**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:**

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

**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!**

**Name – Shubham Shastri**

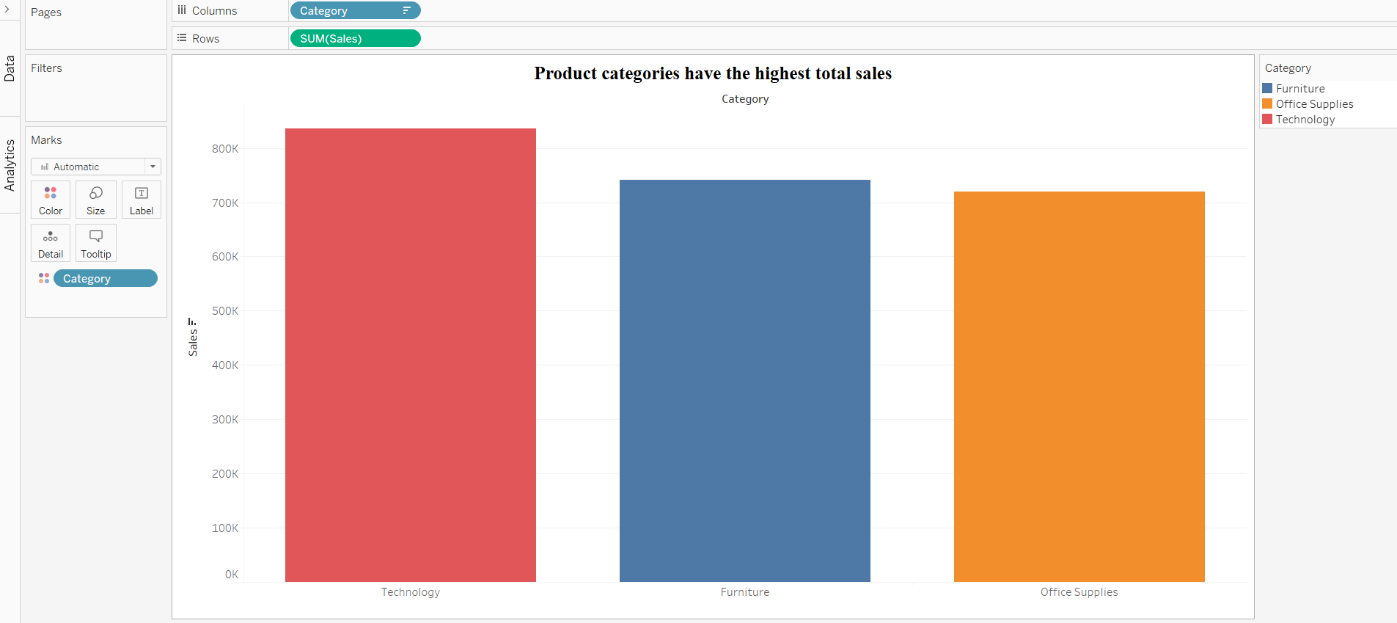
**Link of tableau public:** [Sample Superstore Data Q and A dashboard | Tableau Public](https://public.tableau.com/app/profile/shubham.shastri3068/viz/SampleSuperstoreDataQandAdashboard/Dashboard1?publish=yes)

**Link of Github:** [sdshastri/Sample-Superstore-Data-Q-A (github.com)](https://github.com/sdshastri/Sample-Superstore-Data-Q-A)

**Questions:**

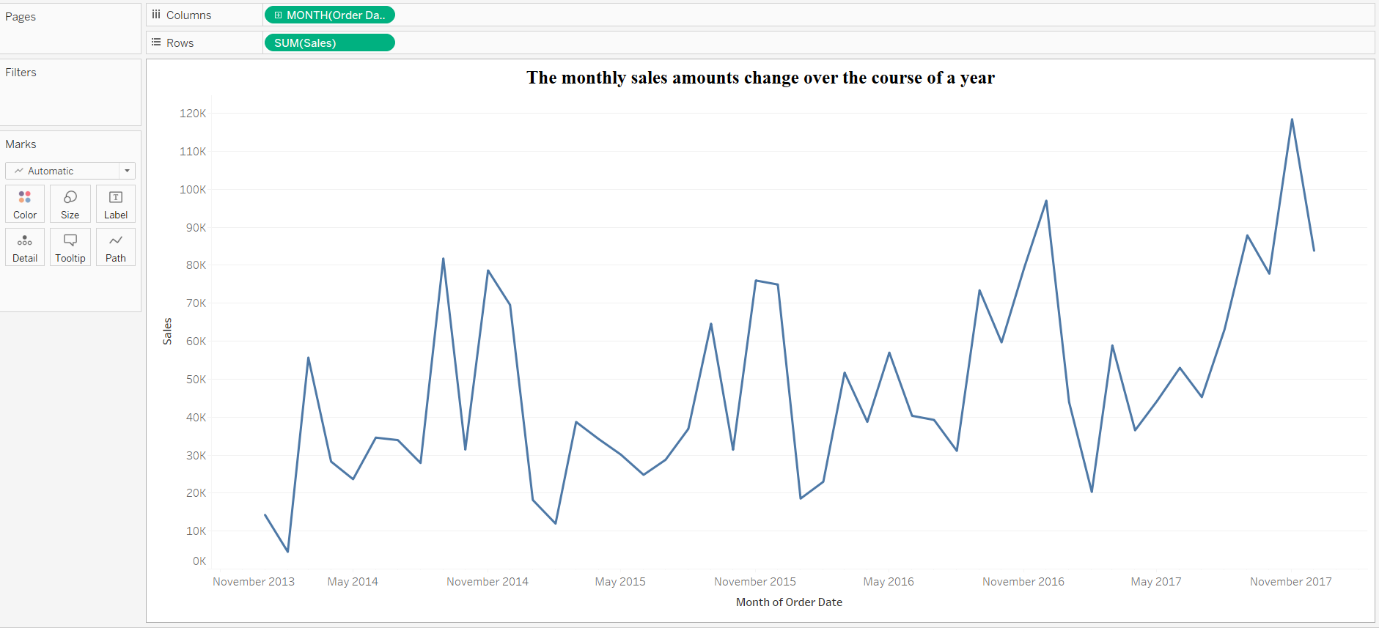
1. Which product categories have the highest total sales in the "Superstore" dataset?

* The Vertical bar chart represents product categories along with the x-axis. The total sales are represented by the y-axis. The size of the bar gives us which product categories have the highest sales.



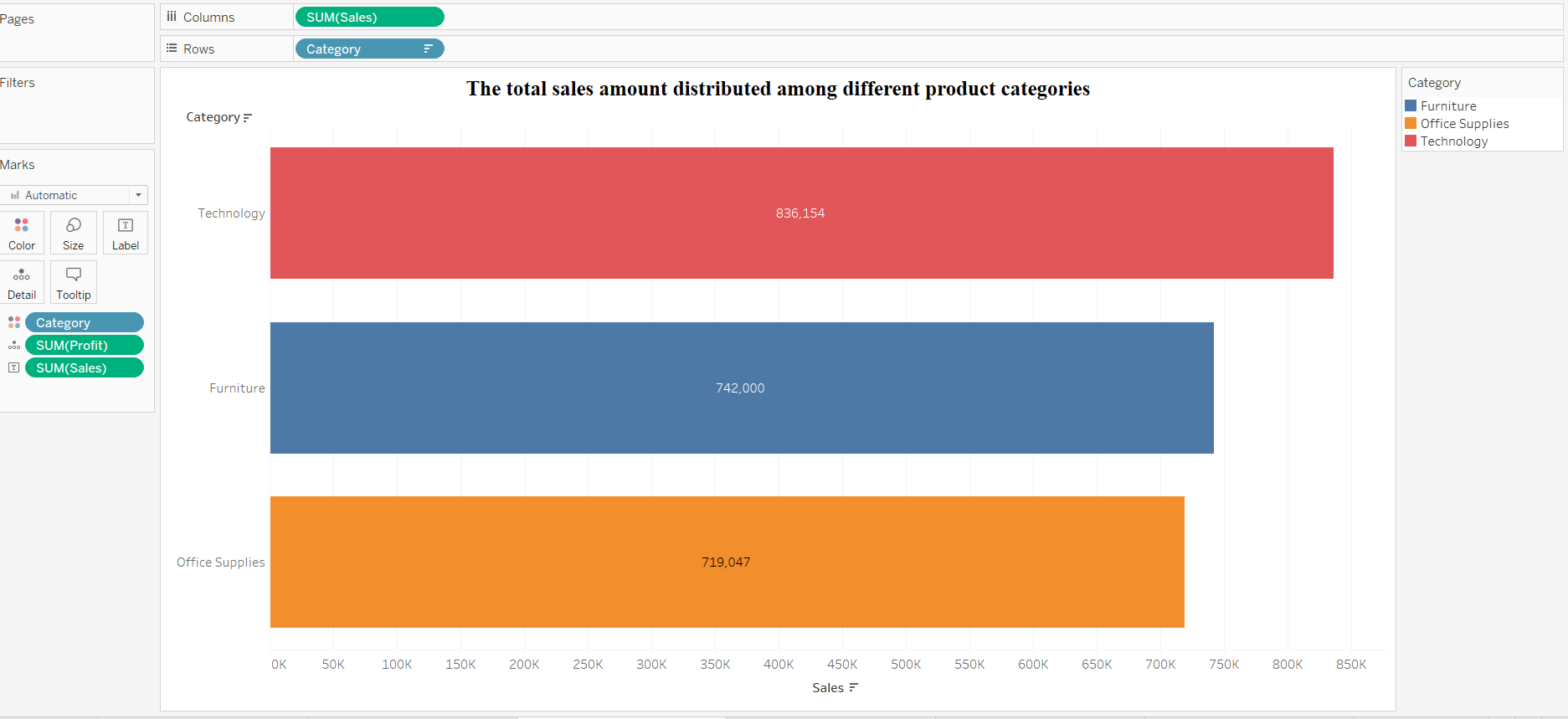
1. How do the monthly sales amounts change over the course of a year?

* The line chart is best used to showcase any trend. The X-axis is denoted by months while the sales amount is put on the Y-axis. We can easily visualize the change of sales over time.



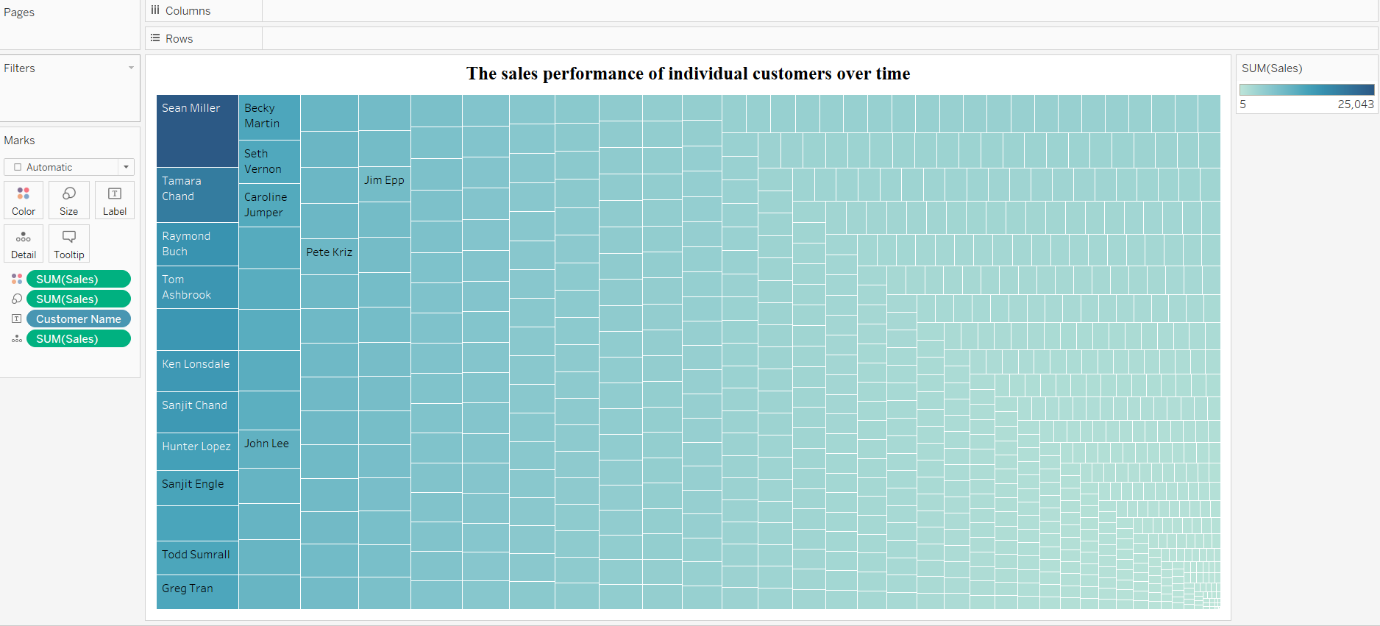
1. How is the total sales amount distributed among different product categories?

* The horizontal bar chart is used for the distribution of total sales among different product categories. The product categories are denoted on the y-axis while the total sales are on the x-axis. The bar shows the distribution of sales.



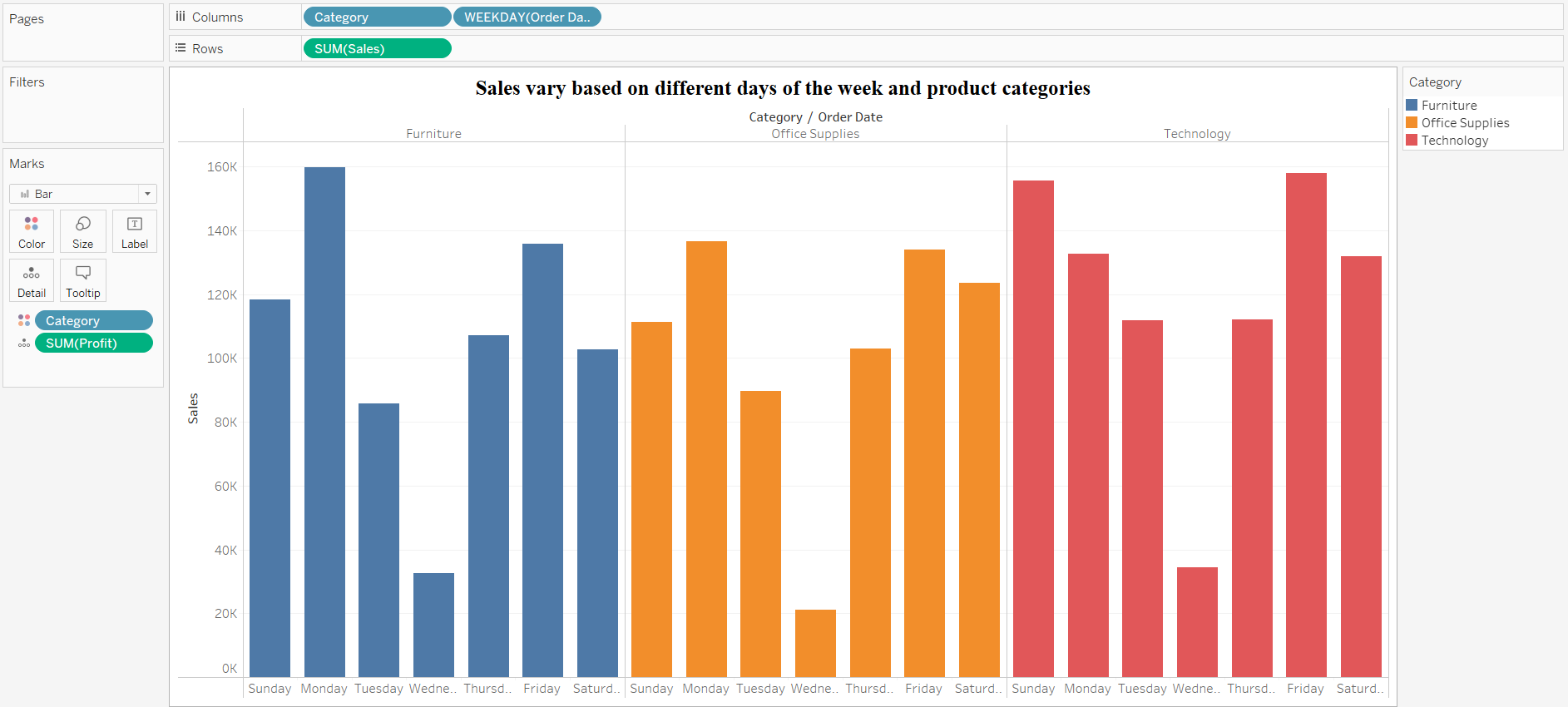
1. Can we analyse the sales performance of individual customers over time?

* The tree chart is best when we want to show the distribution of the large size of the category. The individual customers have a large category column. Sales are represented by the size of the rectangle. We can see details of any customer by hovering over the corresponding rectangle.



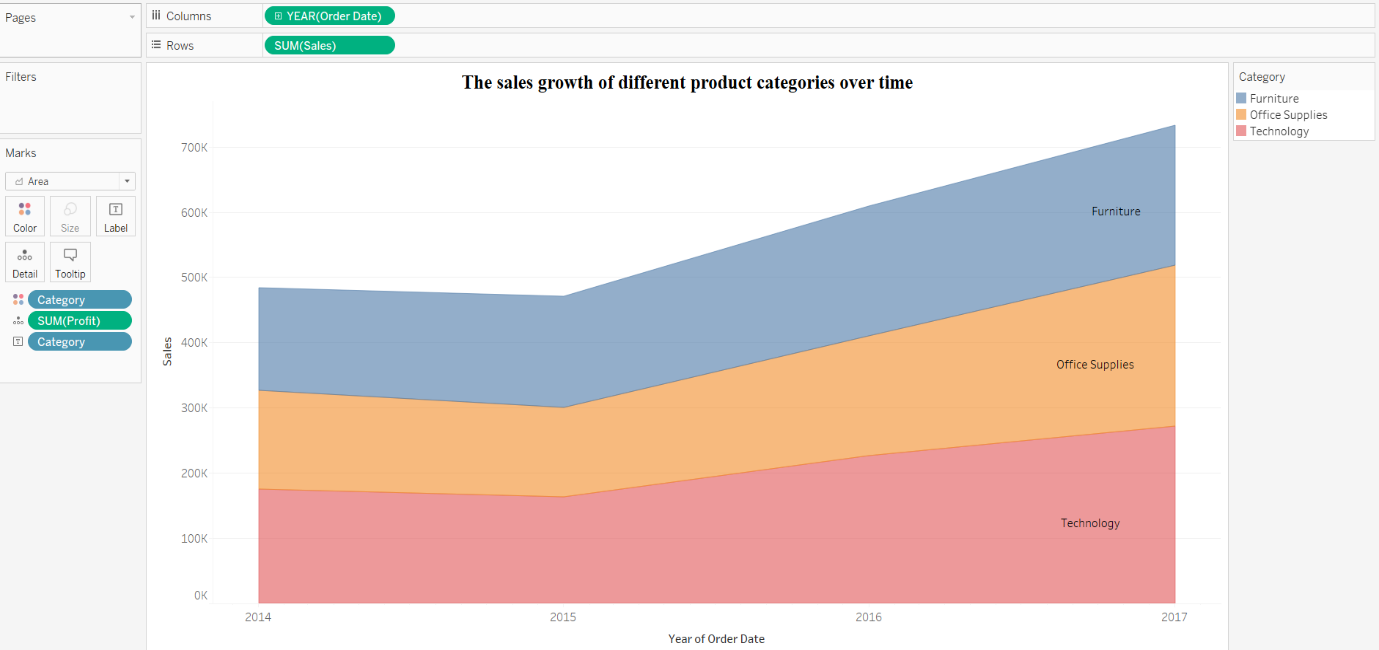
1. How do sales vary based on different days of the week and product categories?

* The column chart is selected for the given question. We divide the column chart into three based on product categories. The x-axis represents the day of the week while the y-axis represents the sales. All the charts is side by side with the same y-axis, so we can compare easily.



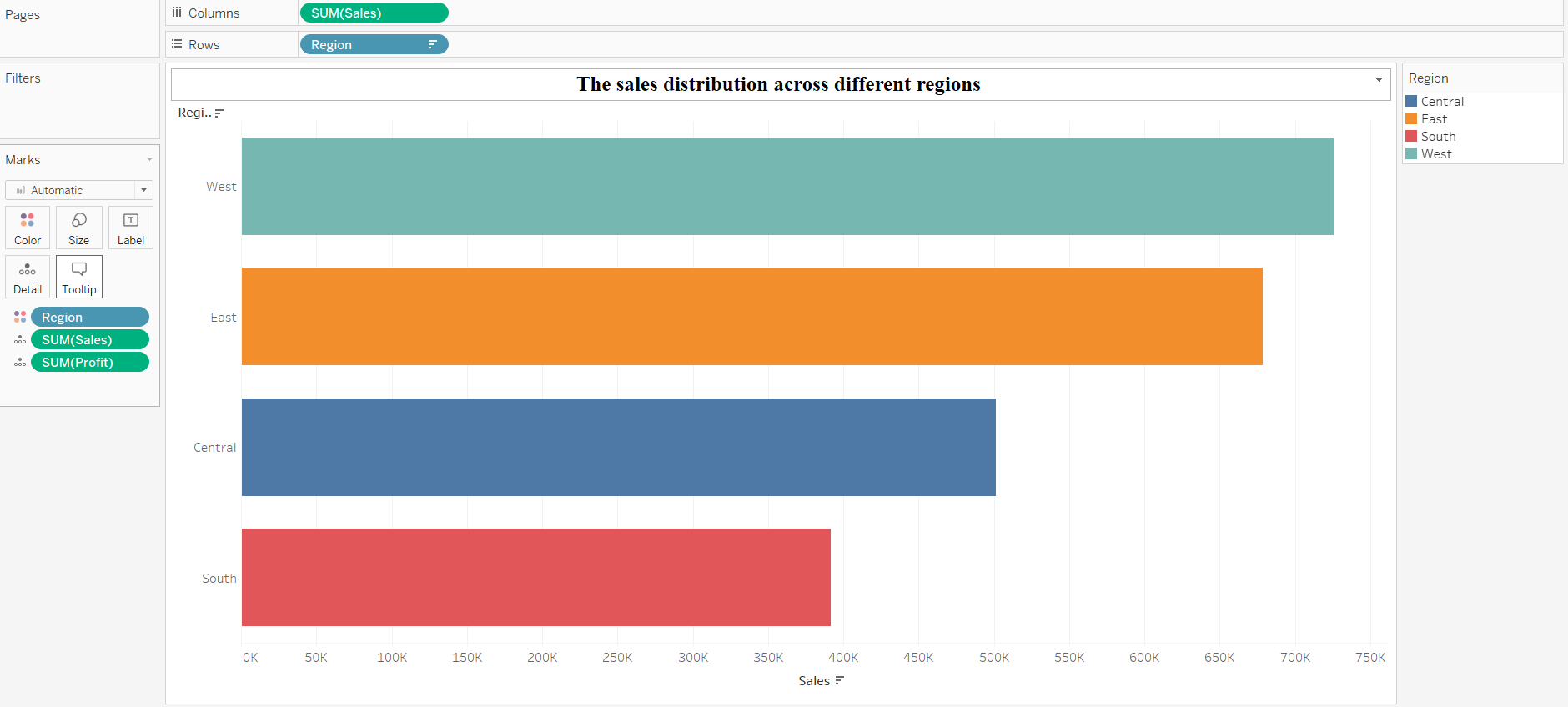
1. Can we visualise the sales growth of different product categories over time?

* The area graph is selected for that question. The different area or colour represents the product categories. The time is represented by the x-axis while the sales are represented by the y-axis.



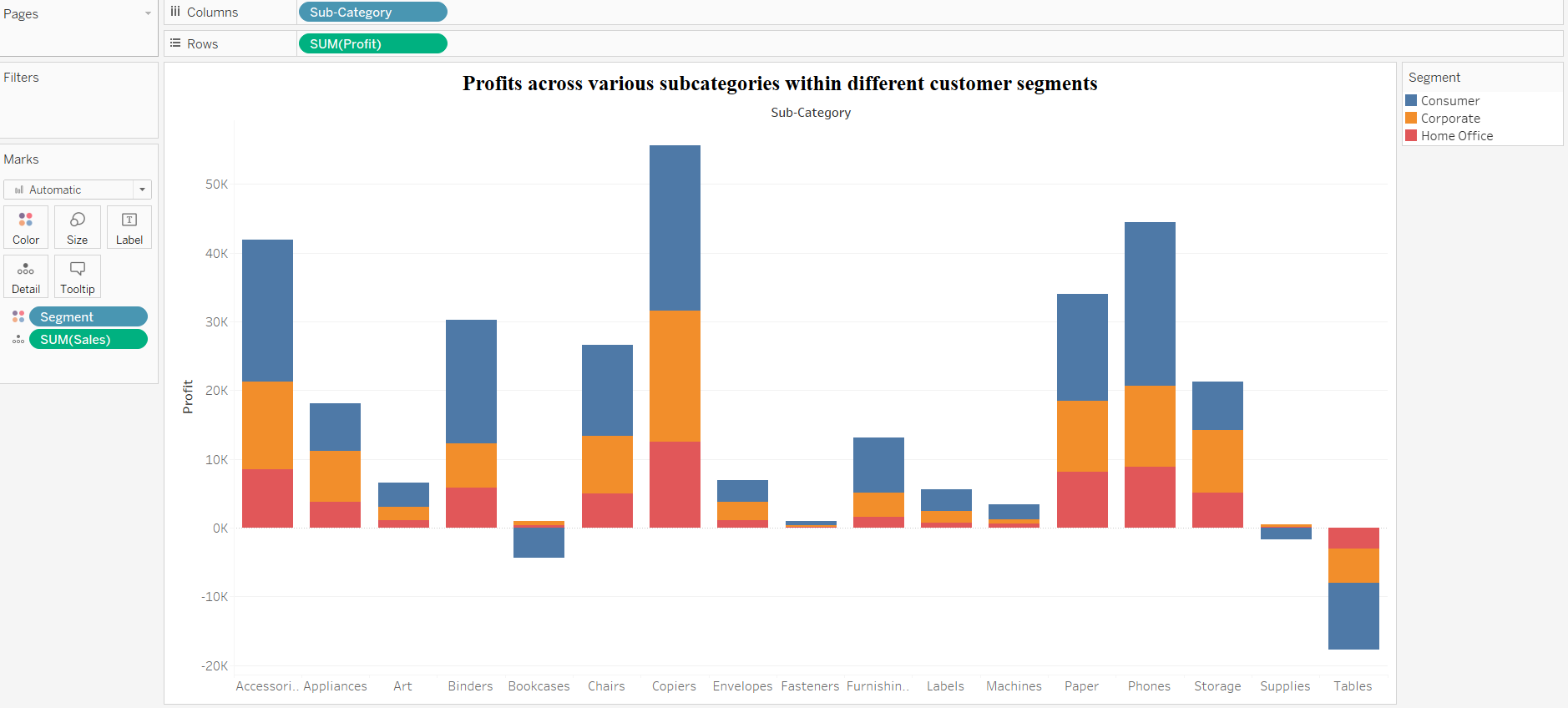
1. How does the sales distribution vary across different regions in the "Superstore" dataset?

* The bar is best to represent the categorical vs numerical graph. Here we denoted the y-axis by the region while the sales denoted by the x-axis. We can visualize the distribution by just the size of bar.



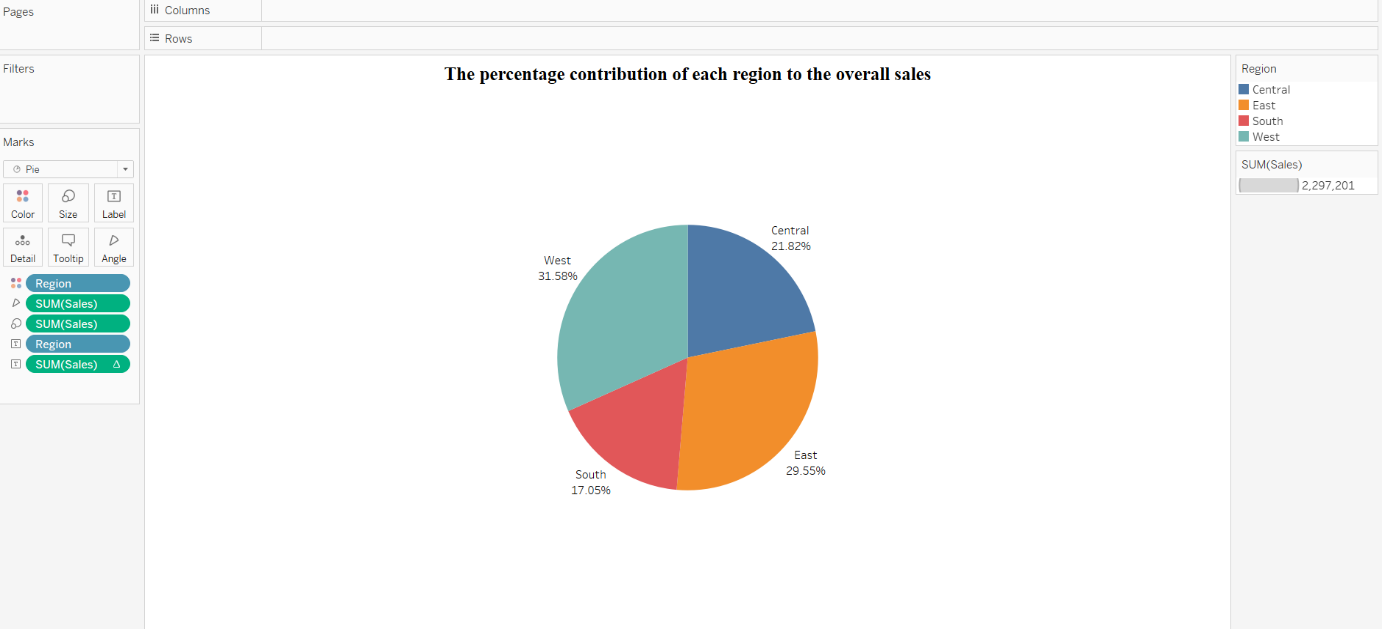
1. Can we visualise the composition of profits across various subcategories within different customer segments?

* The column chart represents the profits across the sub-categories. The column is divided into three customer segments. The y-axis represent the profits while the x-axis represents the sub-categories.



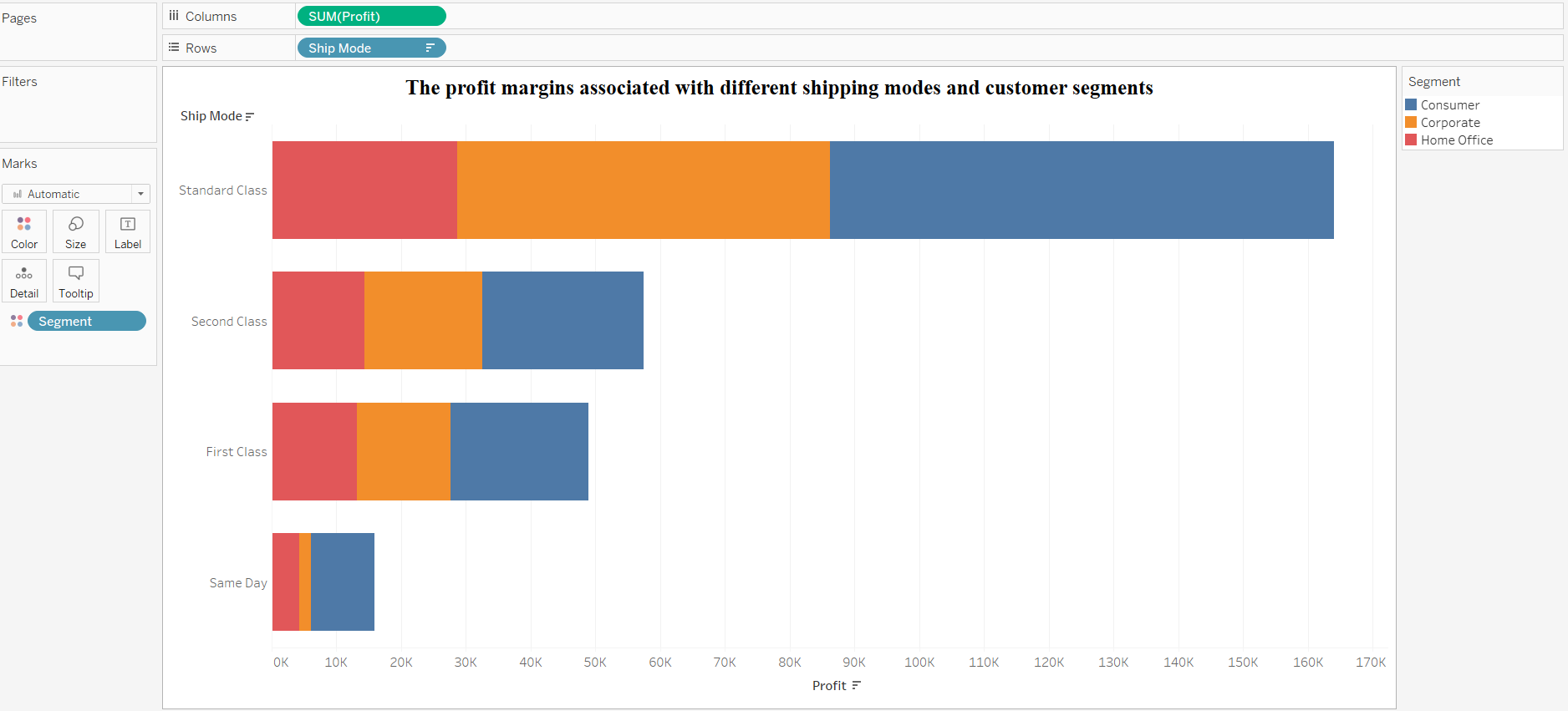
1. What is the percentage contribution of each region to the overall sales?

* The pie chart represents the percentage of sales for each region. Each pie tells us how much the particular region contributes to the overall sales.



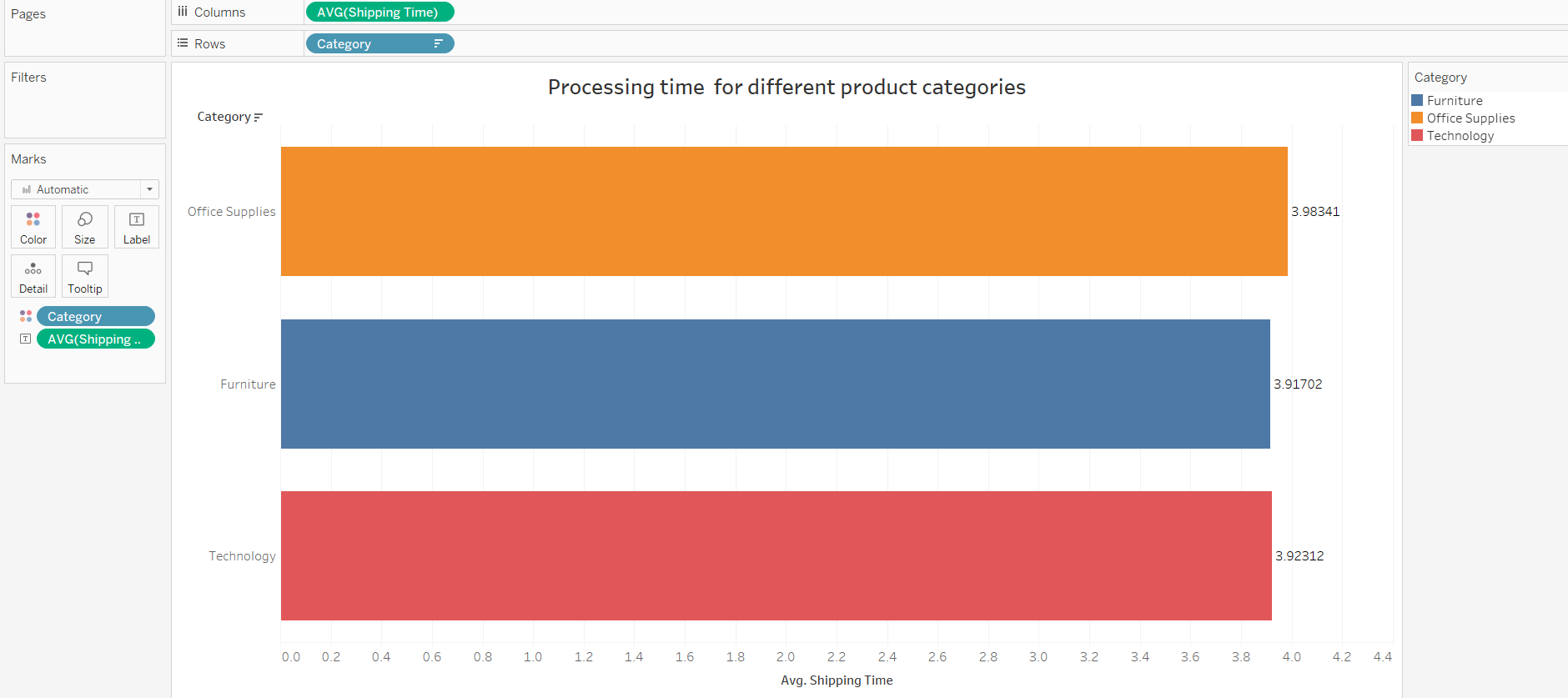
1. Can we visualise the profit margins associated with different shipping modes and customer segments?

* In the bar chart, the bar represents the profit by different ship mode. The bar is divided by the colour based on the customer segments. The X-axis represents the profit while the y-axis represents the ship mode.



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

* The bar chart represents the average processing time (days) for the different product categories. The x-axis represents the average shipping time while the y-axis represents the product category.



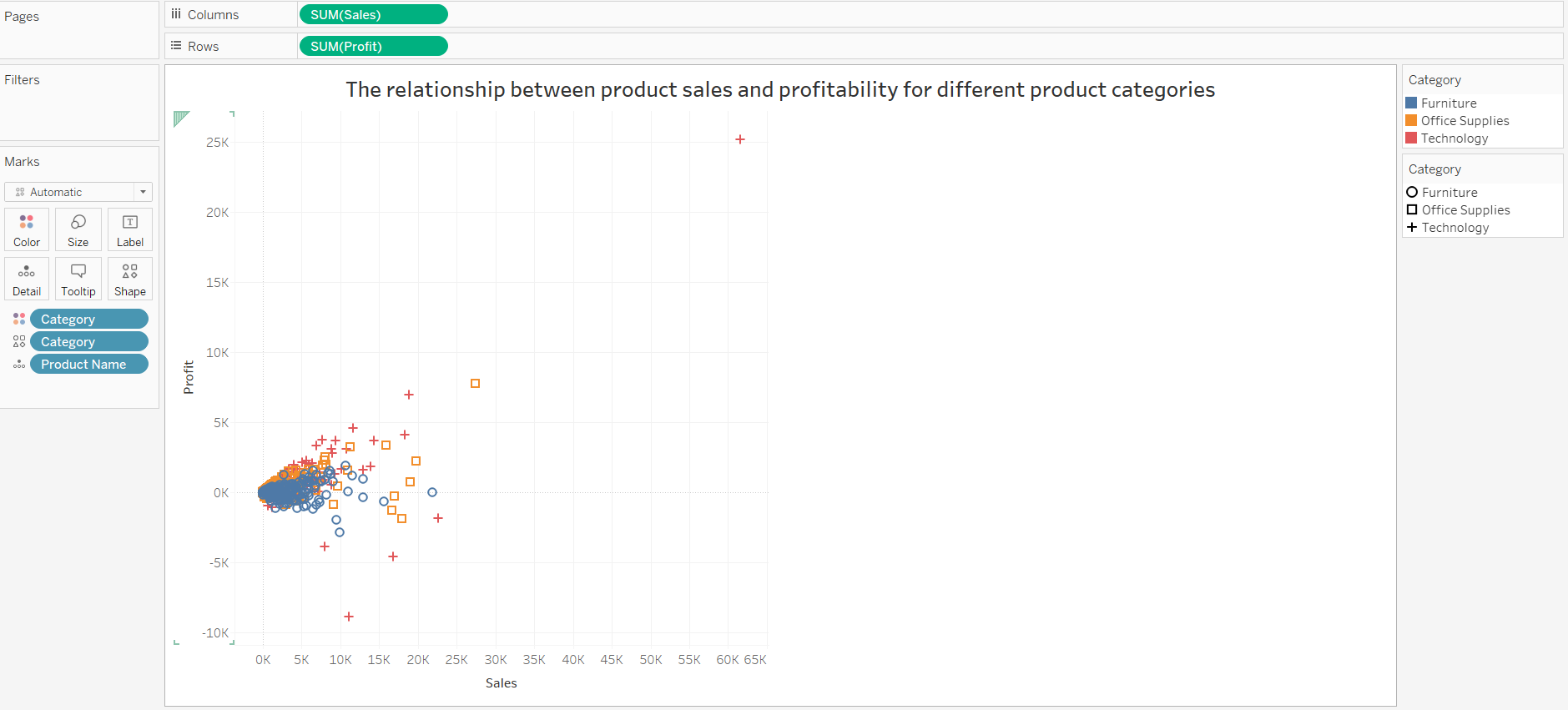
1. How does the performance of different salespeople compare in terms of sales, quantity, and profitability?

* We combine a 3-column chart of sales, quantity, and profitability. On the y-axis first is sales, second is quantity, and last is profit. On the x-axis the name of the salespeople.



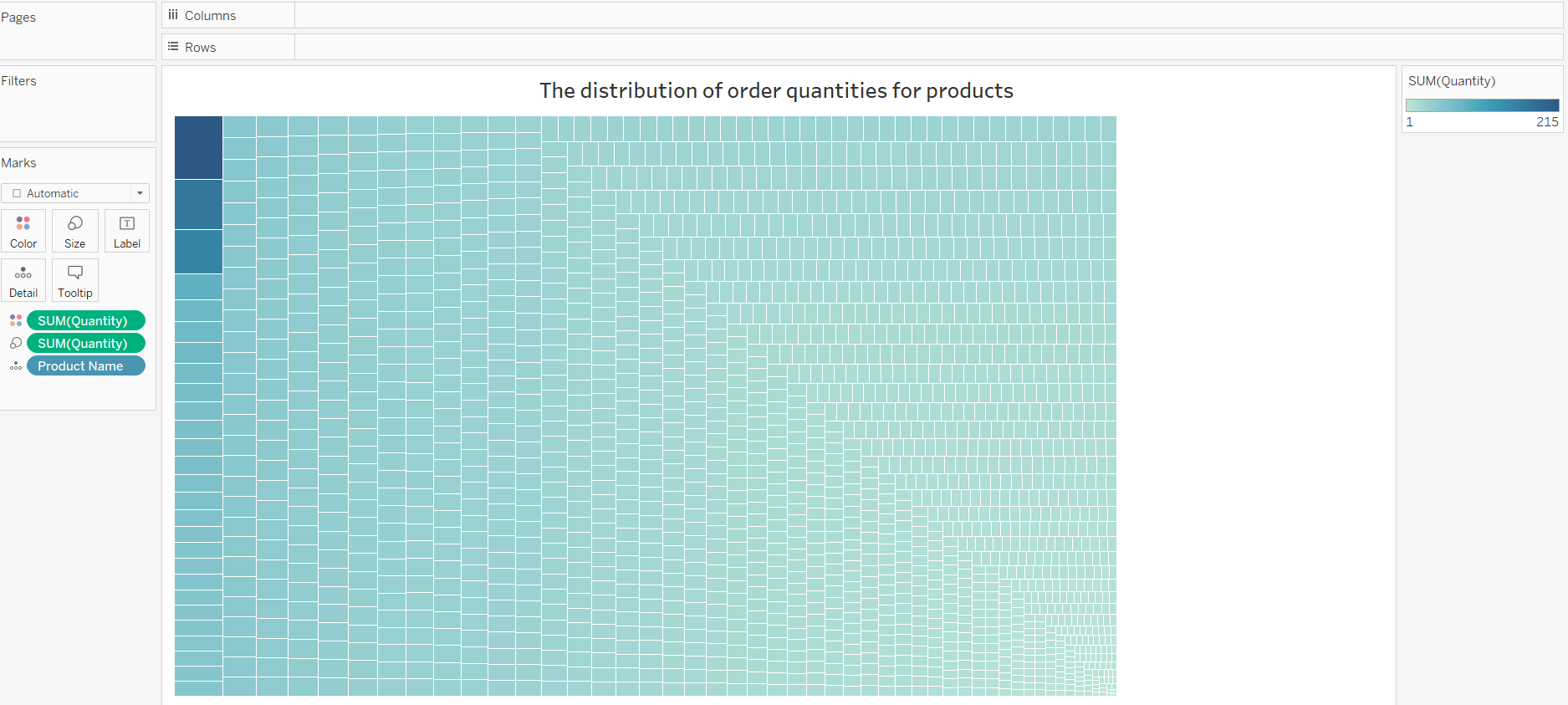
1. Can we visualise the relationship between product sales and profitability for different product categories?

* The scatter plot is best known for the visualisation of 2 numerical values. Here we visualize the product sales and profit.



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

* If the number of categories is large we use three charts to show distribution. The size of the rectangle represents the order quantities. Each rectangle represents the product name.



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

* The box plot shows the distribution of the overall profit. The product categories are represented by the colour or hover over the chart we can see the details.



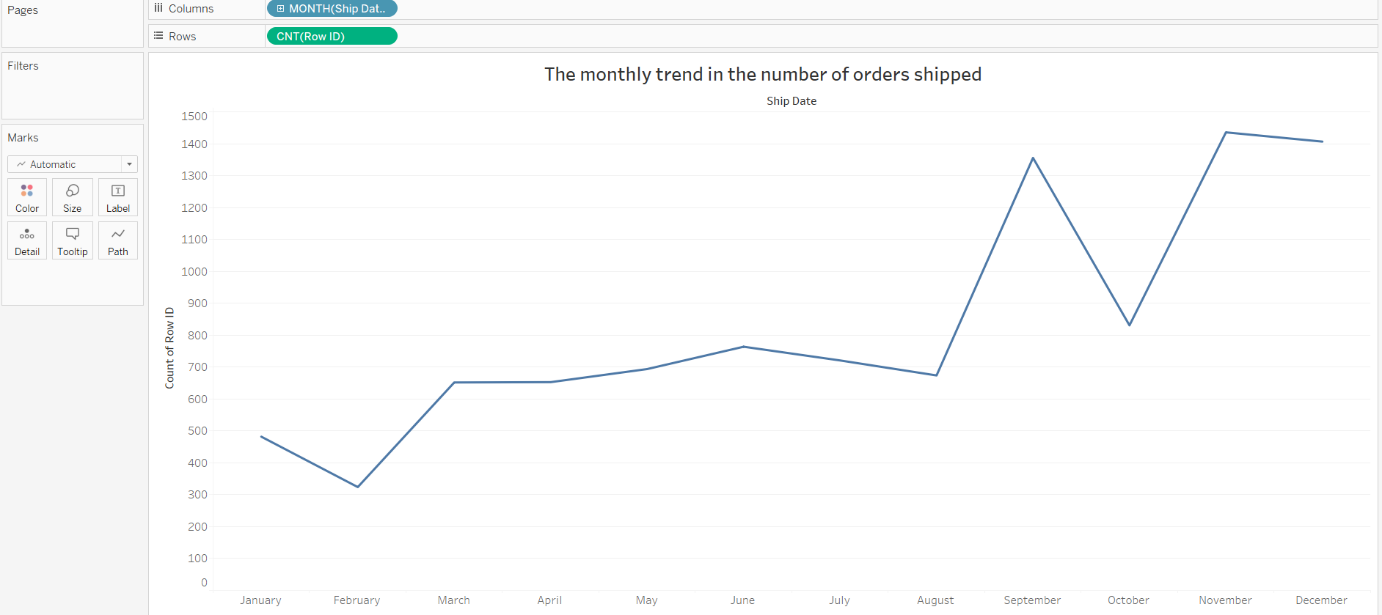
1. Can we compare the shipping time distributions for different shipping modes?

* The bar chart shows the shipping time for the different shipping modes. The x-axis of the bar chart denotes the average shipping time while the y-axis denotes the ship mode.



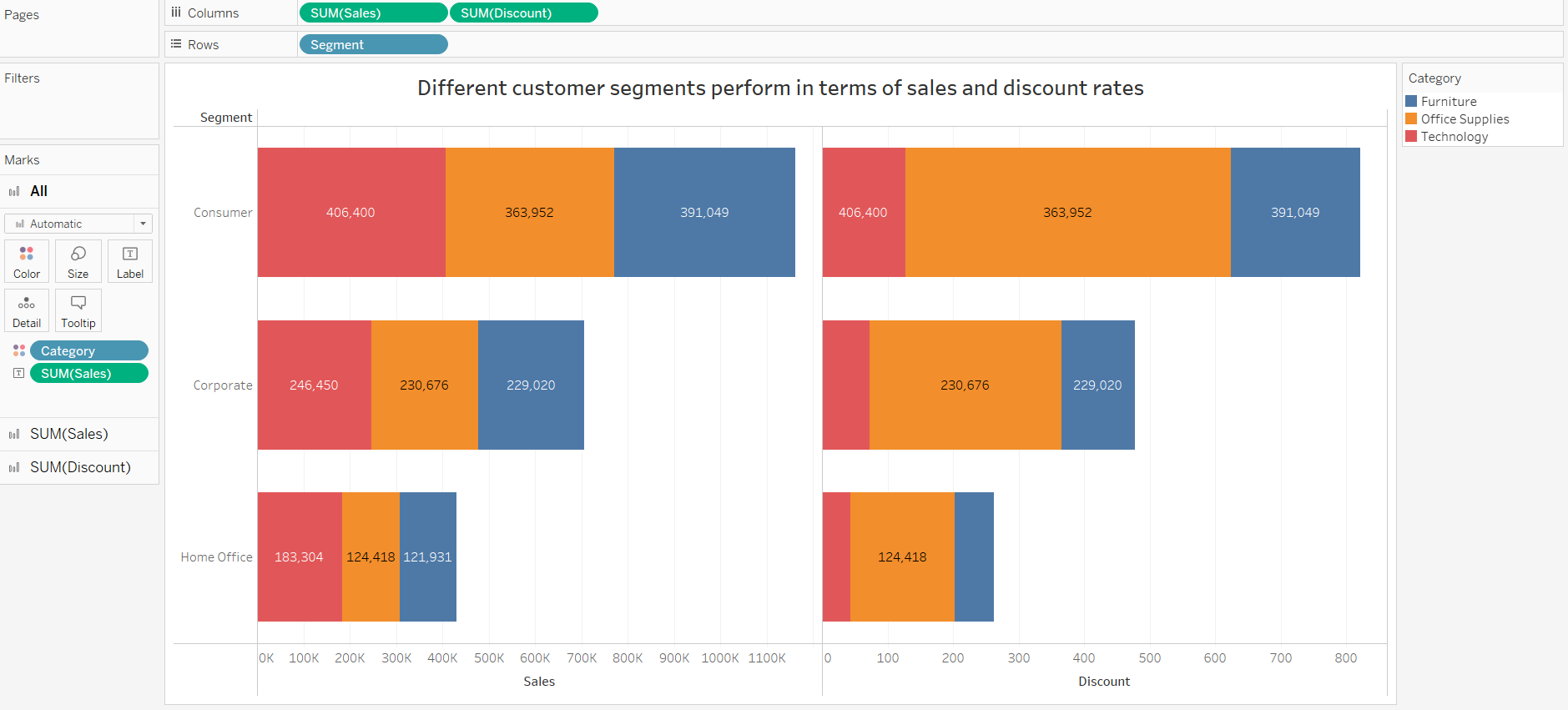
1. What is the monthly trend in the number of orders shipped?

* The line chart is best to show the trends over time. Here, we show the trend of the number of orders shipped over the month. The x-axis shows the month while the y-axis shows the number of orders shipped.



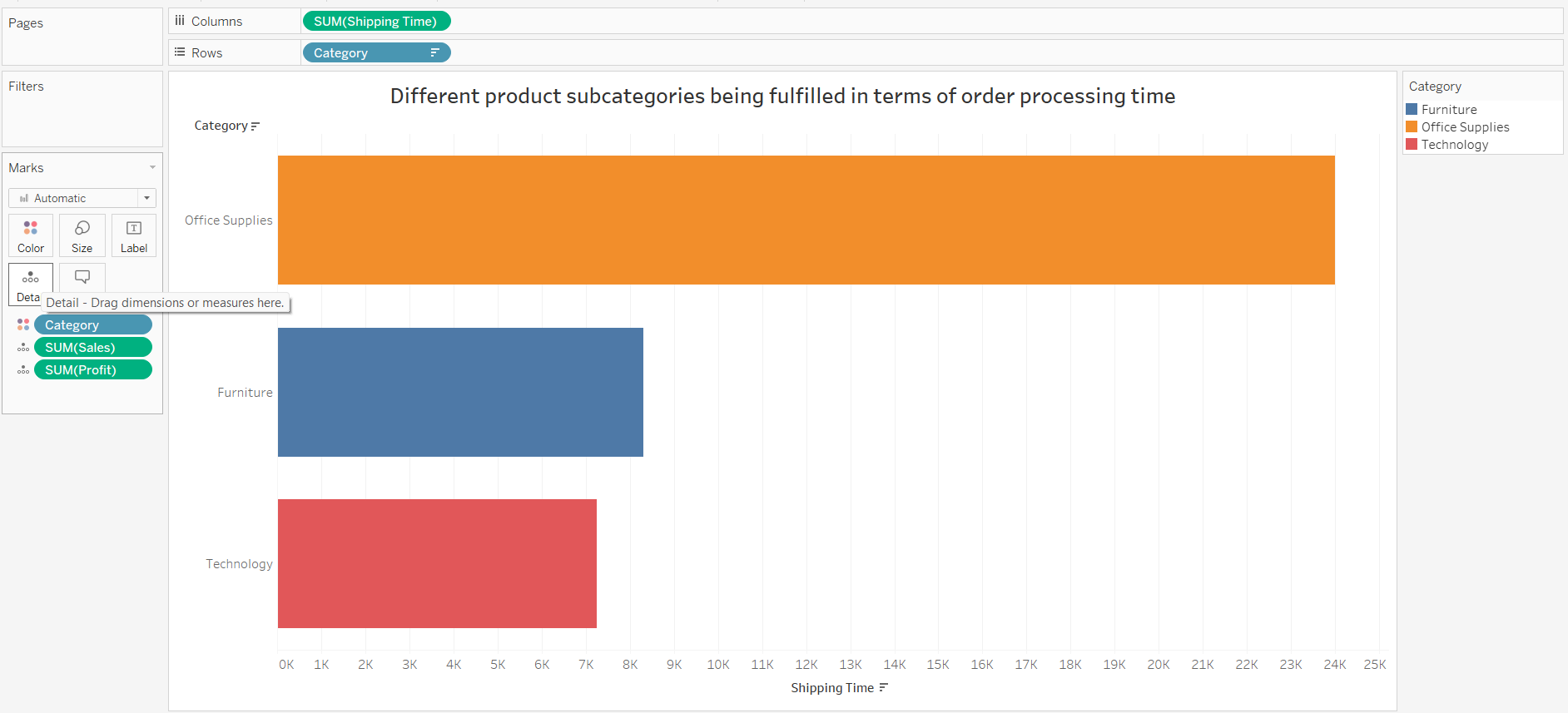
1. How do different customer segments perform in terms of sales and discount rates?

* We select a 2-bar chart for the particular question. The first bar chart shows the sales distribution over different segments. The second bar chart shows the discount over segments. The bars are divided based on the product category.



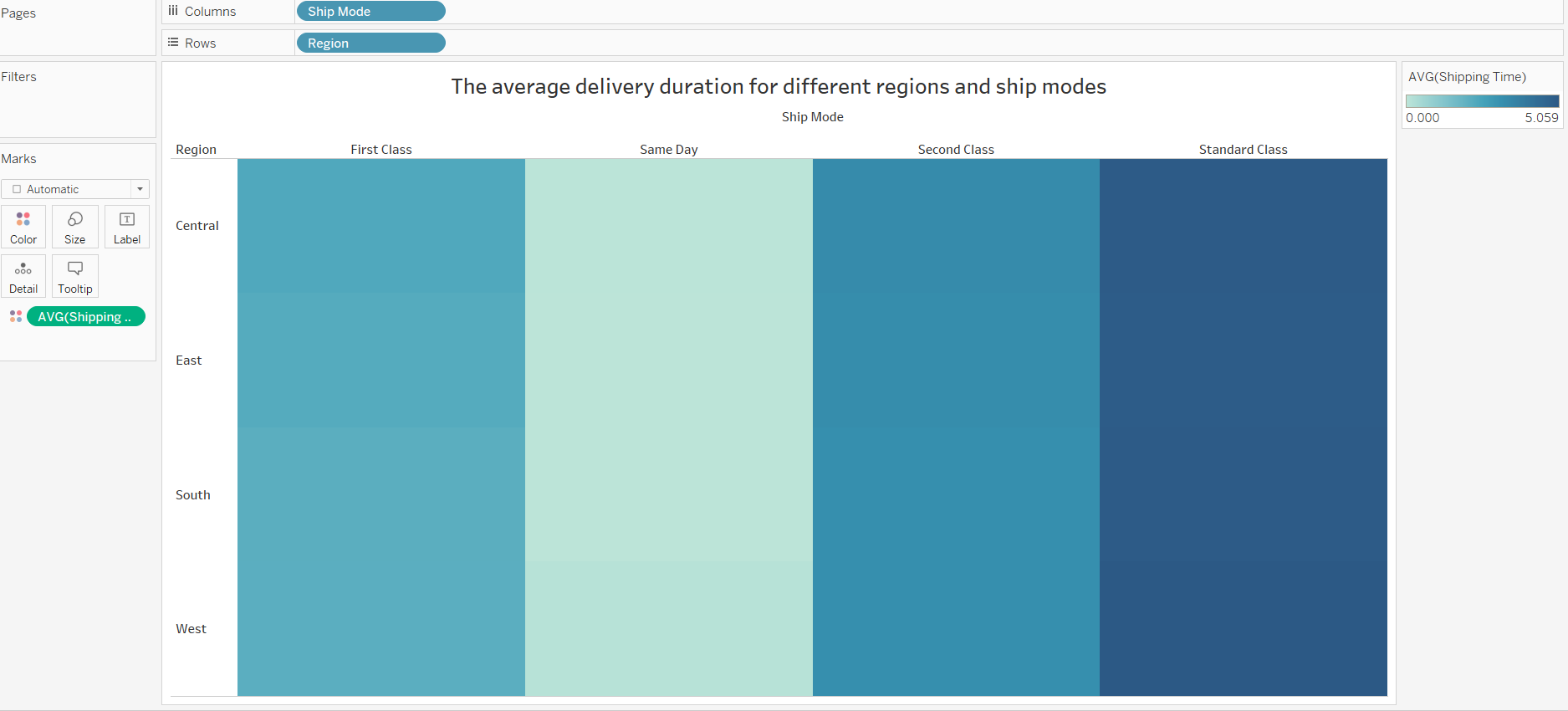
1. How efficiently are different product subcategories being fulfilled in terms of order processing time?

* The bars of the bar chart show the shipping time for different product categories. The x-axis shows the product category while the y-axis shows the shipping time.



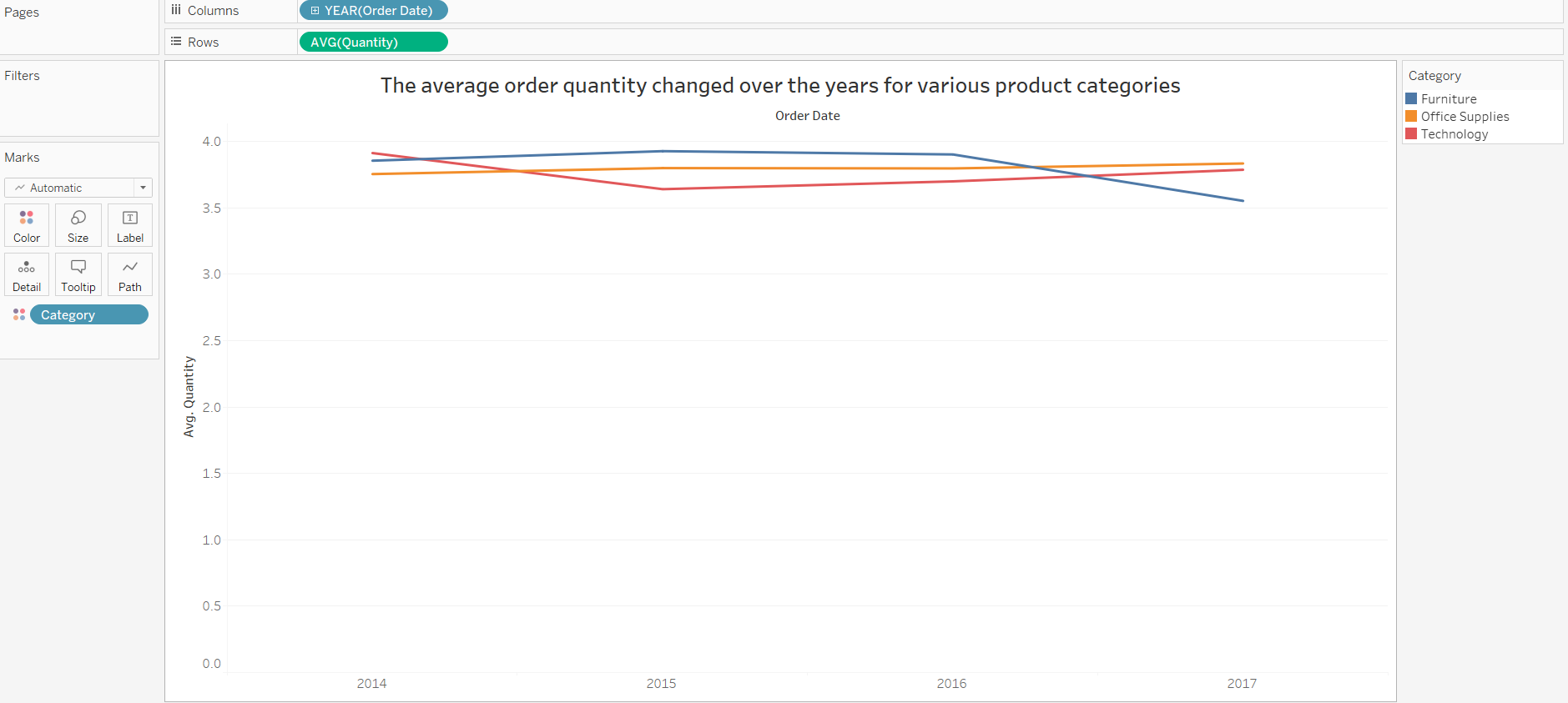
1. What is the average delivery duration for different regions and ship modes?

* We select a matrix with rows with regions and columns with ship mode. The color or hover over the matrix shows the average shipping time.



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

* The average order quantity changed over the years for product categories represented by the line chart. There are 3 lines for each category. The x-axis and y-axis show the time and quantity respectively.



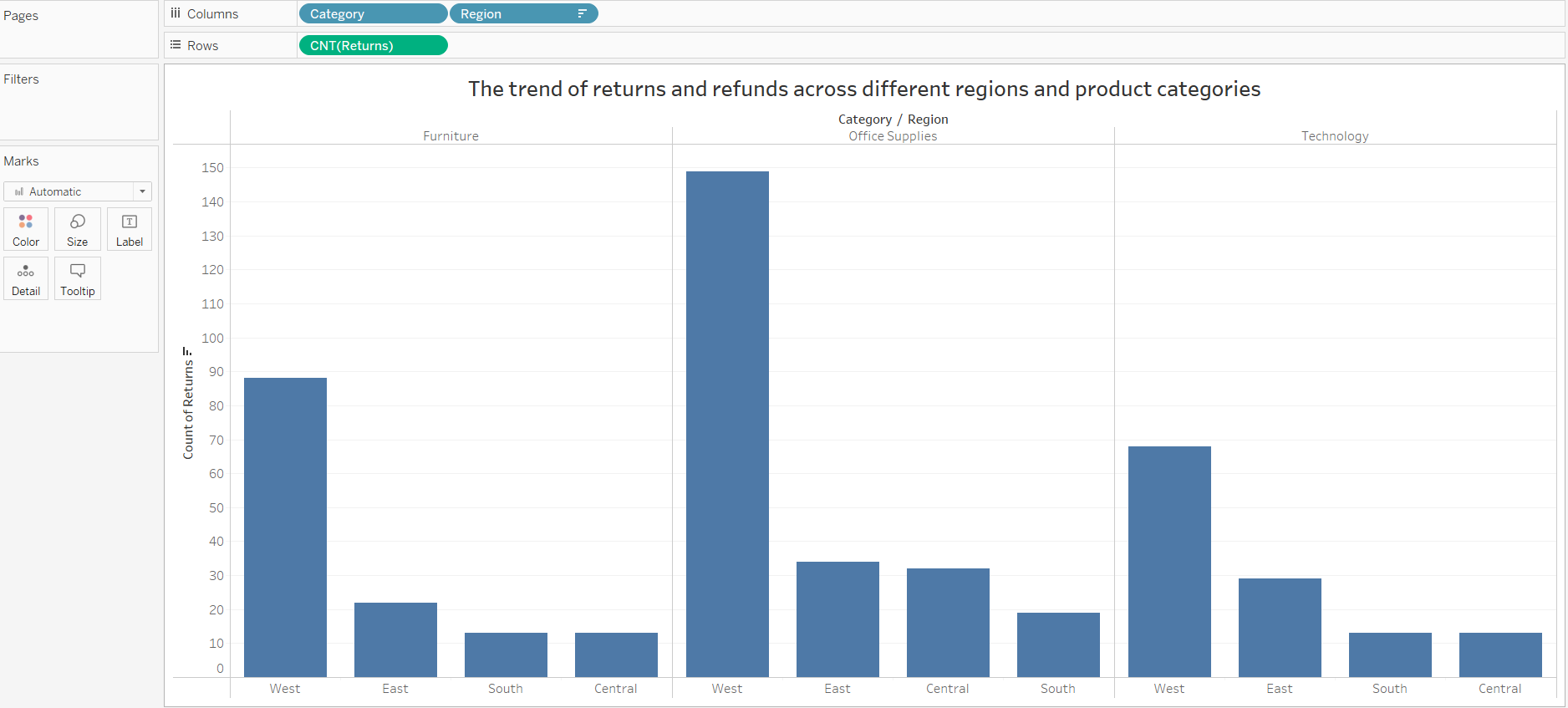
1. Can we visualise the correlation between discount rates and order quantities for different customer segments?

* The scatter plot shows the positive correlation between discount rates and order quantities for different customer segments.



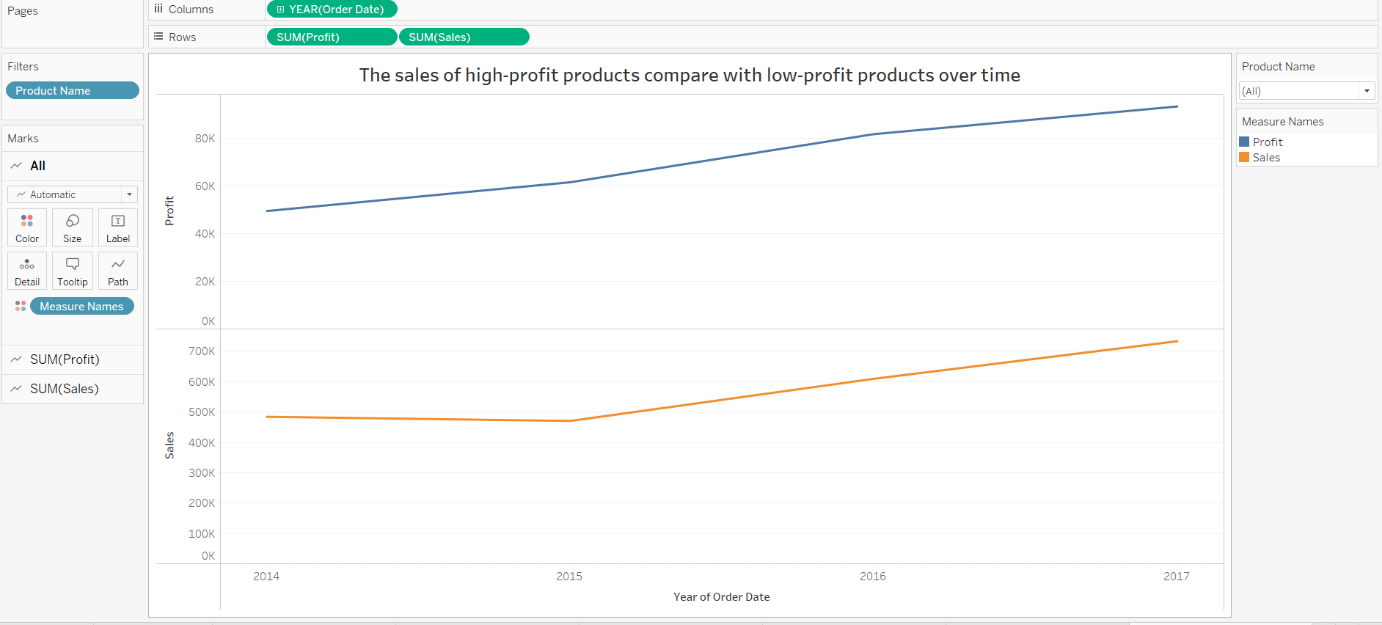
1. What is the trend of returns and refunds across different regions and product categories?

* The column chart shows the number of returns for the region. There are 3 column charts for each category.



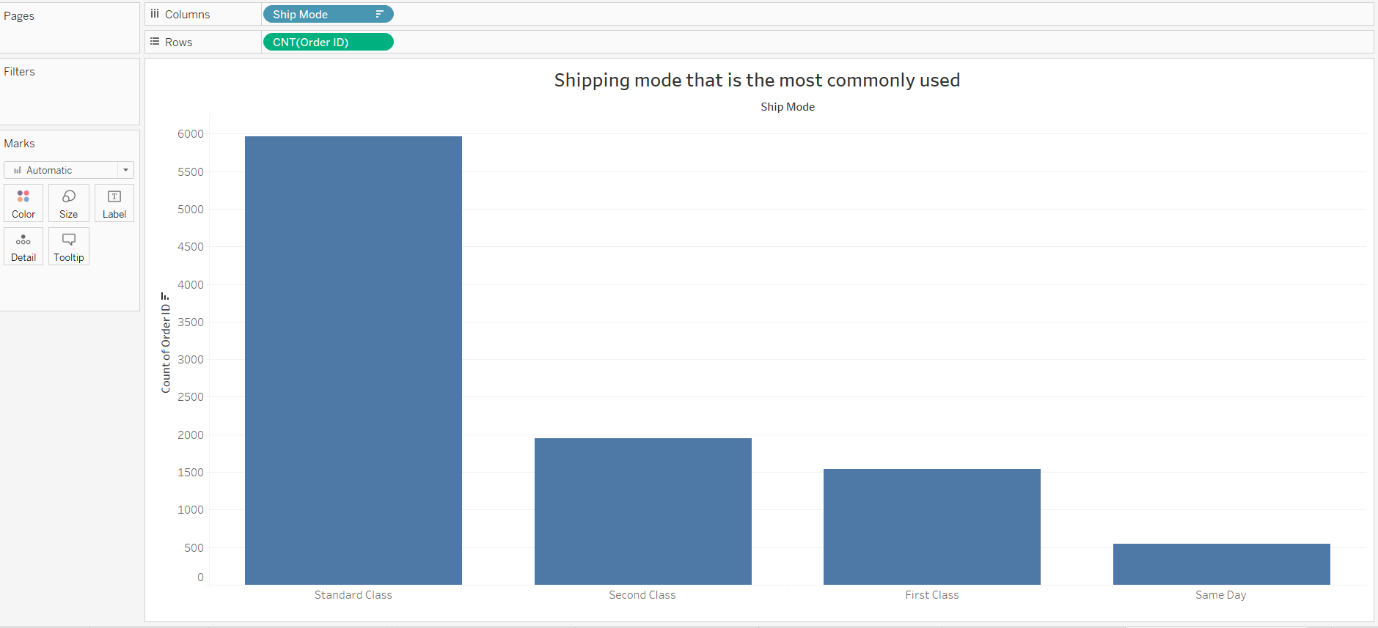
1. How do the sales of high-profit products compare with low-profit products over time?

* The line chart shows the profit and sales of products. The x-axis shows the year. The up y-axis shows the profit while the down y-axis shows the sales. The slicer is used for product names to filter.



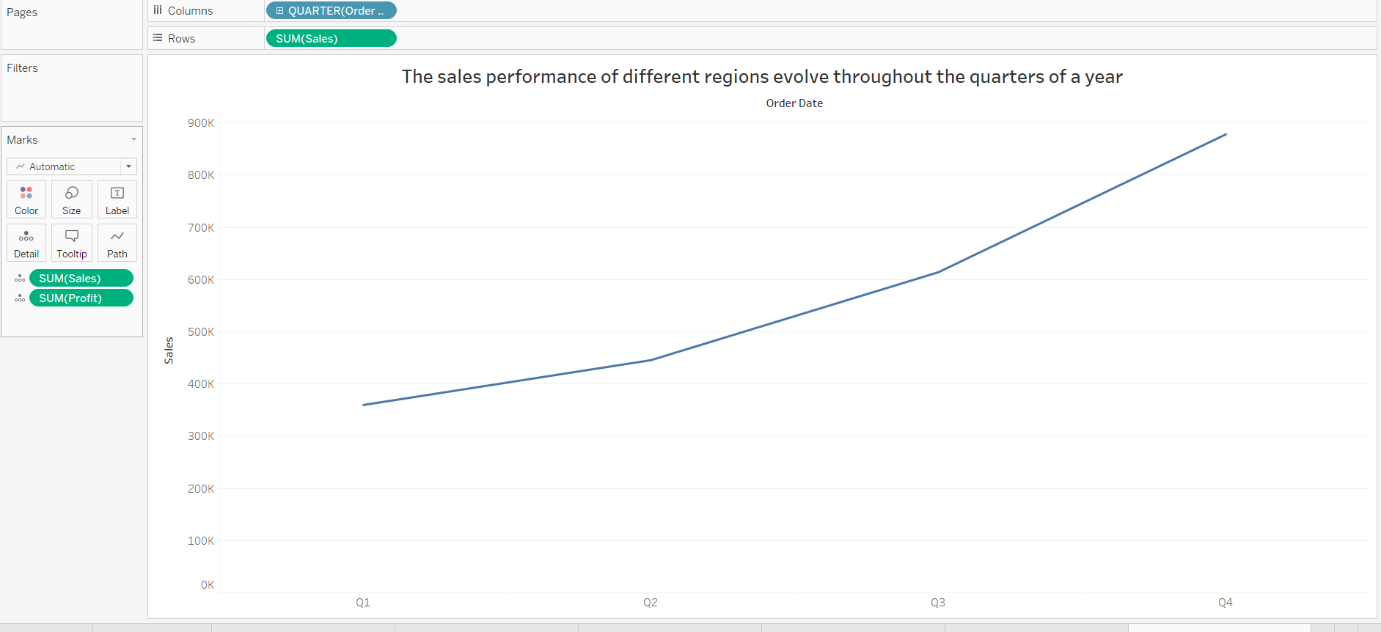
1. Which shipping mode is the most commonly used in the Sample Superstore dataset?

* The column chart represents the shipping mode that is most used. Set the x-axis by the shipping mode while setting the y-axis by the number of orders. The size of the column tells us which shipping mode is mostly used.

y

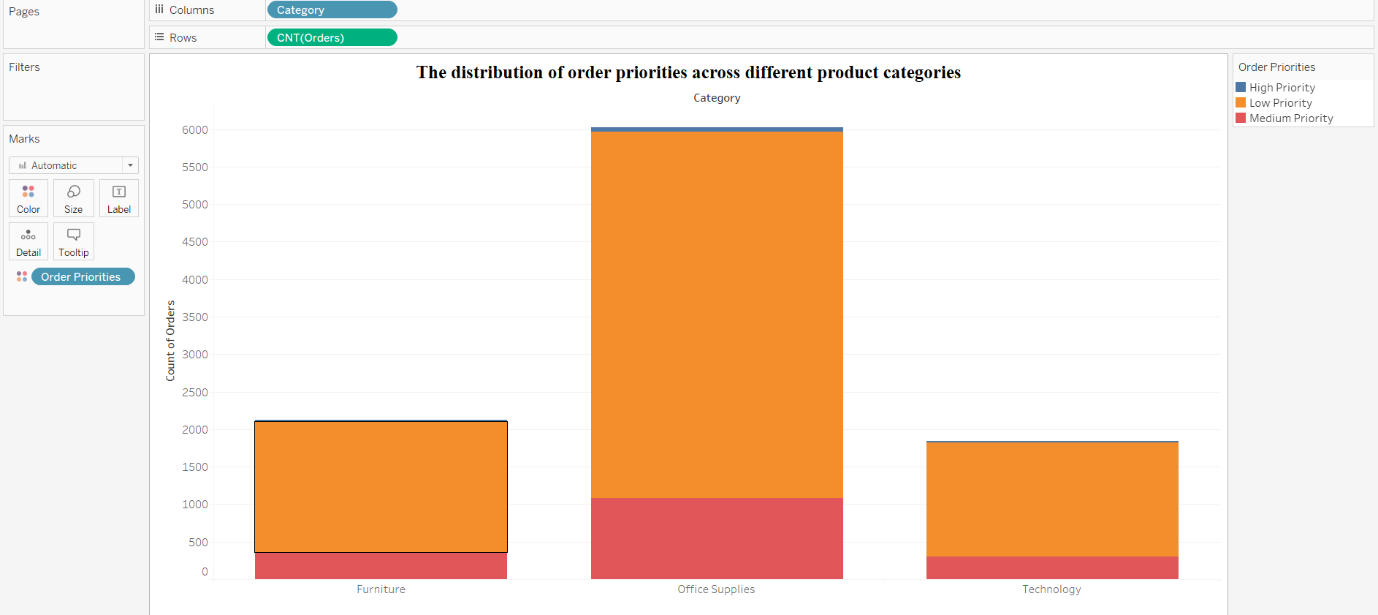
1. How does the sales performance of different regions evolve throughout the quarters of a year?

* The line chart represents the trend of sales over the quarters of the years. The x-axis tells the quarters of the year while the y-axis tell about the sales.



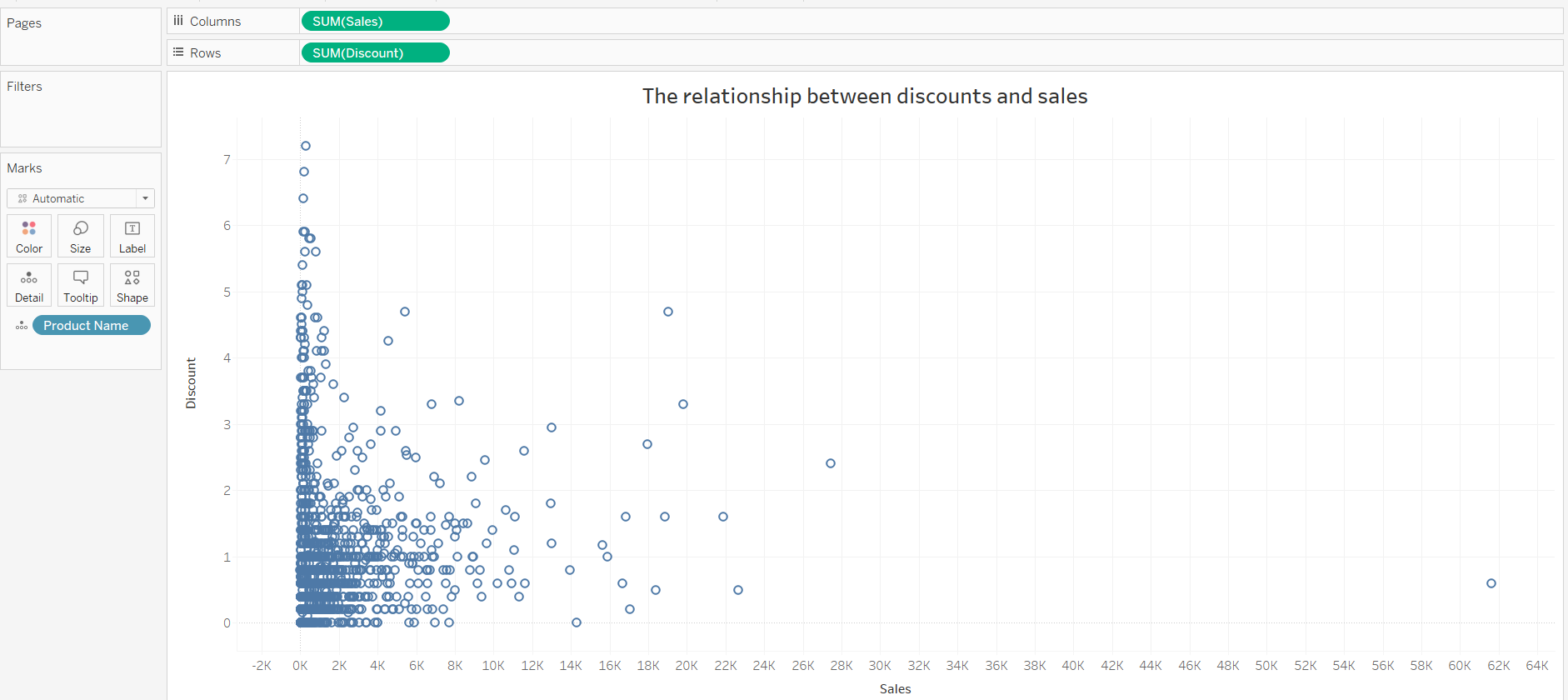
1. What is the distribution of order priorities across different product categories?

* The column chart shows the order of priorities across the different product categories. The split of the column tells order priorities. The x-axis shows product categories and the y-axis shows the number of orders.



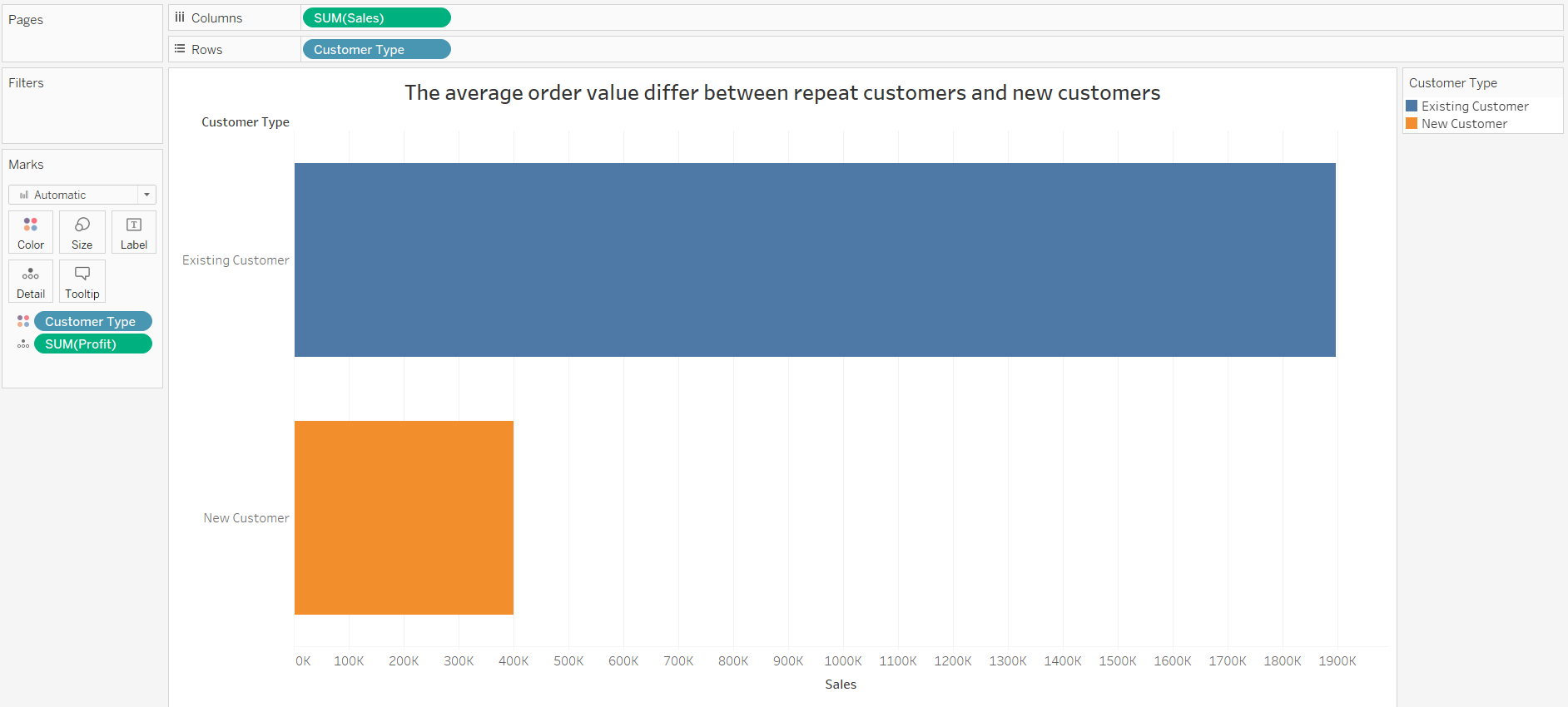
1. What is the relationship between discounts and sales?

* The scatter plot shows the relationship between the discounts and the sales. The x-axis shows the sales and the y-axis shows the discounts.



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

* The bar chart represents the sales of repeat and new customers. The x-axis shows the sales column and the y-axis shows the existing and new customers.



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

* The map shows the area where the store is located. The hover over the city tells us the city name, region, profits, and quantity.

