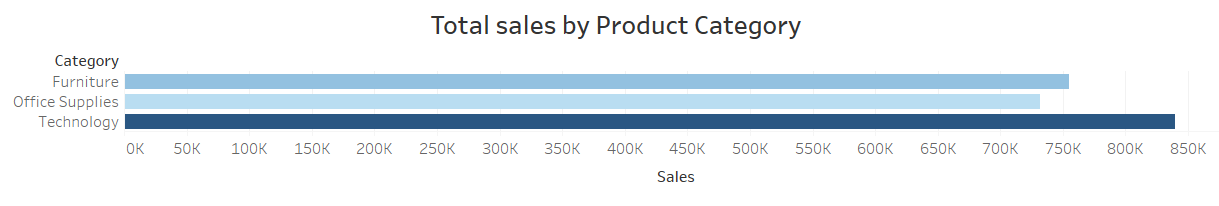
Superstore Sales Analysis EDA using Tableau

Submitted by – Anurudra Jena

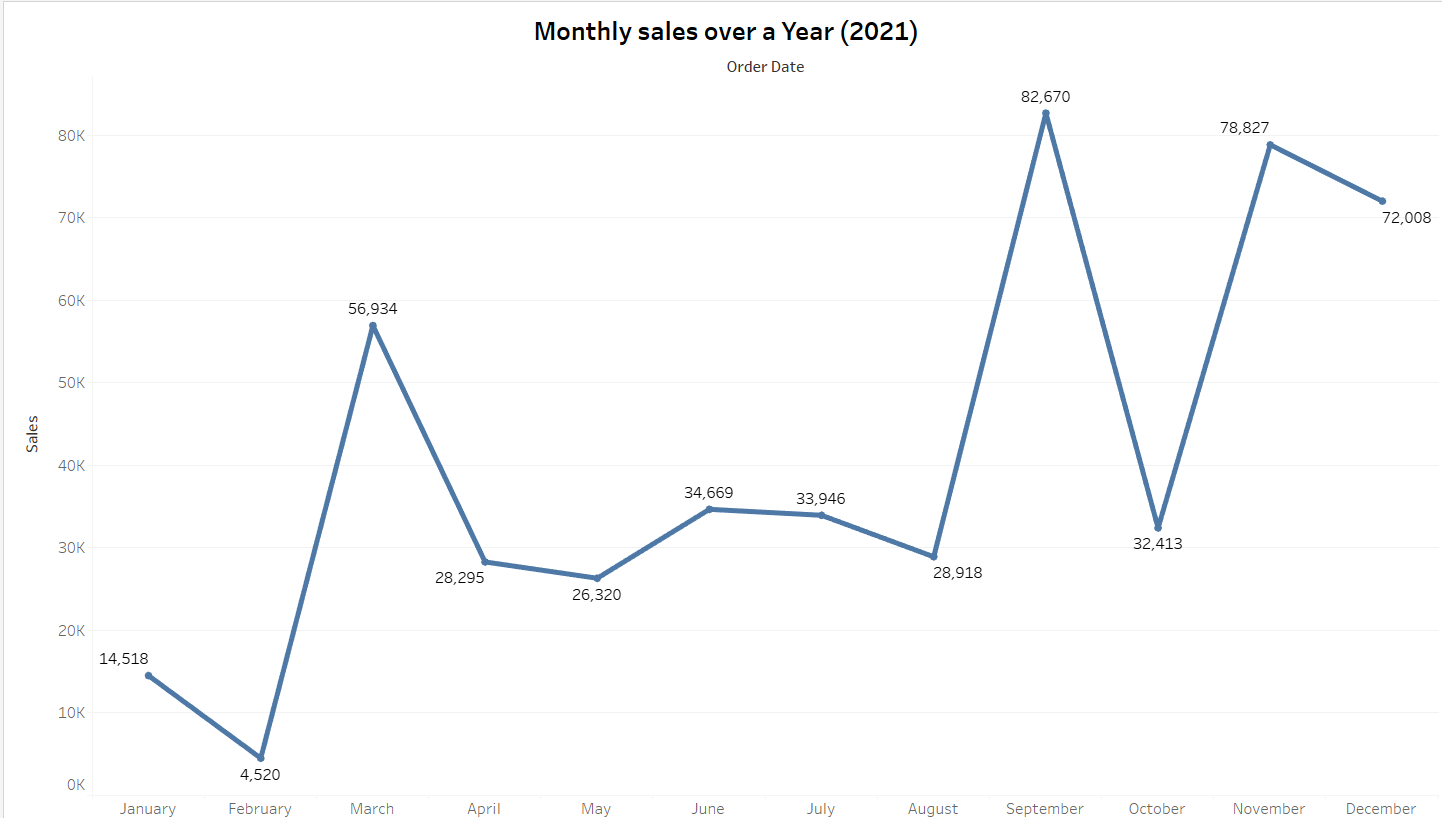
As we’re provided with a number of questions to be answered upon gaining insights by performing different visualizations in the Tableau application, below are the results of the analysis taken and answers to those questions.

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



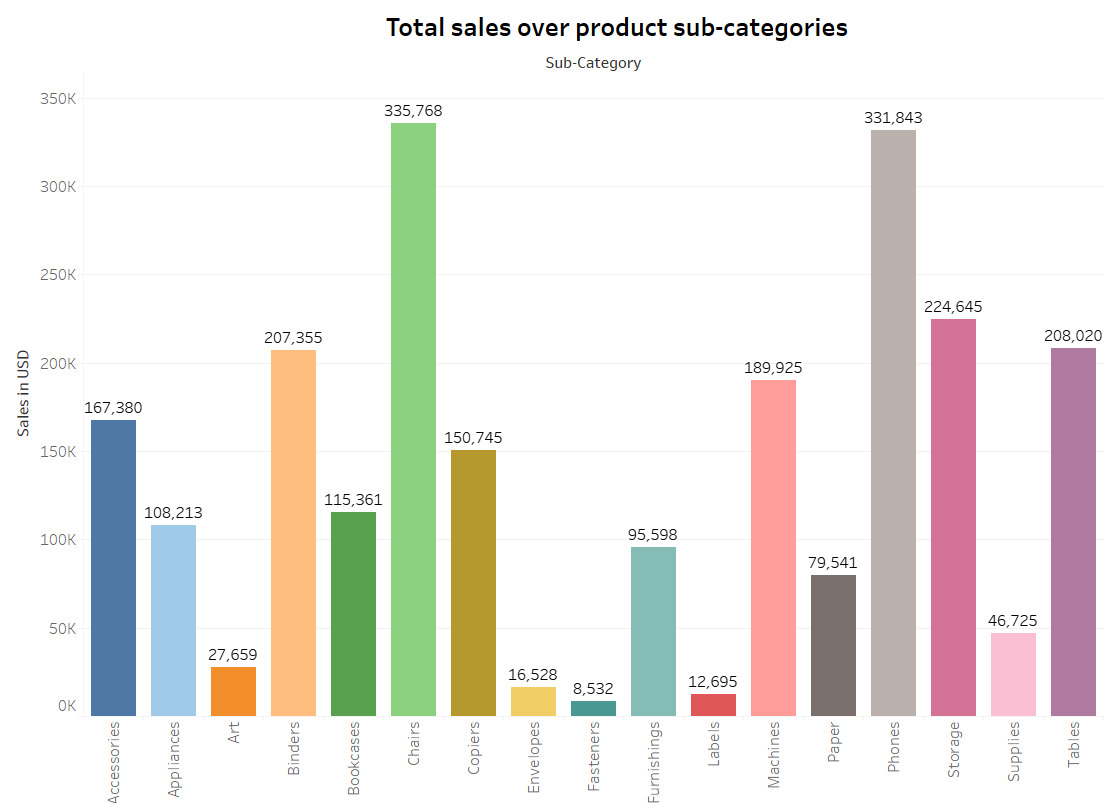
The sales value of different (3) product categories have been depicted through a bar graph since we needed to compare categorical data.

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



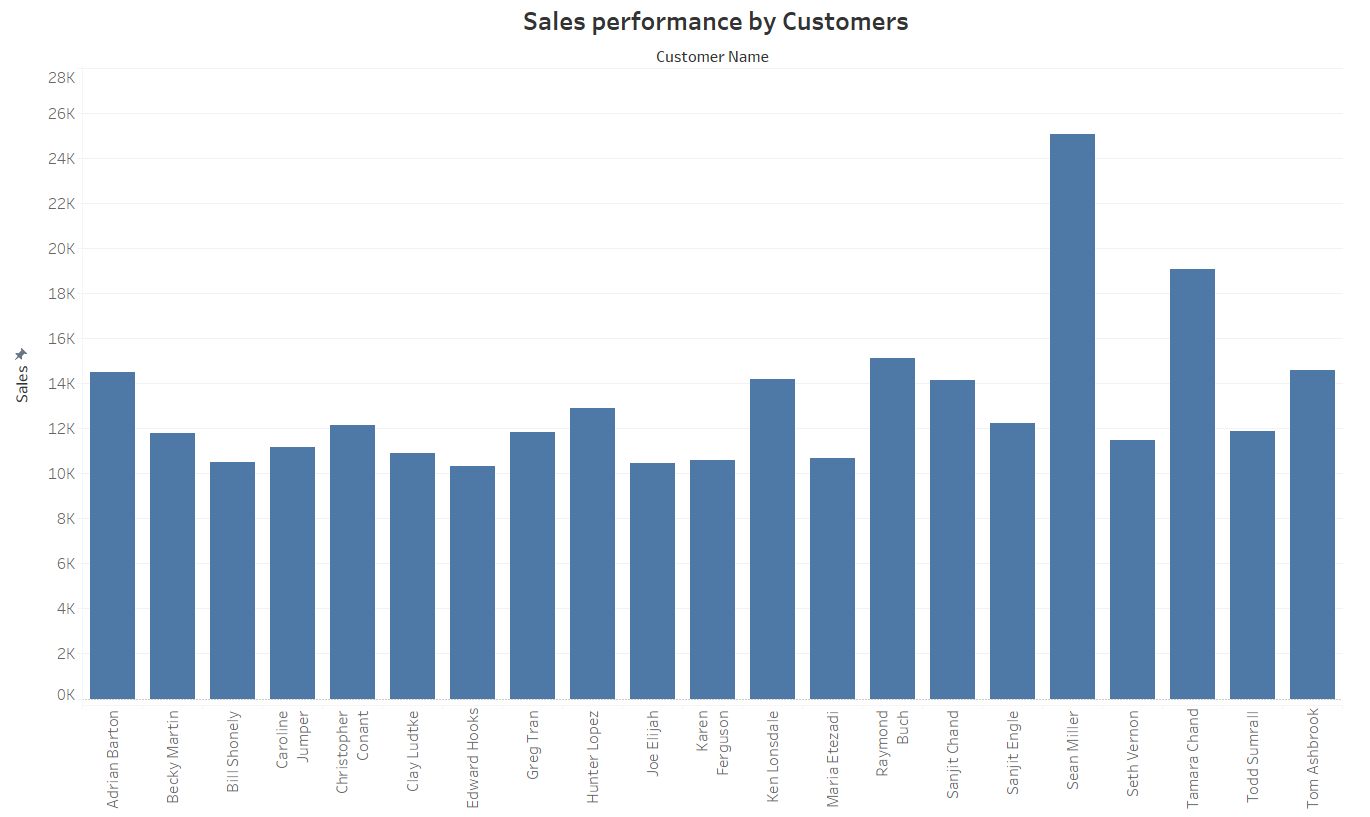
The total sales value distributed over the months of a year (2021) has been depicted above by a line graph as the data is continuous and time based.

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



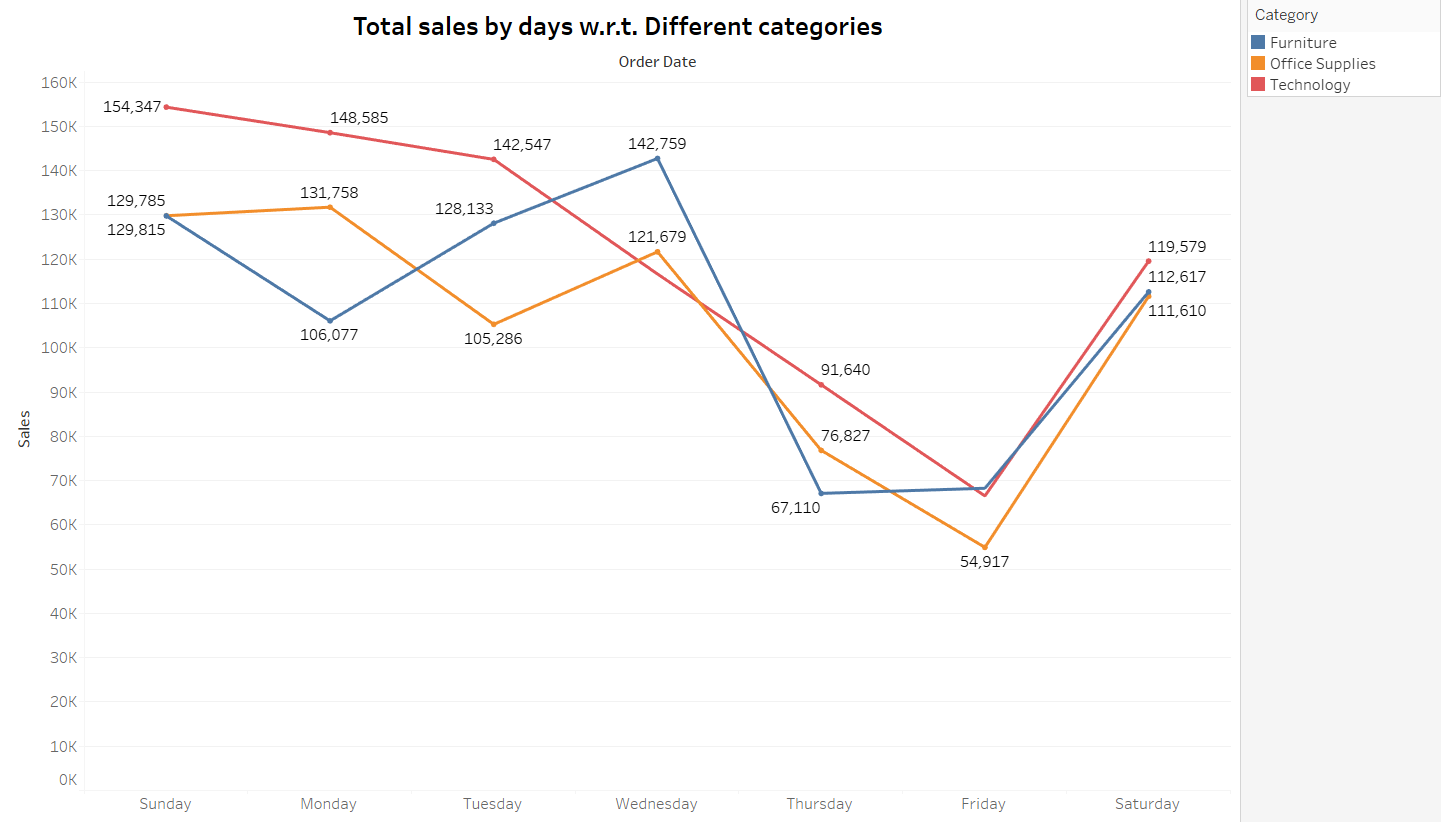
Above visualization depicts the overall sales value with respect to different product sub-categories. Since the data was categorical, it could be simply visualized by bars.

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



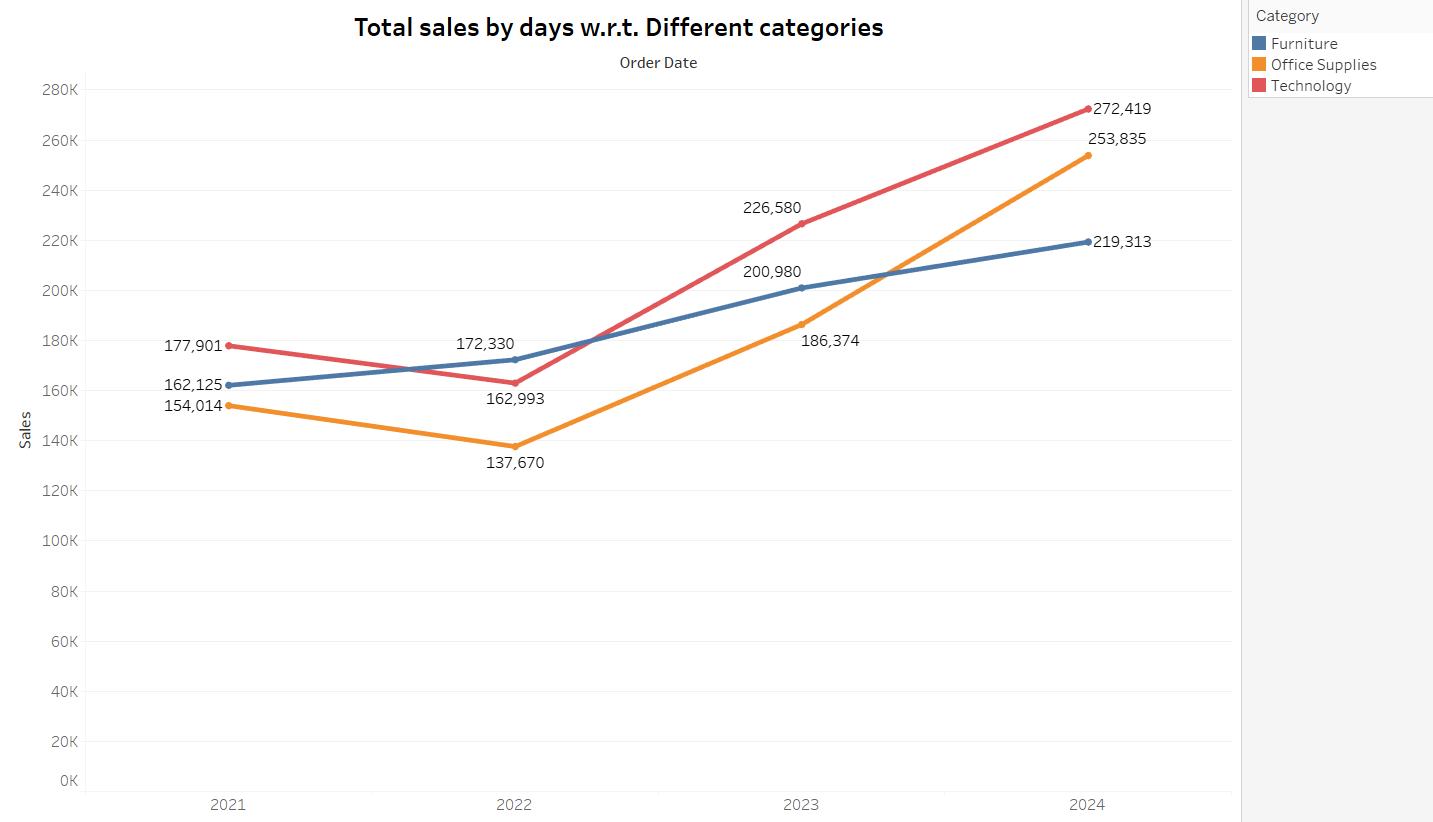
Visualizing total sales value for individual customers would not produce any value as there are innumerable customers present in the database. The result is shown above in the visualization.

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



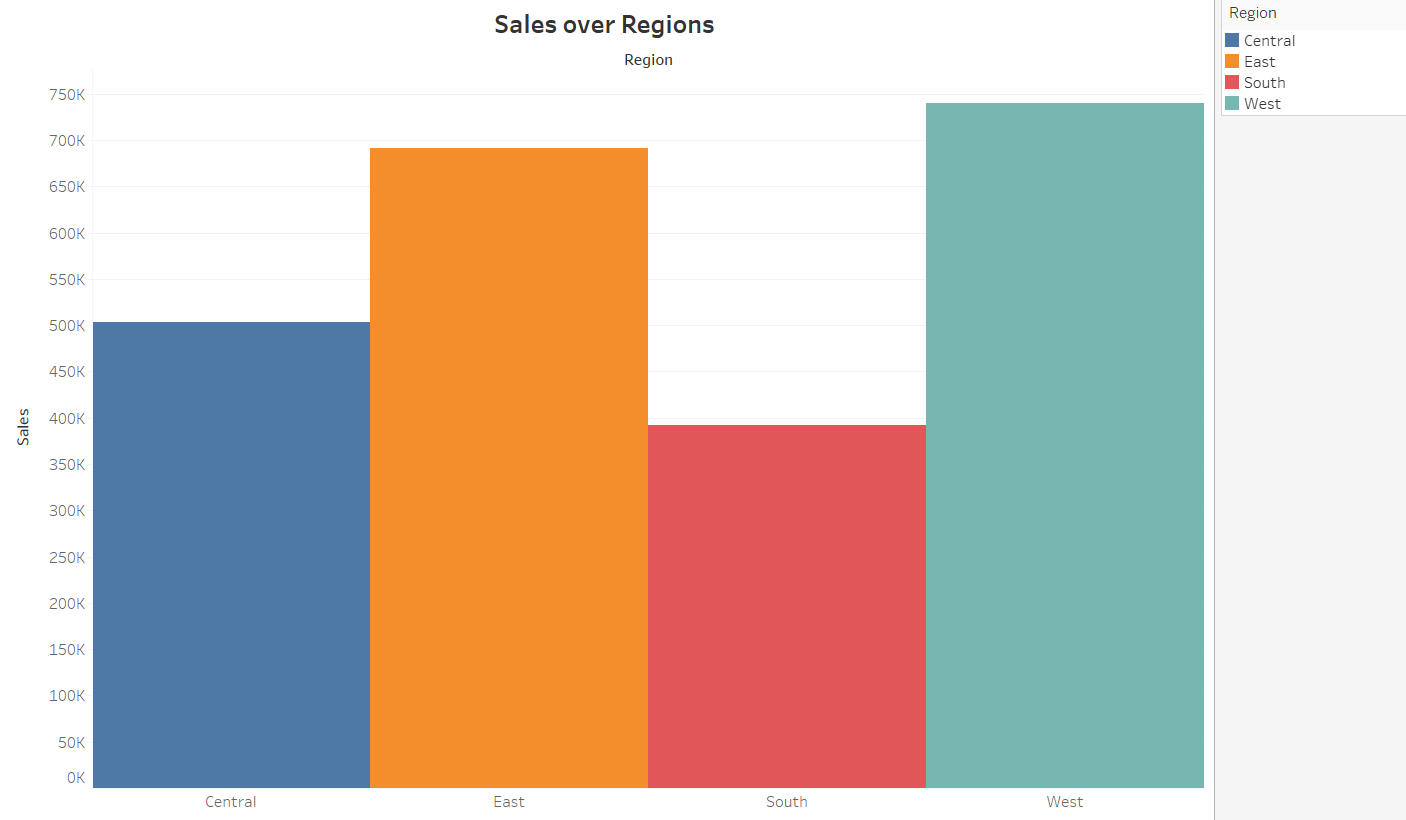
The above visualization speaks of the total sales value over different days of a week according to the entire dataset. It is depicted by a line graph as the data is continuous and time varying.

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



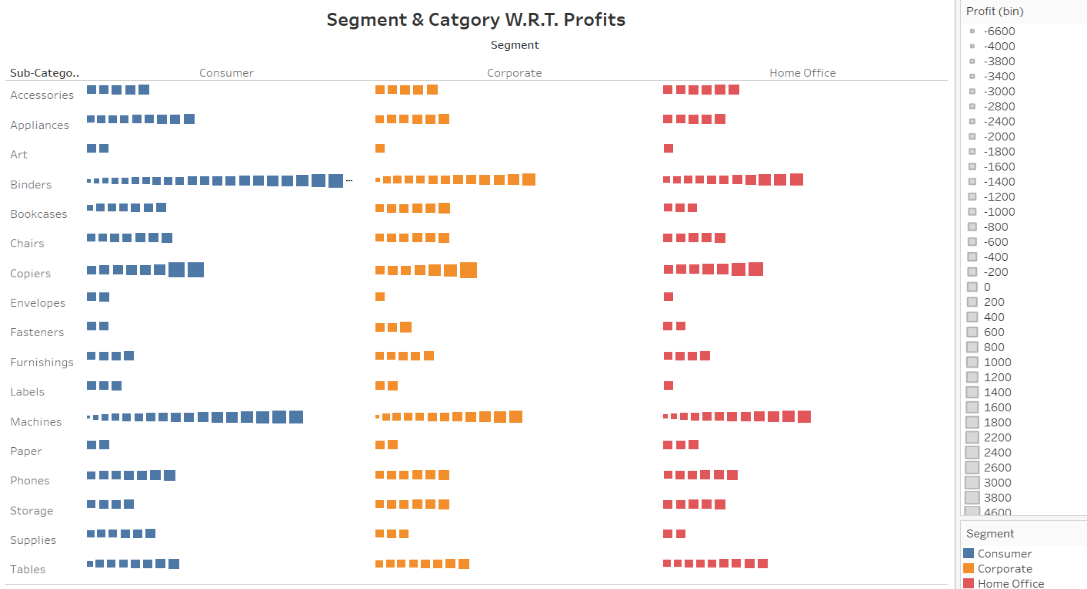
In the above visualization the total sales value of different product categories has been depicted w.r.t. time (in situ. 4 years). The data was continuous and time dependent and hence has been depicted in a line graph.

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



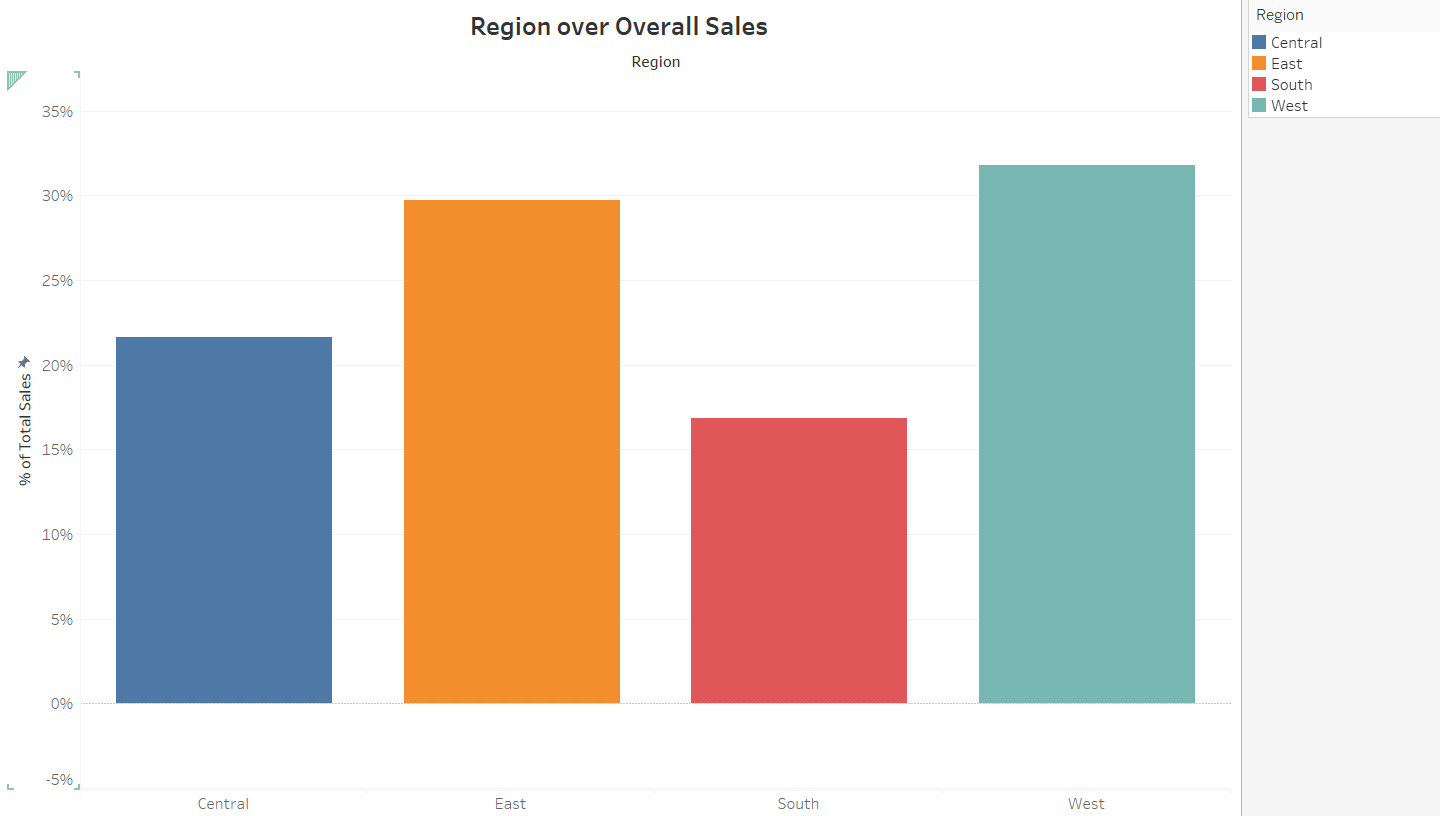
The above depiction says about the distribution of sales across 4 different regions, and since the data was categorical, bars come to the rescue.

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



Although the visualization was made for different subcategories and customer segments, it is advisable to investigate the segment individually by bar graphs as it becomes quite meticulous to render conclusions from above.

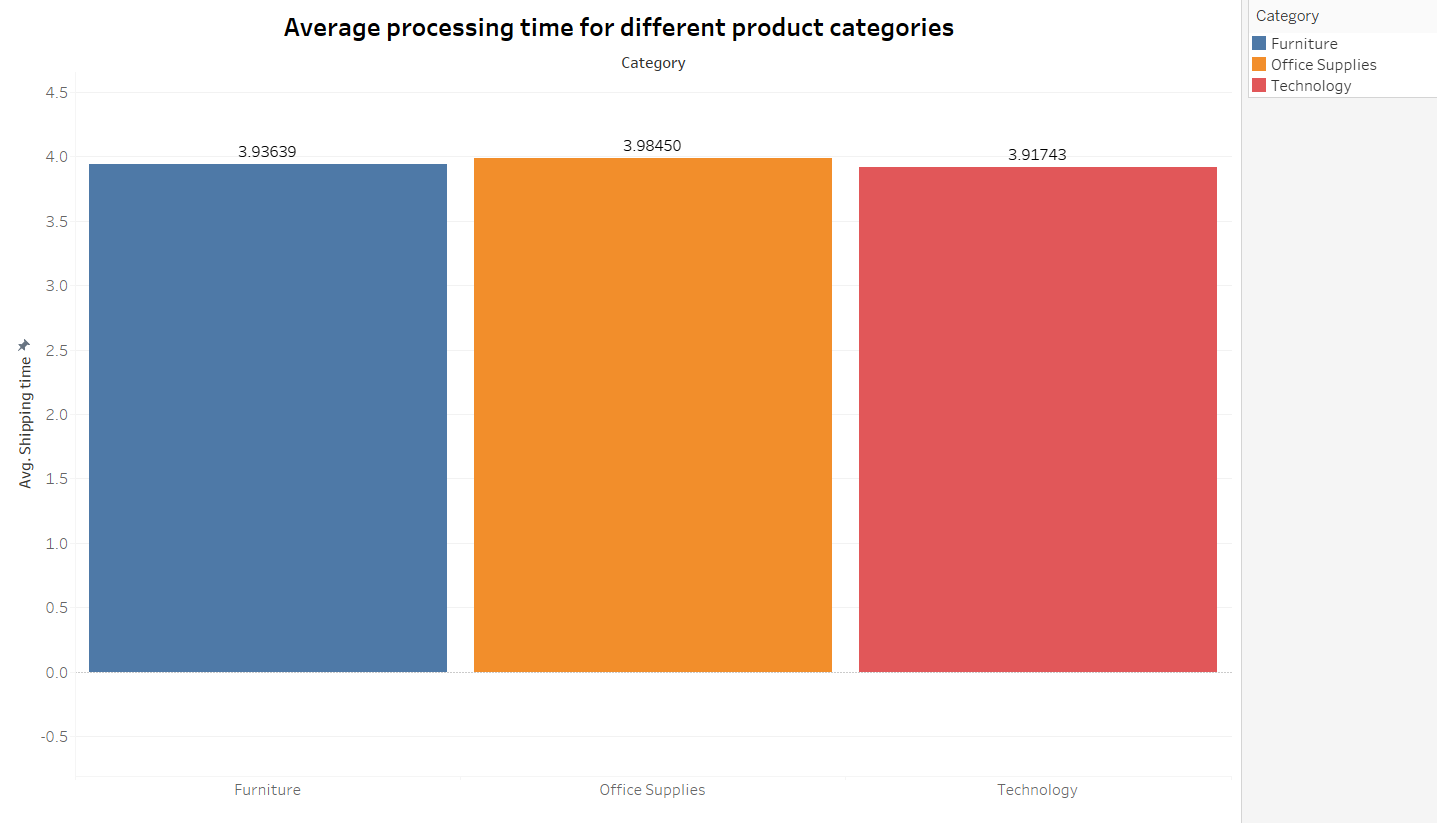
* What is the percentage contribution of each region to the overall sales? Can we visualize the profit margins associated with different shipping modes and customer segments?



The above bar chart shows the percentage of sales value over the whole across different regions.

Since the data was categorical, it was depicted with bars.

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



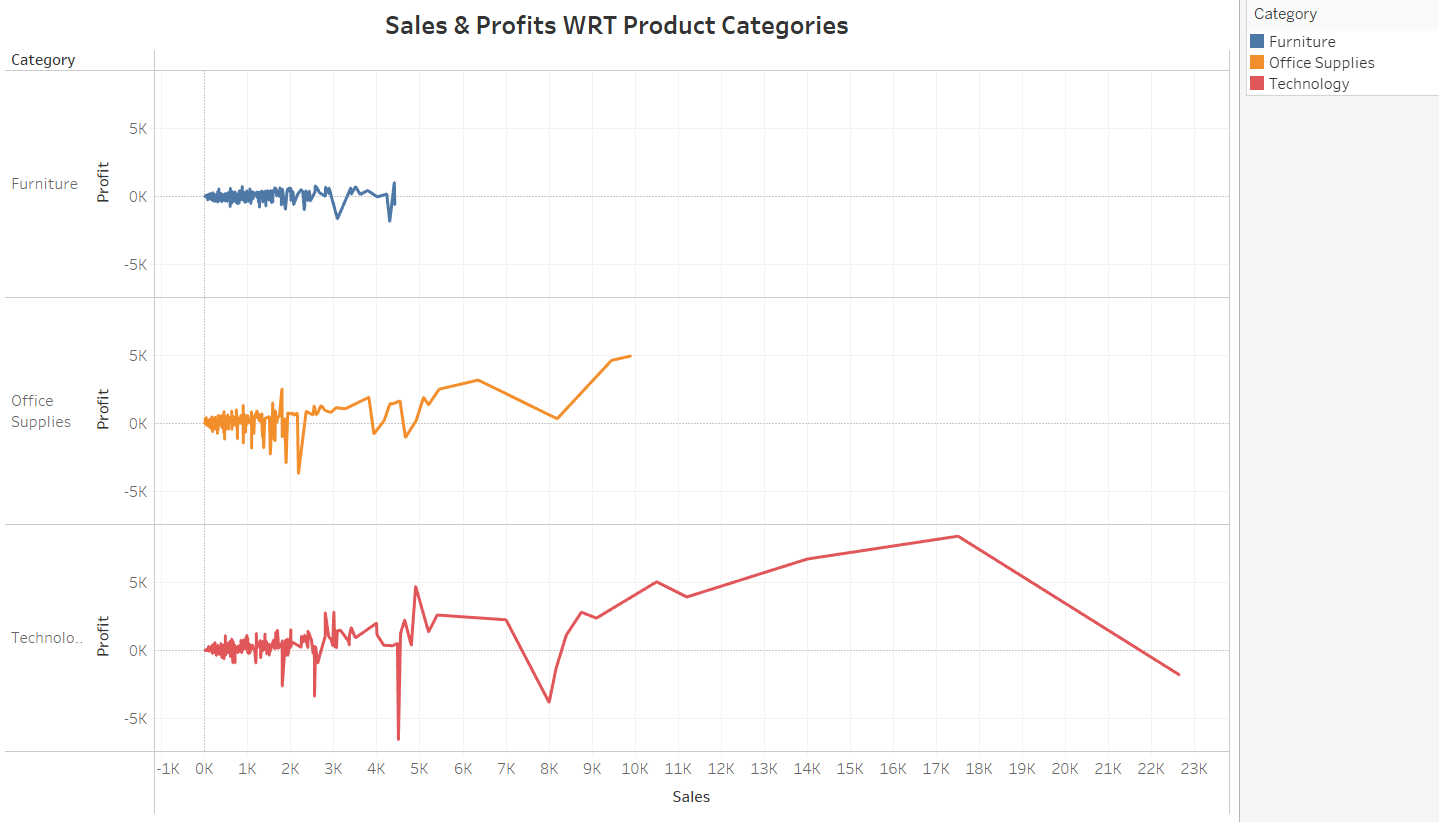
The above chart represents the average processing time for each product category; and it has been depicted by graphs as the data is categorical.

* How do discounts affect overall profit?



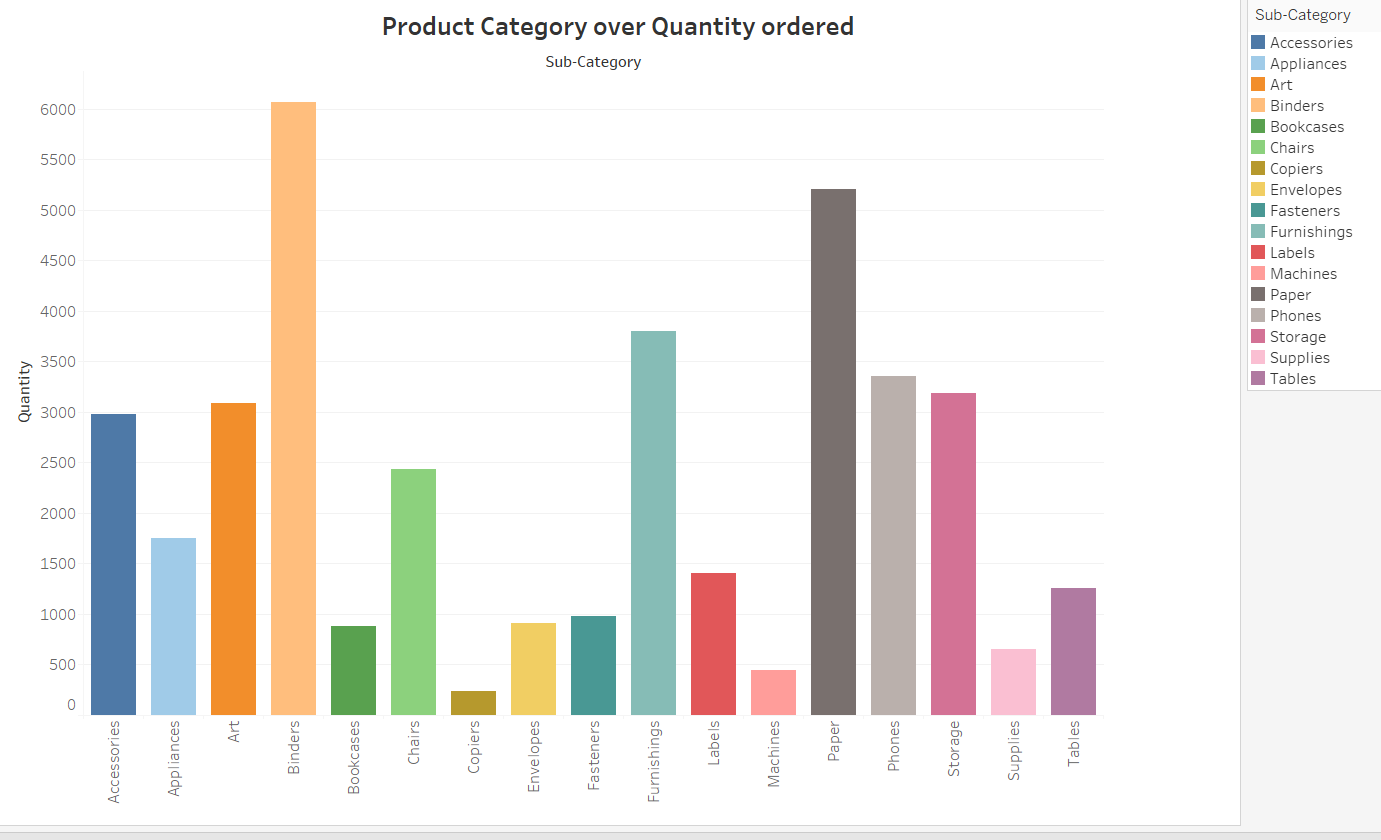
Above pictorial depiction is not quite figurative as the data is not discrete is distributed across the while dataset.

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



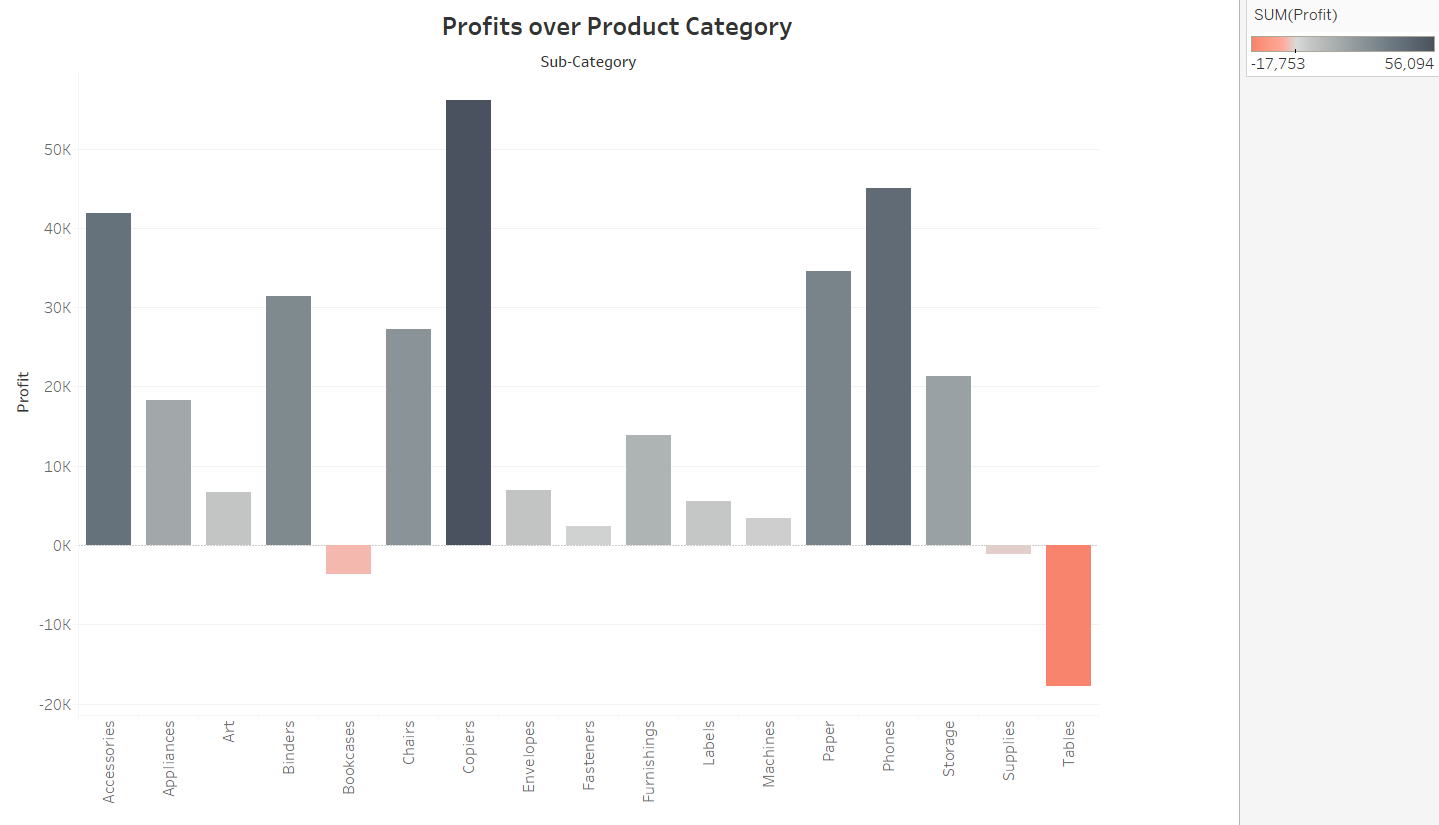
Since the profit data is continuous, it has been depicted in line charts, but since the count is high no definitive result can be sourced from it. Some insights can be gained from the data but it won’t be much informative.

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



The above visualization depicts the overall ordered quantities of different sub-categories and has been represented by bars as the data is categorical.

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



The above visualization depicts the overall profits generated by different sub-categories.  
Bars represent the categories and the colors represent the profits.

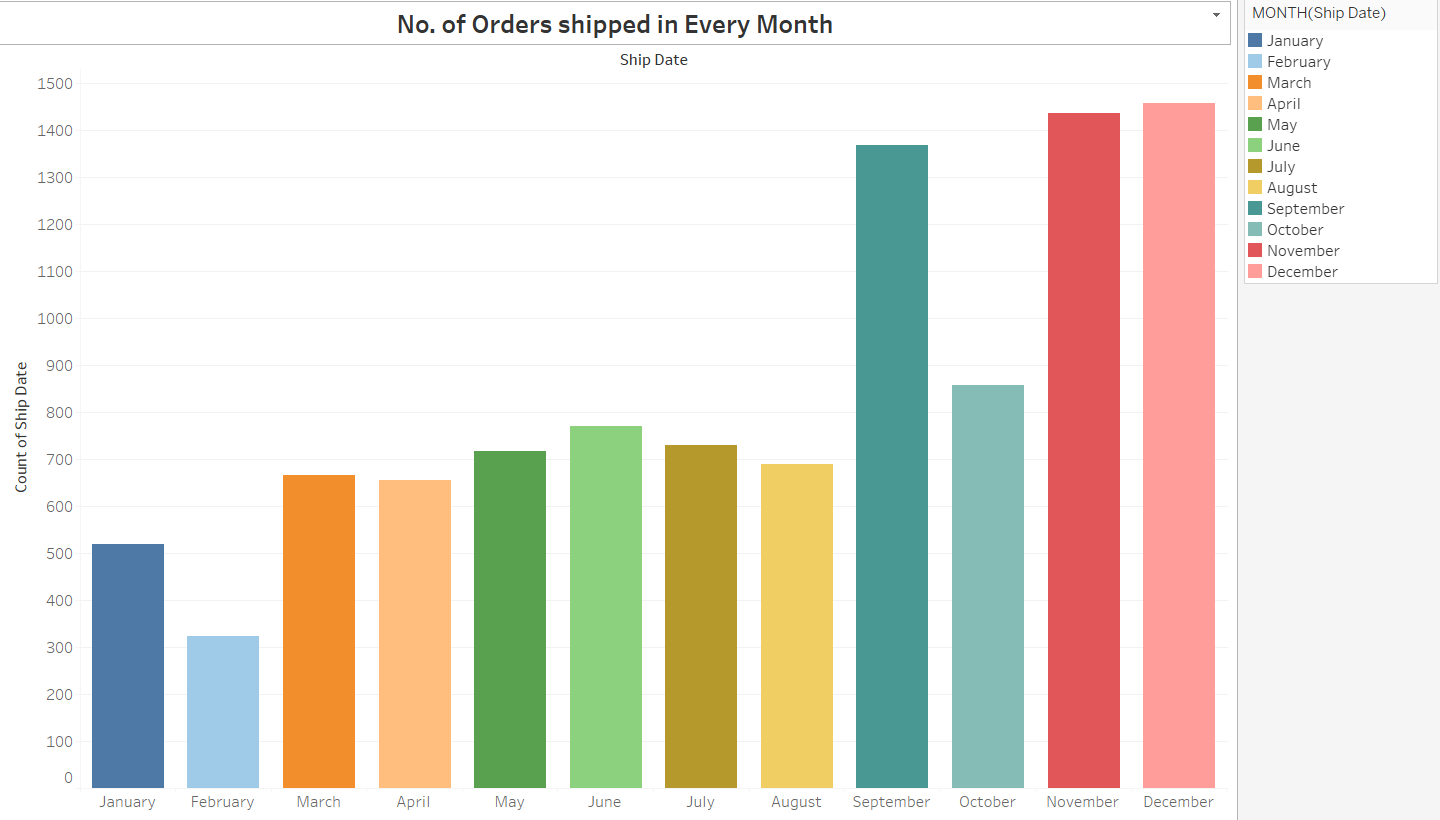
* Can we compare the shipping time distributions for different shipping modes?



The above chart represents the average shipping time with respect to diff. shipping mode.

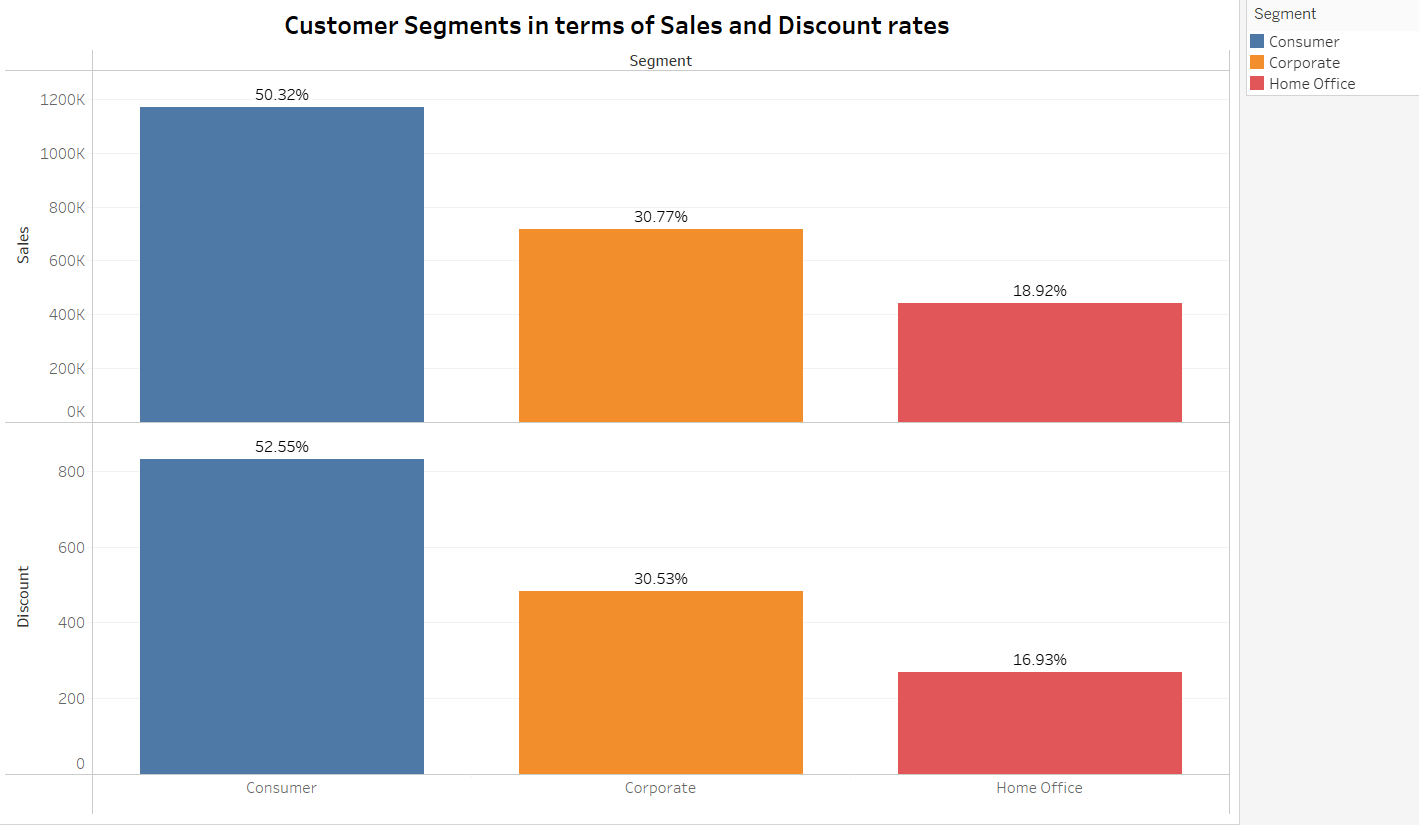
The bars represent the categories of shipping mode.

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



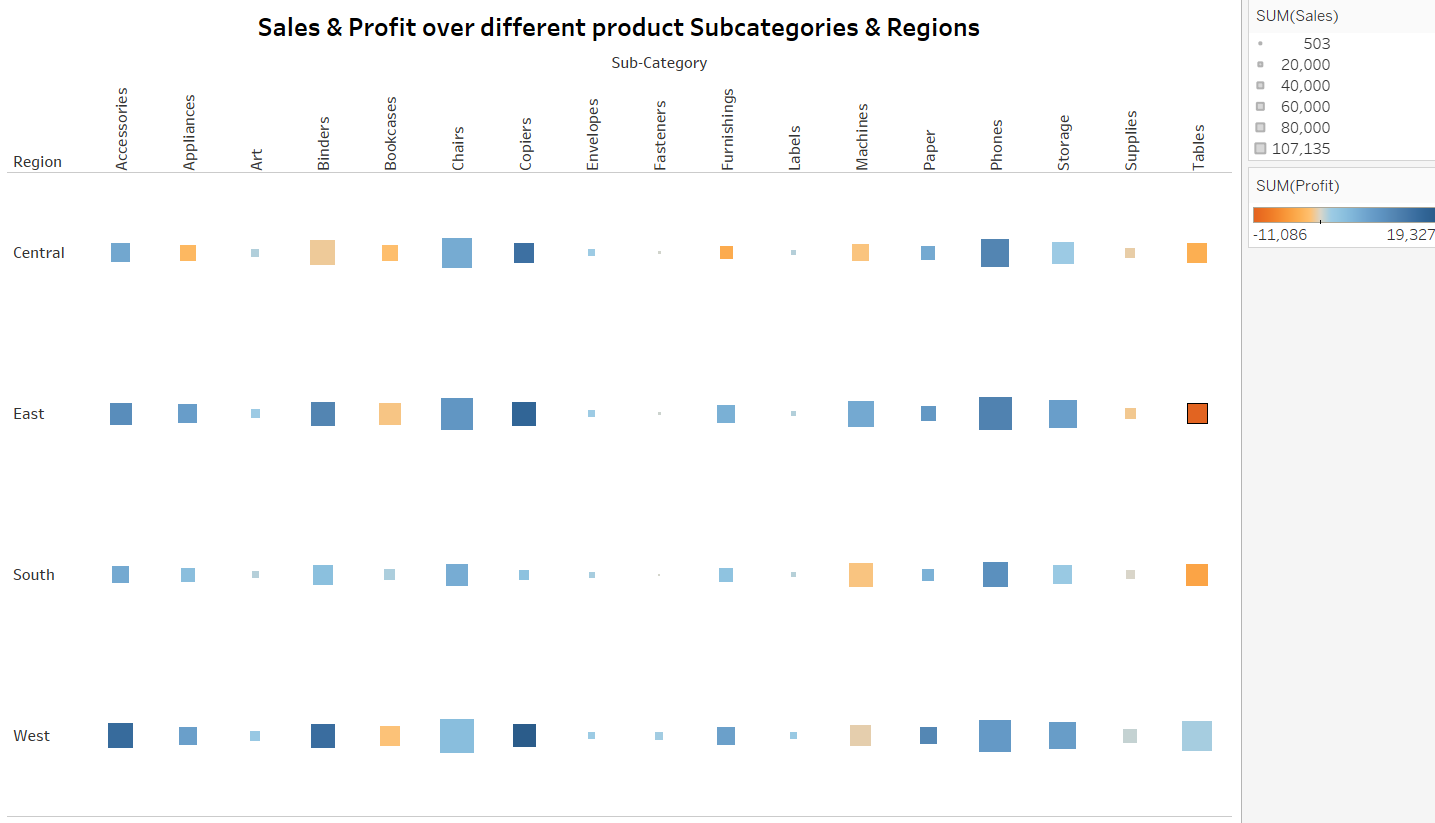
The above chart represents the count of orders shipped in different months in the entire dataset. Since the data is categorical, bars can represent the months.

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



The above chart represents the customer segments and its relationship to discount and sales.  
since the sales and discount data is continuous, and the segments are categories, they have been dissected into two separate bar charts.

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



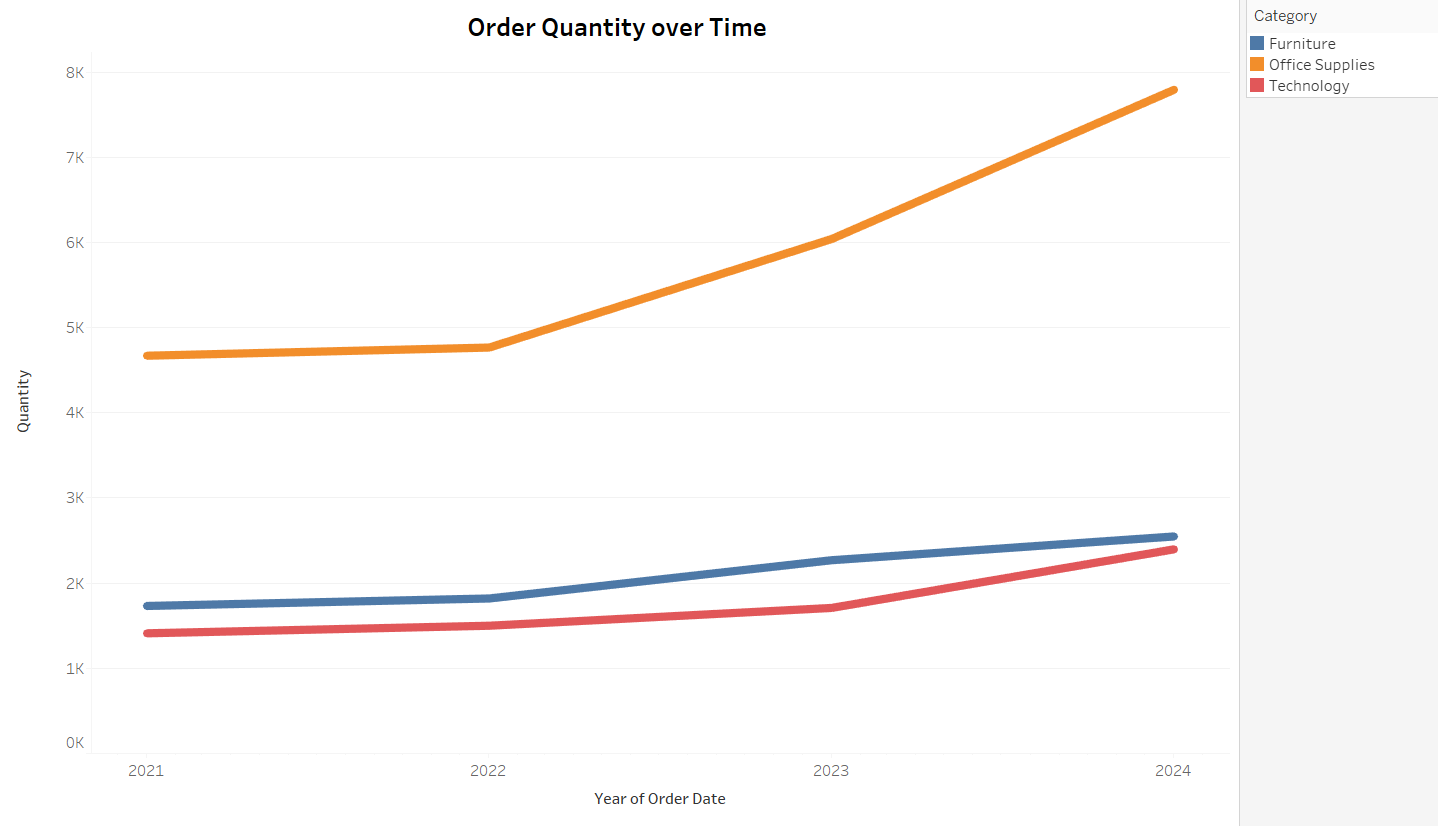
Although the visualization was made for different subcategories and regions, it is advisable to investigate the segment individually by bar graphs as it becomes quite meticulous to render conclusions from above.

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



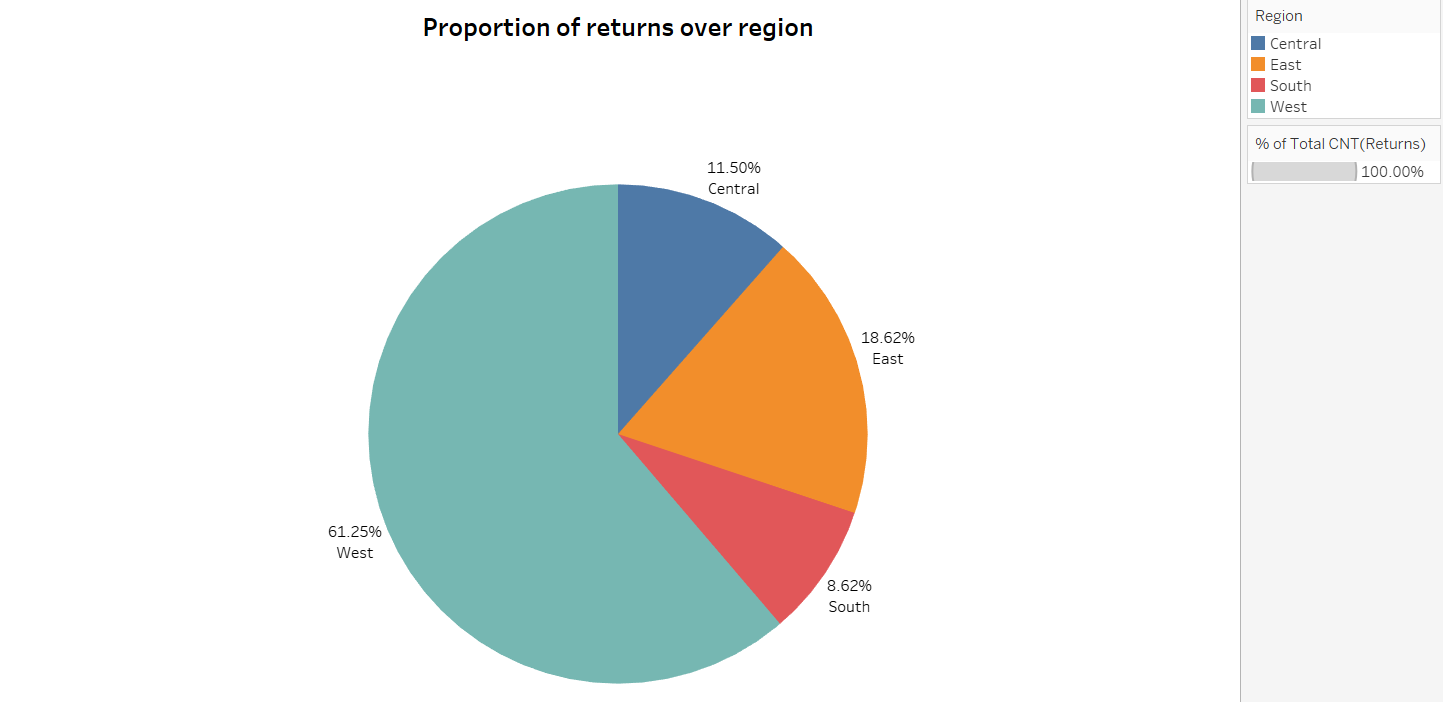
The above visualization pictures the average delivery duration for different regions vs. ship modes. Since most of the data is categorical and shipping time is discrete, bar chart was used.

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



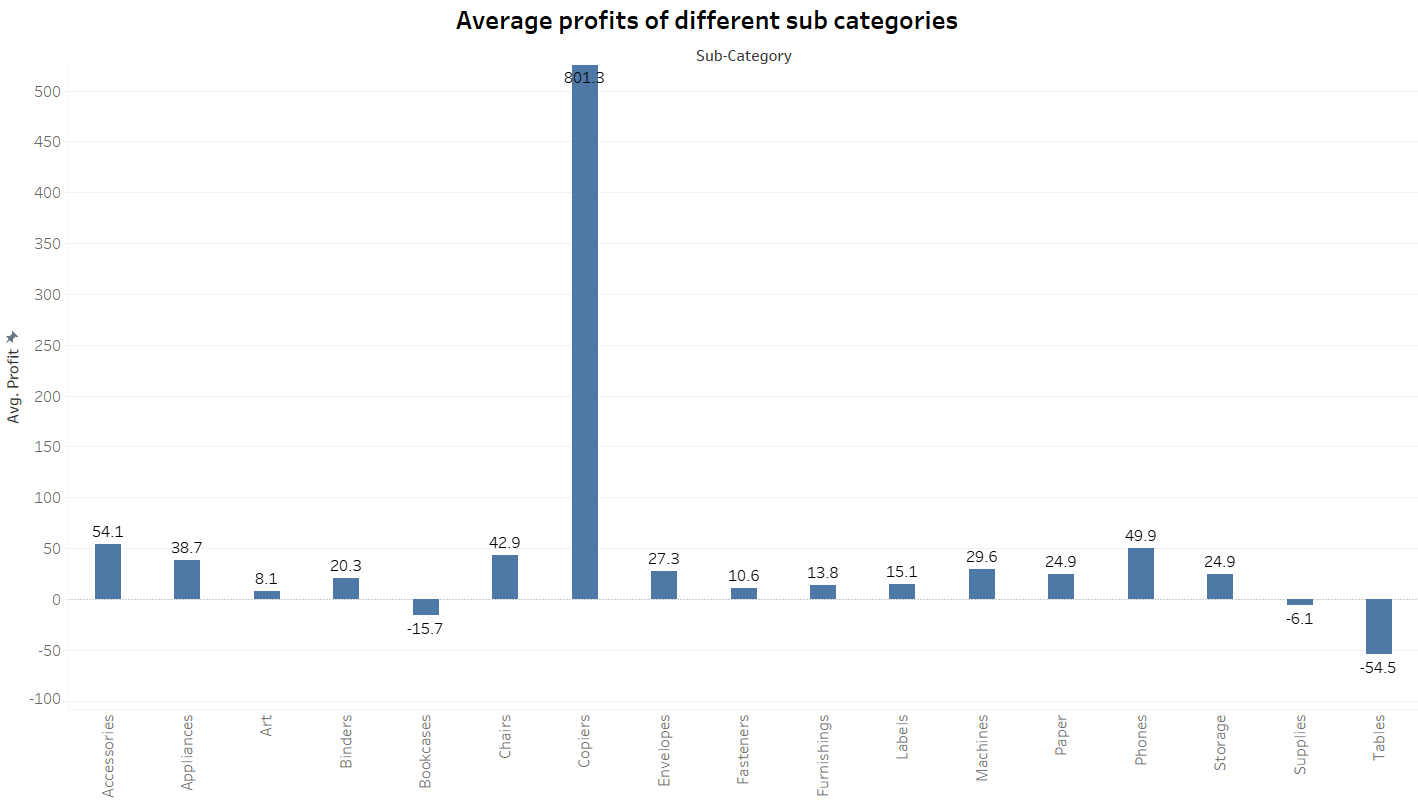
The above line chart paints the various order quantity of different product categories across the time period of the entire data set and since the data is continuous, line chart was used.

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



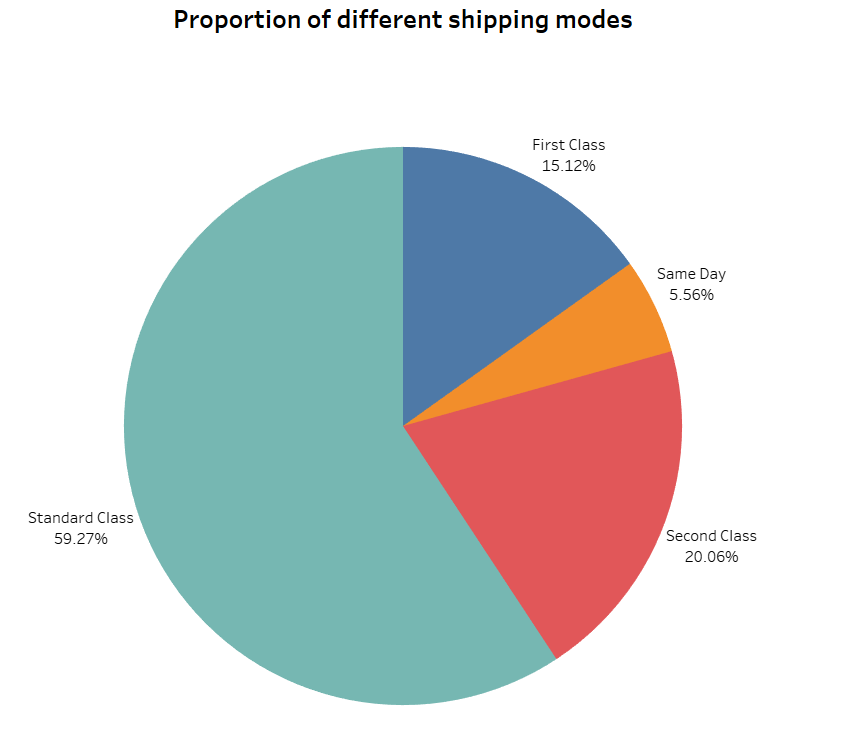
The above pie chart shows how the returns were distributed across different regions in the entire dataset, and since the proportions are considered, a pie chart would be most appropriate.

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



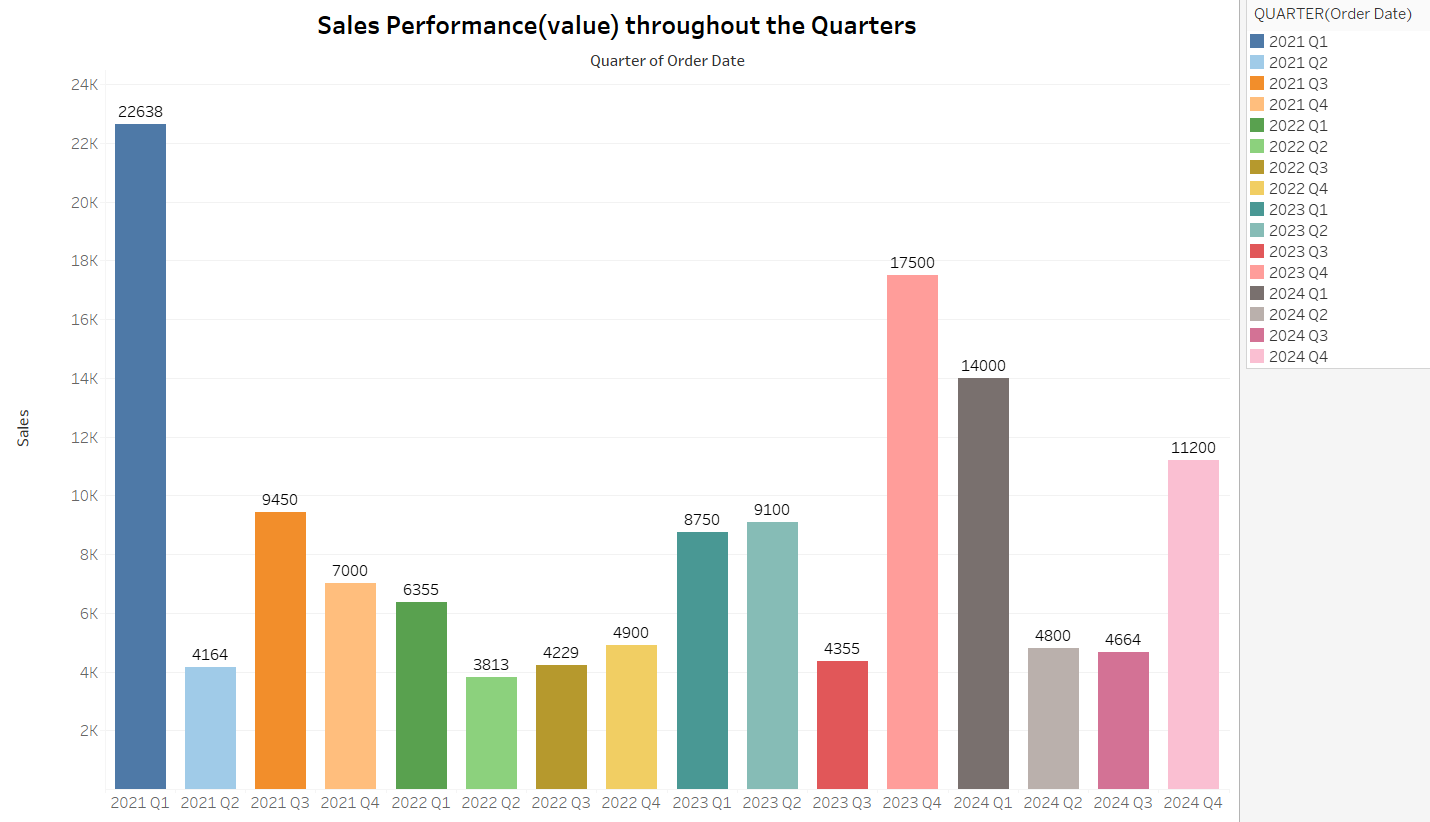
The above bar chart shows the profits generated by different sub categories of products and as the values are categorical, bar charts would be appropriate.

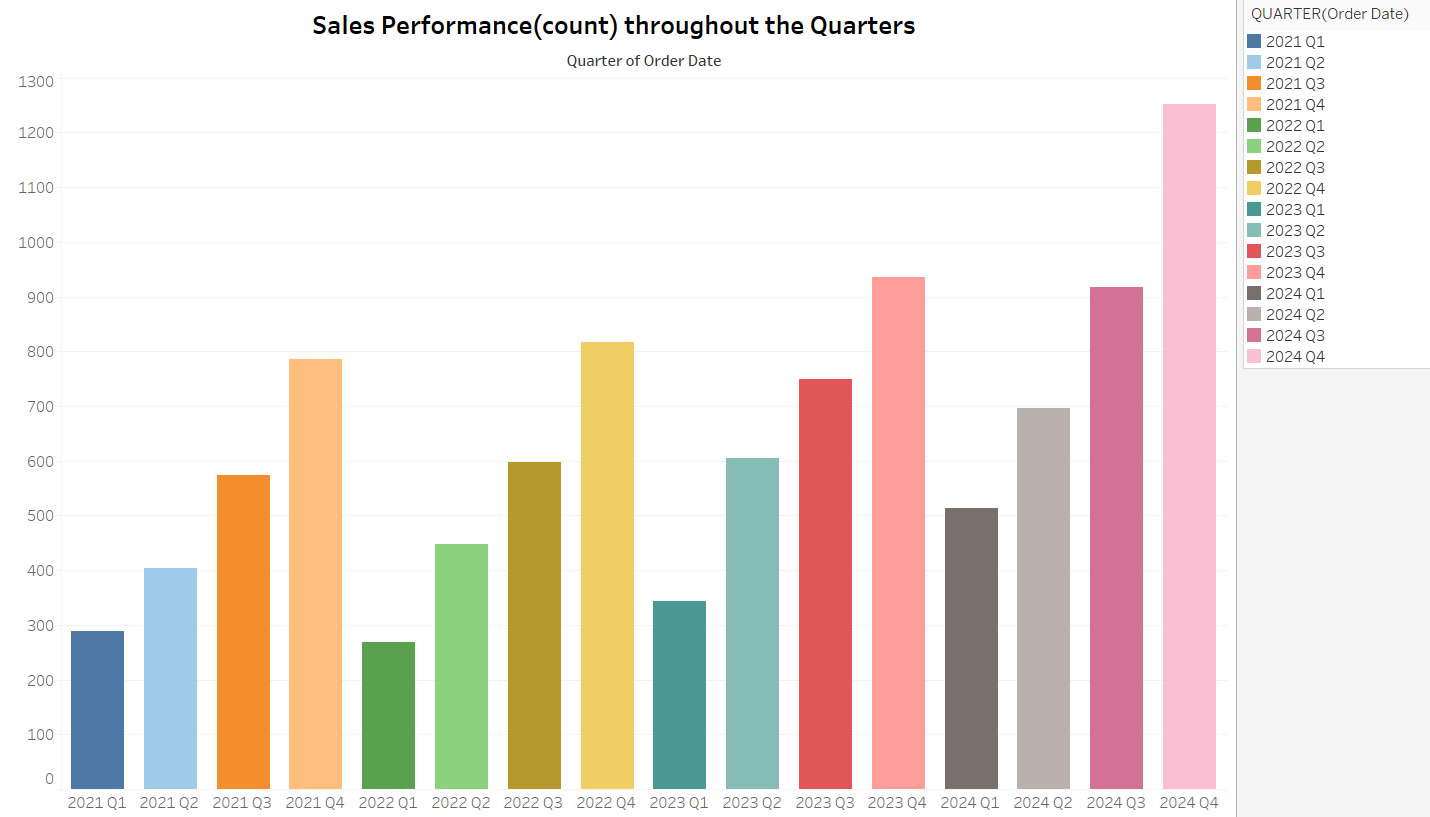
* Which shipping mode is the most commonly used in the Sample Superstore dataset?



The above pie chart successfully shows that standard class shipping mode is the most popular shipping mode in the entire dataset.

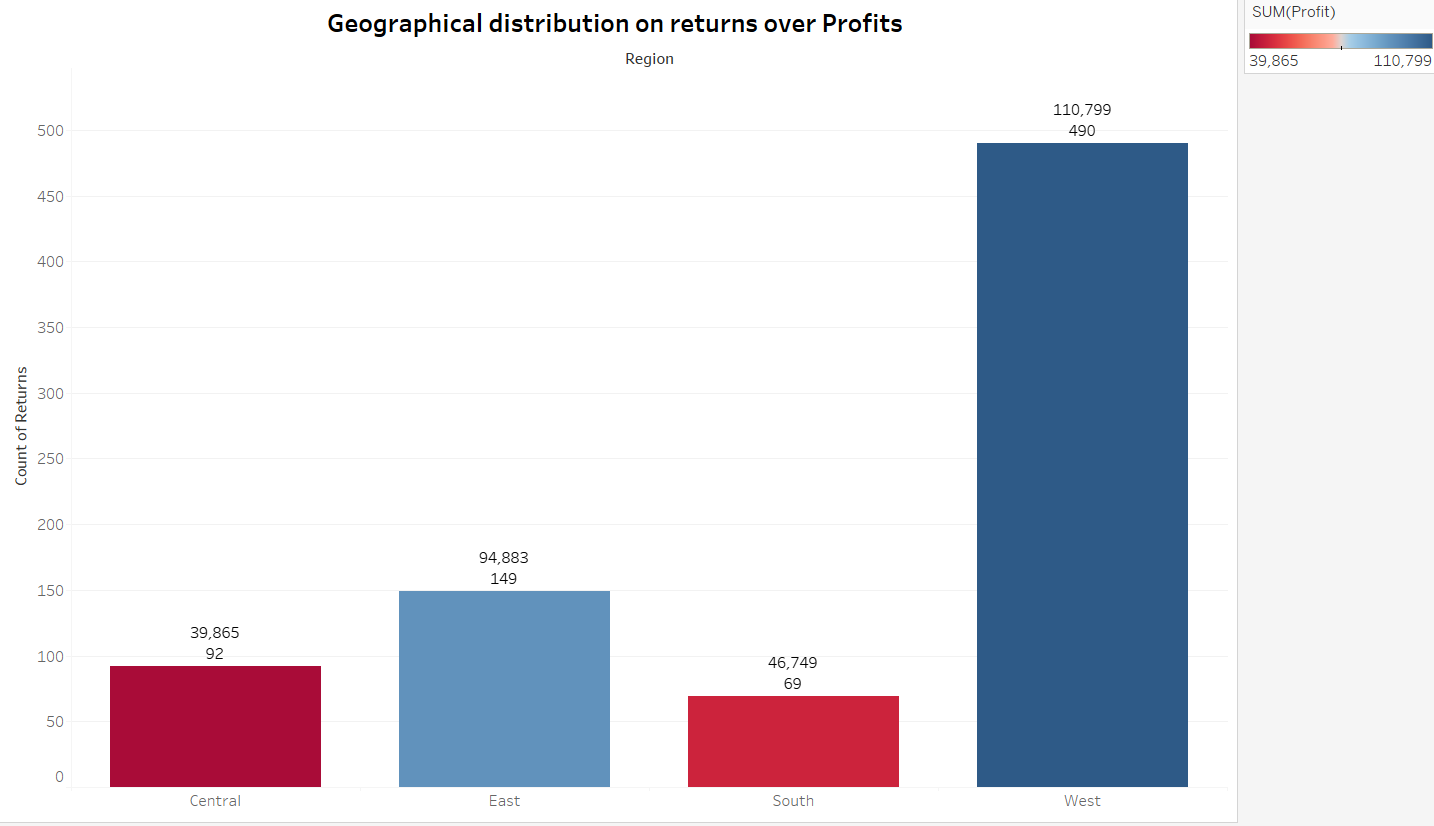
* How does the sales performance of different regions evolve throughout the quarters of a year?





In the above two bar charts, sales performance (both count and value sum) has been depicted wrt. different quarters of the years of data given in the dataset.

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



The above chart represents the returns across different regions and the color of the bars represent the profits made by those orders. Although the data and picture seems accurate, any kind of deemed results is inconclusive.