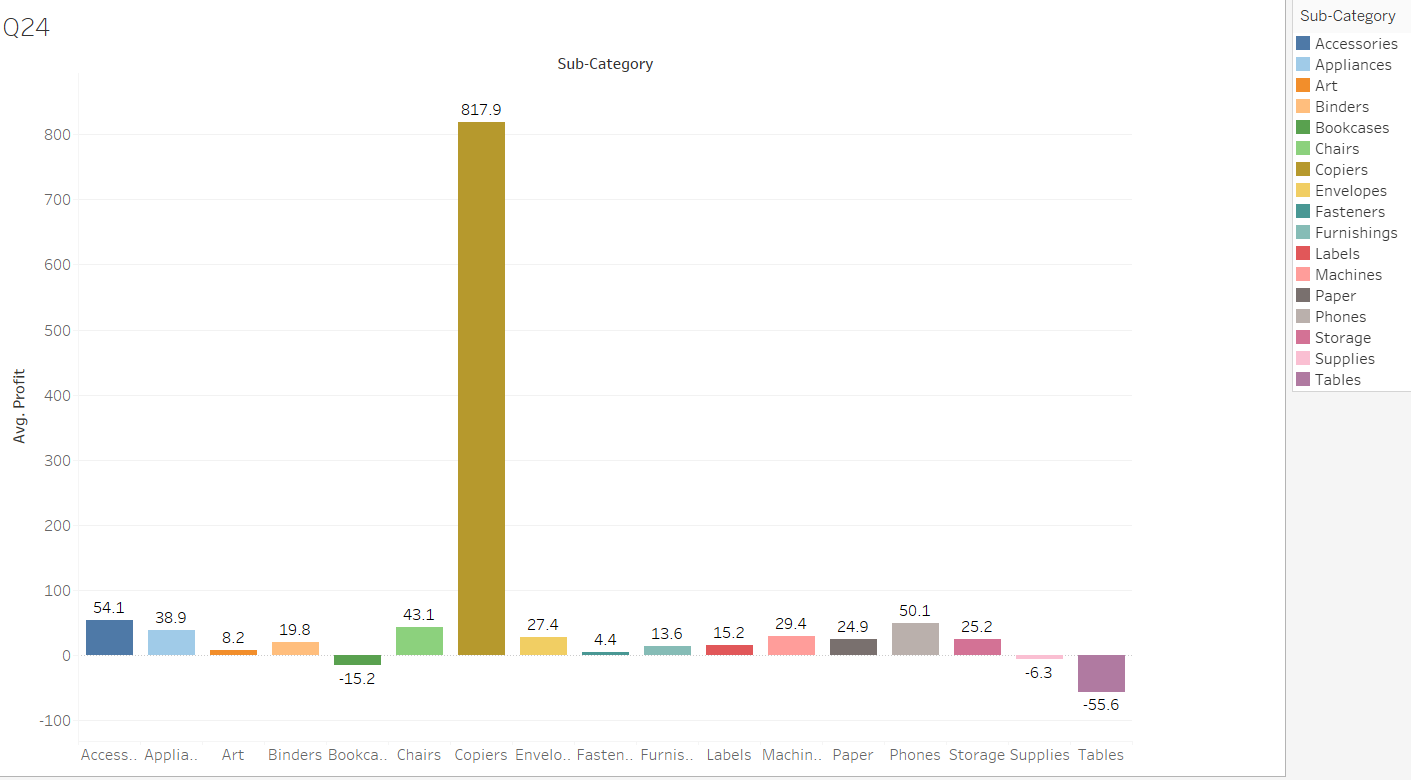
## What is the proportion of orders returned in each region within the Superstore dataset?

The proportion of orders returned is visualised in the graphical representation and the superstore data shows that the East, West, North, and South regions have different numbers of orders returned.



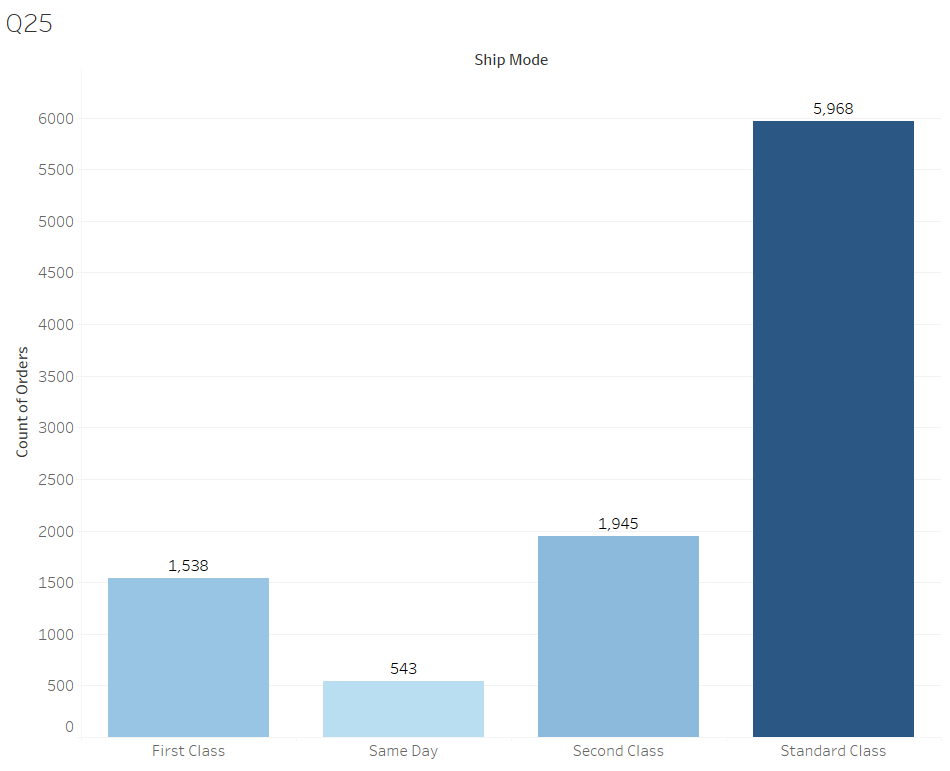
The choice of graph is justified in terms of comparative analysis in different regions and the data shows that the West Region has the highest quantity of orders returned. This is to deliver the idea of comparative data in different regions.

## Can you compare the profit of different products for different subcategories?



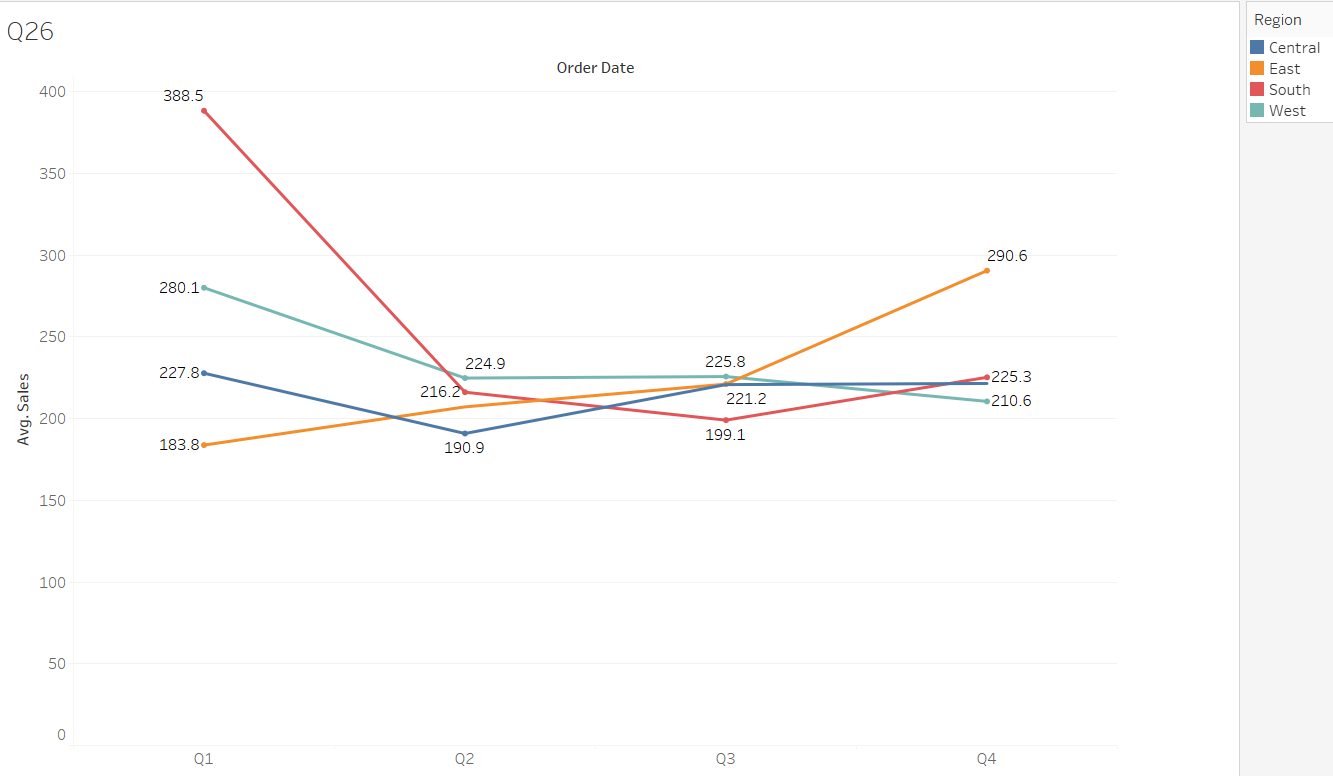
Yes, the average profit of different product subcategories can be effectively observed and compared in the following visualisation. This bar plot shows the values of profit for different subcategories and the bars clearly show the comparison. There are also labels showing the values of the profits for different subcategories.

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



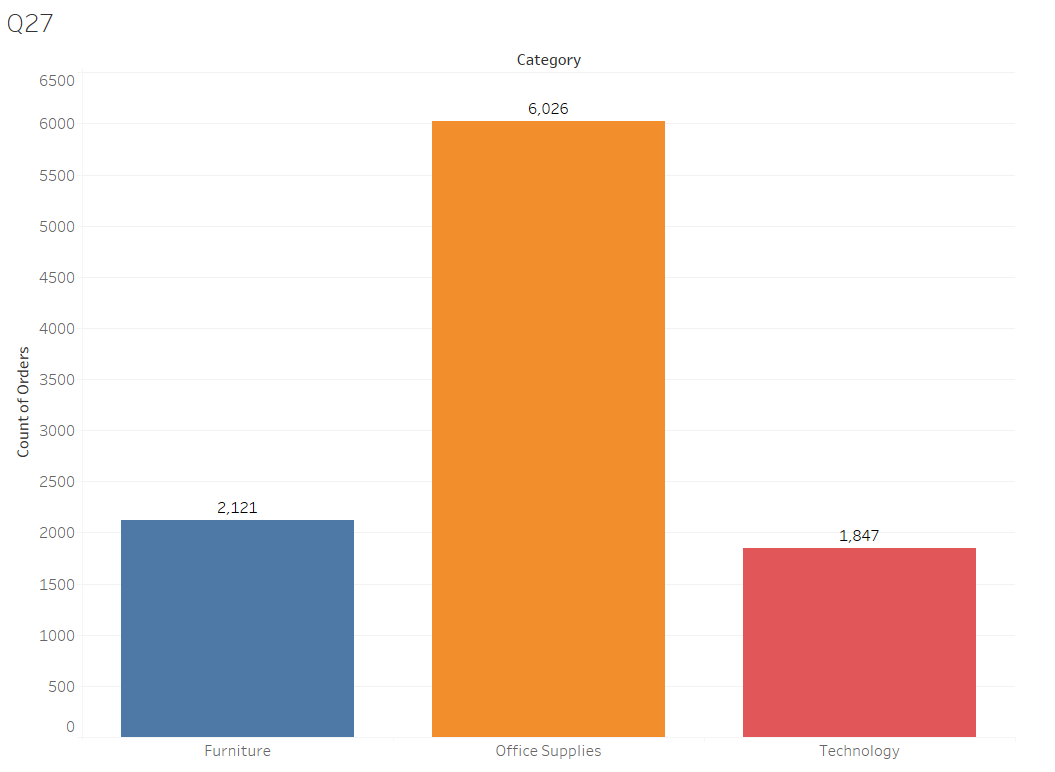
The bar plot clearly shows the count of orders for different shipping modes and it shows that the standard class has the highest count of orders. The bar plot is also effective in showing this comparison and it also shows the count of orders through the labels.

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



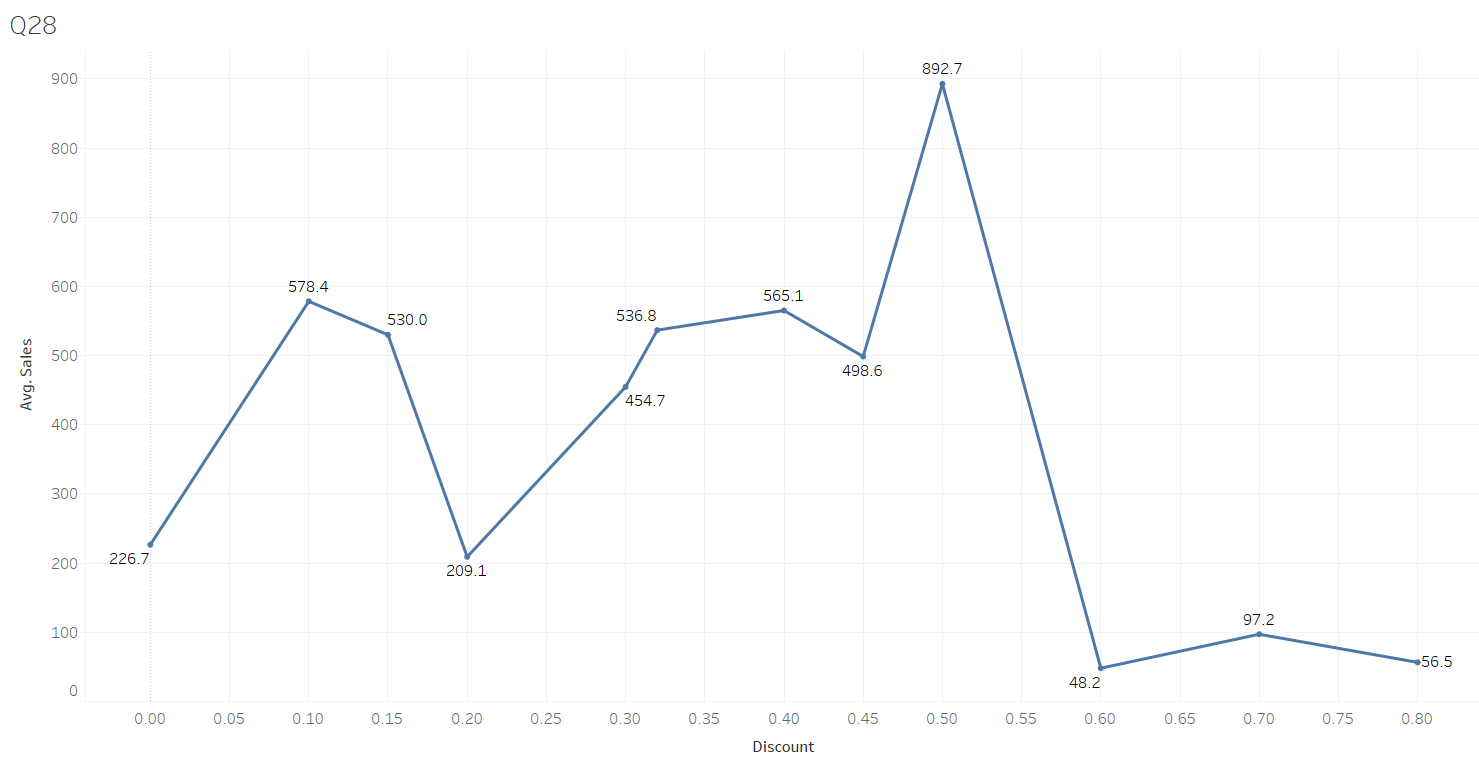
The average sales in different quarters of years are effectively visualised with the line plot and the choice of the graph is justified to observe the trends over quarters. This shows that in the South region, the sales have been high in the first quarter and it has been lowered and have gone up slightly in the fourth quarter. The other lines also show the ups and downs of the sales in different quarters and some effective trends are observed in this context.

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



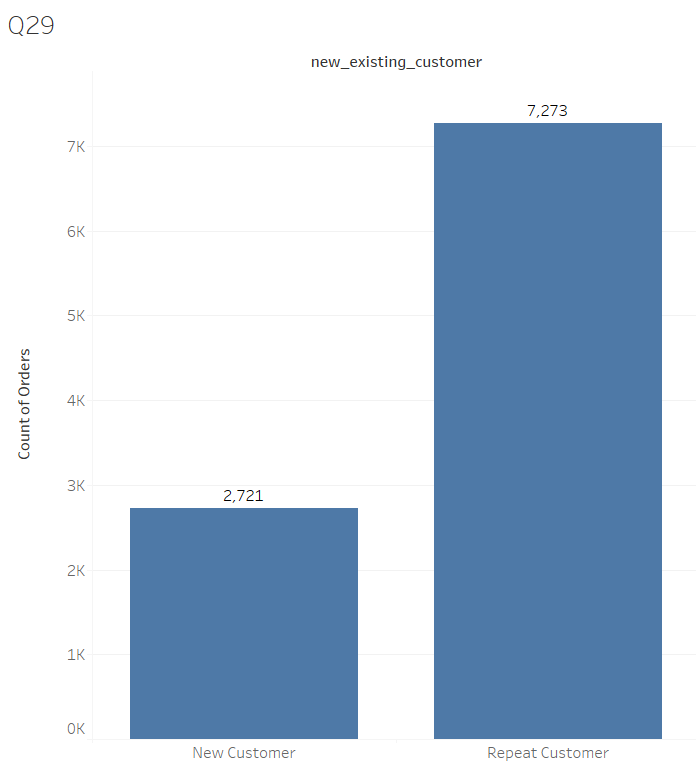
The distribution of the count of orders for different categories can be effectively observed in the graphical representation. The comparative bar plot shows that the office supplies have the highest order count. The choice of the graph is justified to understand the distribution of the count of orders.

## What is the relationship between discounts and sales?



The simple line plot shows the relationship between the discounts and the sales and it shows that for certain values, the average sales acquire high value and in case of higher discounts, the average sales are low. The sales are highest between the 0.48 and 0.54 value and lowest is between 0.6 to 0.8.

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



Two formulae have been applied to achieve new and repeat customers. The first is

***“{FIXED[Customer ID]:YEAR(min([Order Date]))}”***

This develops a column showing the first year of occurrence for the customers. It is further used in gaining the idea if a customer is new or repeat using the following formulae:

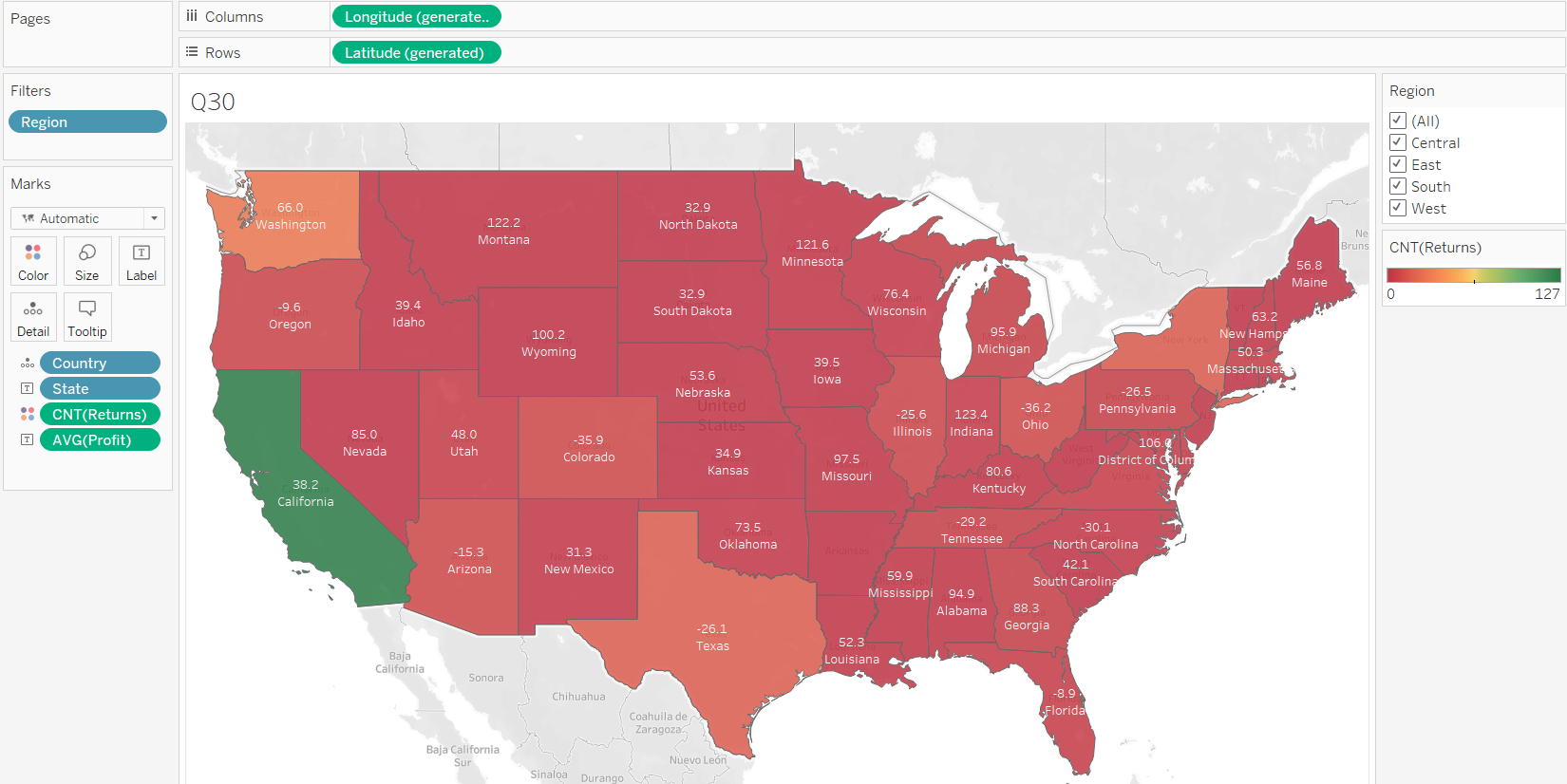
***“IF YEAR([Order Date]) = [FirstCustomerYear] THEN 'New Customer'***

***ELSE 'Repeat Customer'***

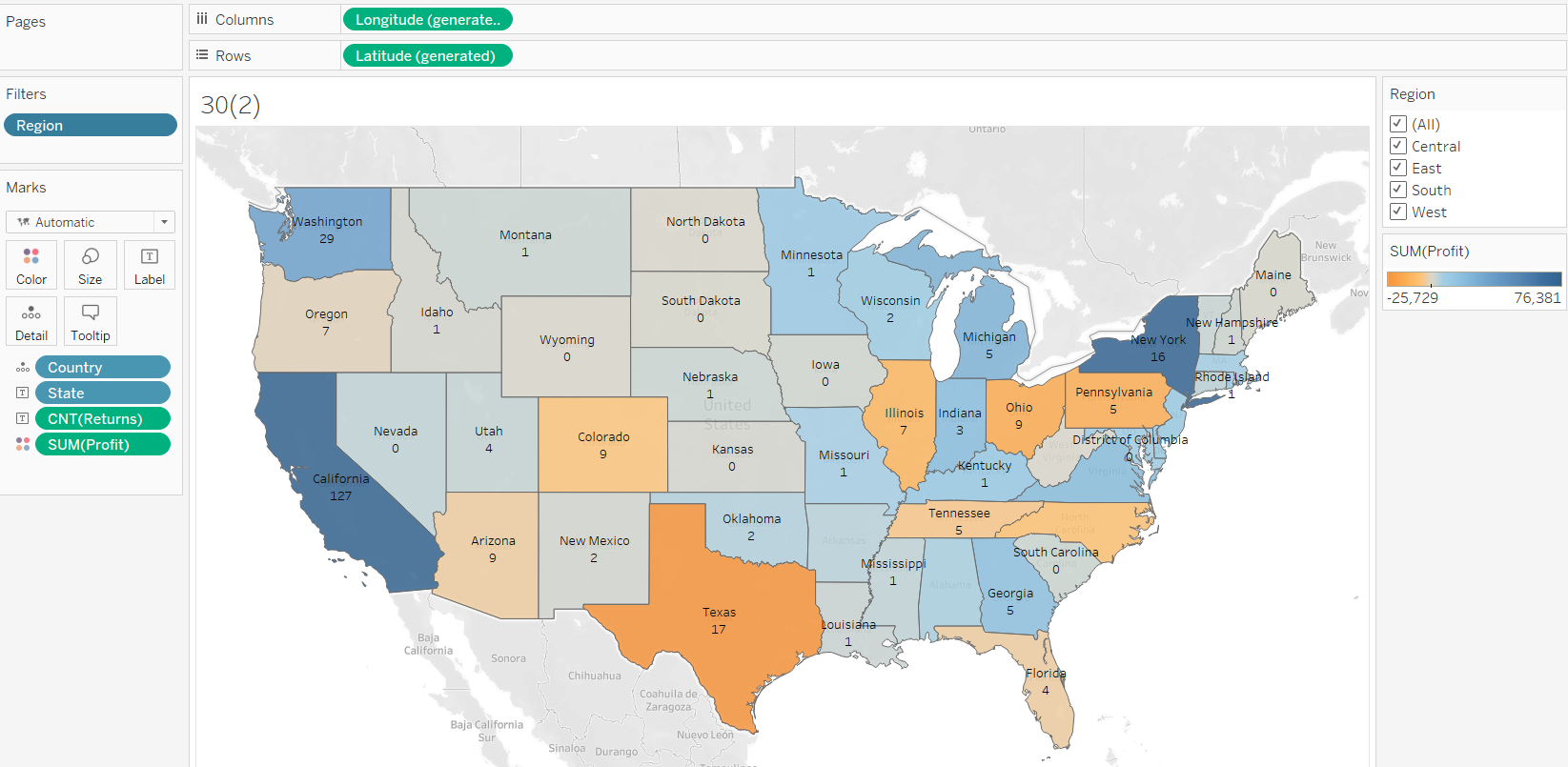
***END”***

This shows if the customer is new or repeat or an order is done by a new or repeat customer. This column is developed with the data and it can be used in this present visualisation. The visualisation shows how the count of orders for new and repeat customers and repeat customers will have a higher number of orders.

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



This graphical representation shows the geographical distribution of the overall returns. The respective profitability is marked and it can help assess its relation with the number of overall returns in different states.



This is another plot in a similar format but showing the scale of profitability with a colour code. The count of returns is also seen on the labels and these are subjected to get the idea. Both charts together show the relation between the returns and profitability.

Both the graphs are capable of showing the returns and profitability over the same geographical distributions and the filters help determine the regions of the plots respectively.