Course Project Report

Smart Supply Chain for Data Analysis

Team Members:

Neeraj Deshingkar
Naved Naik
Shivam Srivastav
Alex Weiner

1.0 Introduction

In recent years, supply chain management has gotten discombobulated by various factors. The primary influence of this disarray has been the Covid-19 pandemic. The countless people losing their jobs and numerous people working from home due to the pandemic has caused mayhem in regards to supply chains around the world. As a team, our motivation to dive deeper into this topic was the implications that this type of data has in the real world of business and supply chain analytics and also the relevance of this topic to our current time. First, for some background on our data. The dataset we analyzed was a supply chain dataset used by the company DataCo Global. The link to this dataset can be found here: https://www.kaggle.com/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis.

Our Dataset contains over 180,000 rows and 45 columns. Each row in our Dataset includes the product purchased, type of product, product category, profit per order, cost of the product, geographical data, delivery mode, and customer segment.

This dataset contains a vast range of data from delivery status, to department name, to even the order city and product id.

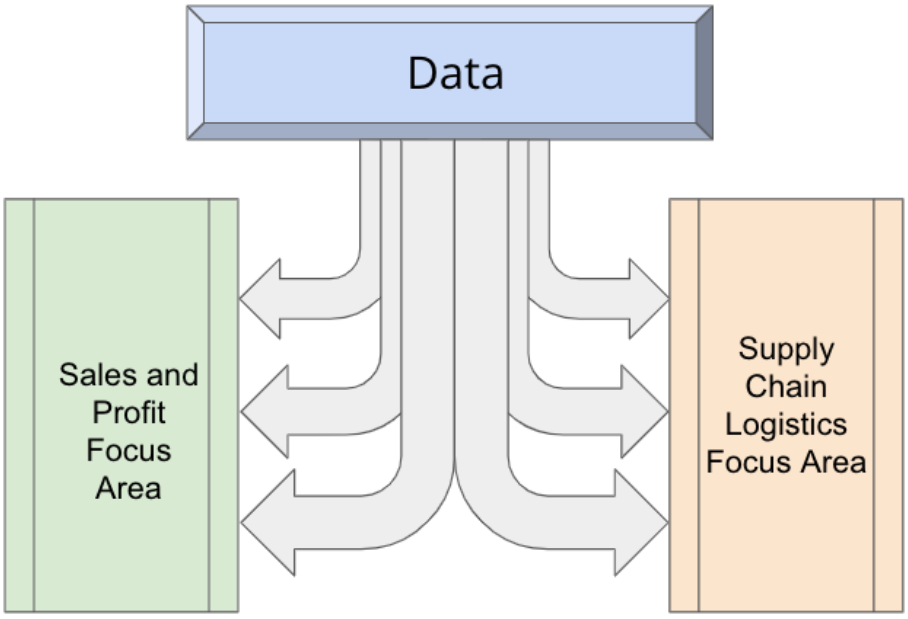
A few questions we set out to answer include:

I. Which market has higher/lower sales?
II. Which product category has higher/lower sales?
III. Which type of shipping mode does each order region prefer?
IV. What is the trend of late orders with respect to region?
V. Which product has a higher chance of stockout with respect to region?
VI. Which products exhibit the highest frequency of late orders?
VII. What is the average profit per department?
VIII. What is the profit per region?
IX. What are the product sales per customer segment?
X. What is the profit per region (Country) ?

Our objective for this data analysis project is primarily to focus on the production, sales, distribution, and provisioning components. From here, we wish to determine where the supply chain is performing adequately and where it is failing to meet supplier and buyer demands. After this, we will summarize our findings and make possible recommendations along with highlighting notable trends in the dataset.

Softwares used: Tableau and Excel
Files type: CSV and xlsx

2.0 Approach



When it comes to data compartmentalization, our game plan was to divide the data into two focus areas: sales and profit related data, and supply chain logistics related data. These two focus areas will be the main fields of concentration throughout the report. Our reasoning for dividing up the data in this manner was because not only were we working with an enormous amount of data, but we also thought it would be much more comprehensible if we compartmentalized the data this way for not only us, the users, but you, the viewer. We believe these two main focus areas capture the essence of what we are trying to present throughout this report.

2.1 Why Analyze Supply Chain related Data?

During such a period of uncertainty there are a lot of Supply Chain related aspects that could go wrong and could have a negative impact on the company. This could be the stockout cost a company has to bear, late order cost, or just cost associated with the type of shipping a company chooses because of unavailability of cheaper options.

We analyzed the supply chain data of DataCo Global. As this company shipped various products worldwide which gave us the ability to analyze various parameters and their effect on the sales/revenue/profit of the company.

3.0 Data Analysis:

The importance of supply chain management and supply chain logistics is paramount, especially in today's day and age. The spontaneous influx that the global supply chain faces every day is a tall task to handle. Through our data analysis, we set out to find trends and useful figures in order to capture the adequacies and inadequacies of the supply chain in recent years. First, we will look at two charts relating to sales and profitability.

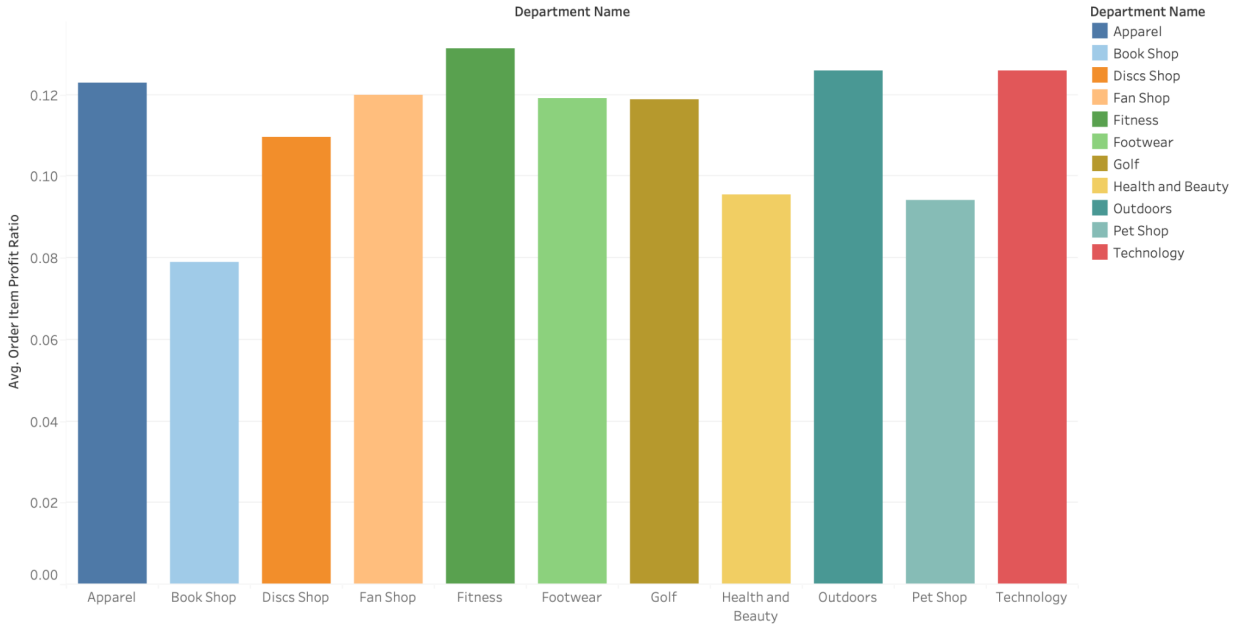


Figure 1: This bar chart shows the average profit ratio per department type.

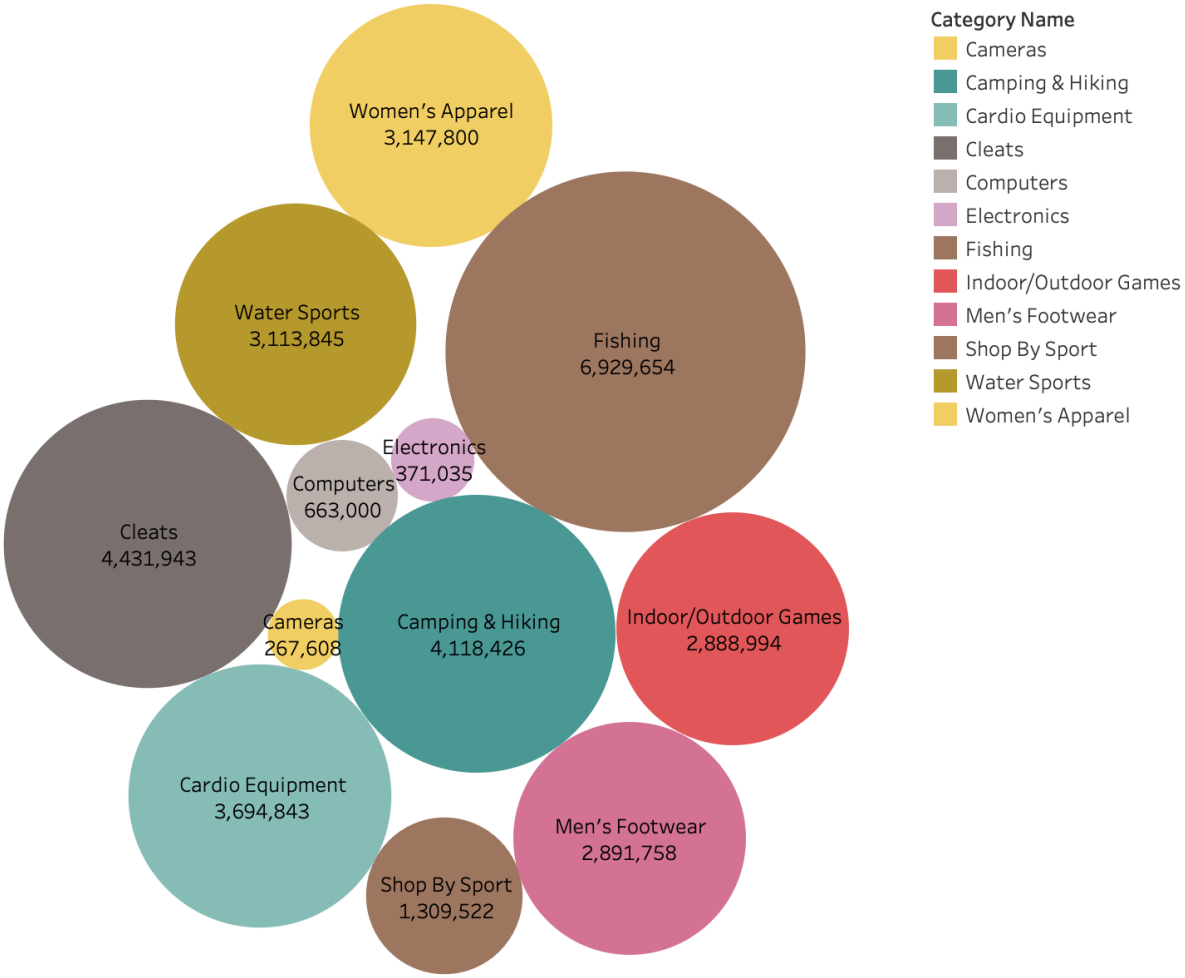


Figure 2: This bubble chart shows the number of sales per product category.

Looking first at the chart relating profit margin to department type, it is evident that the company turns a higher profit in the apparel, outdoors, fitness, and technology departments and turns a significantly lower profit in the book shop, health and beauty, and pet shop departments. Now looking at the chart relating product category to number of sales, it is obvious that fishing, cleats, and camping and hiking dominate the sales portfolio while electronics, computers, and cameras amount to a very small portion of the companies overall sales. Bridging these two data components, we can make statements and inferences regarding where to allocate more advertising, price changes, and even department reduction/expansion.

After understanding where consumer money is being spent and the profit that this company makes on their products, it is important to know who is buying their product. For this, we will look into sales per market and sales per customer segment.

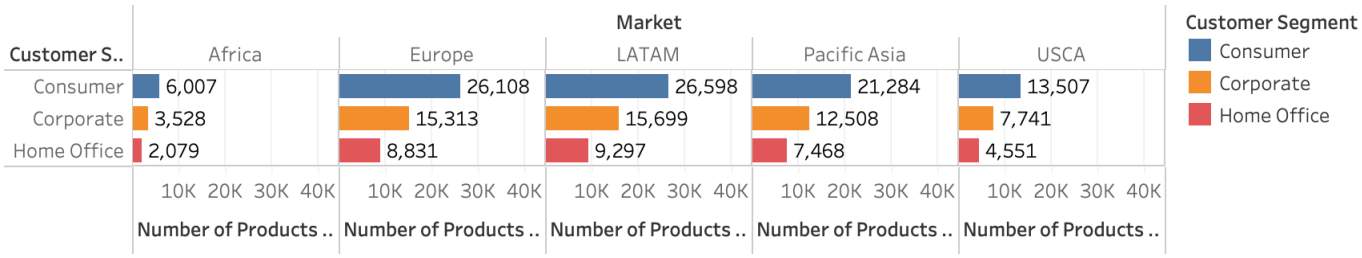


Figure 3: This bar chart shows the sales per type of customer (consumer, corporate, home office) in various markets.

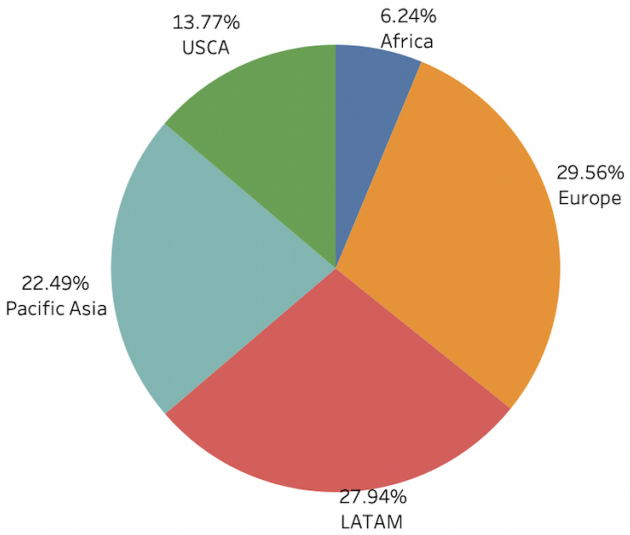


Figure 4: This pie chart shows the number of sales per market.

After having examined these two charts, it is evident that the majority of sales are done in Europe and Latin America. It is also evident that the consumer segment holds a large portion of the sales portfolio. Using this insight, the company can make predictions regarding future demand across various markets and consumer segments. They can also adapt their marketing scheme as they see fit in order to capture the sales in a desired market/consumer segment.

Figure 5: This map shows the profit per region.

Using this world map, we can see where DataCo Global turns the highest profit with respect to the region. While this is not the average profit, rather the total profit, it gives the company a good sense of where they are generating the most revenue. This visual allows DataCo Global to adjust their marketing scheme and product prices accordingly to meet their company objectives. We see that the United States has the highest profit amongst other countries.

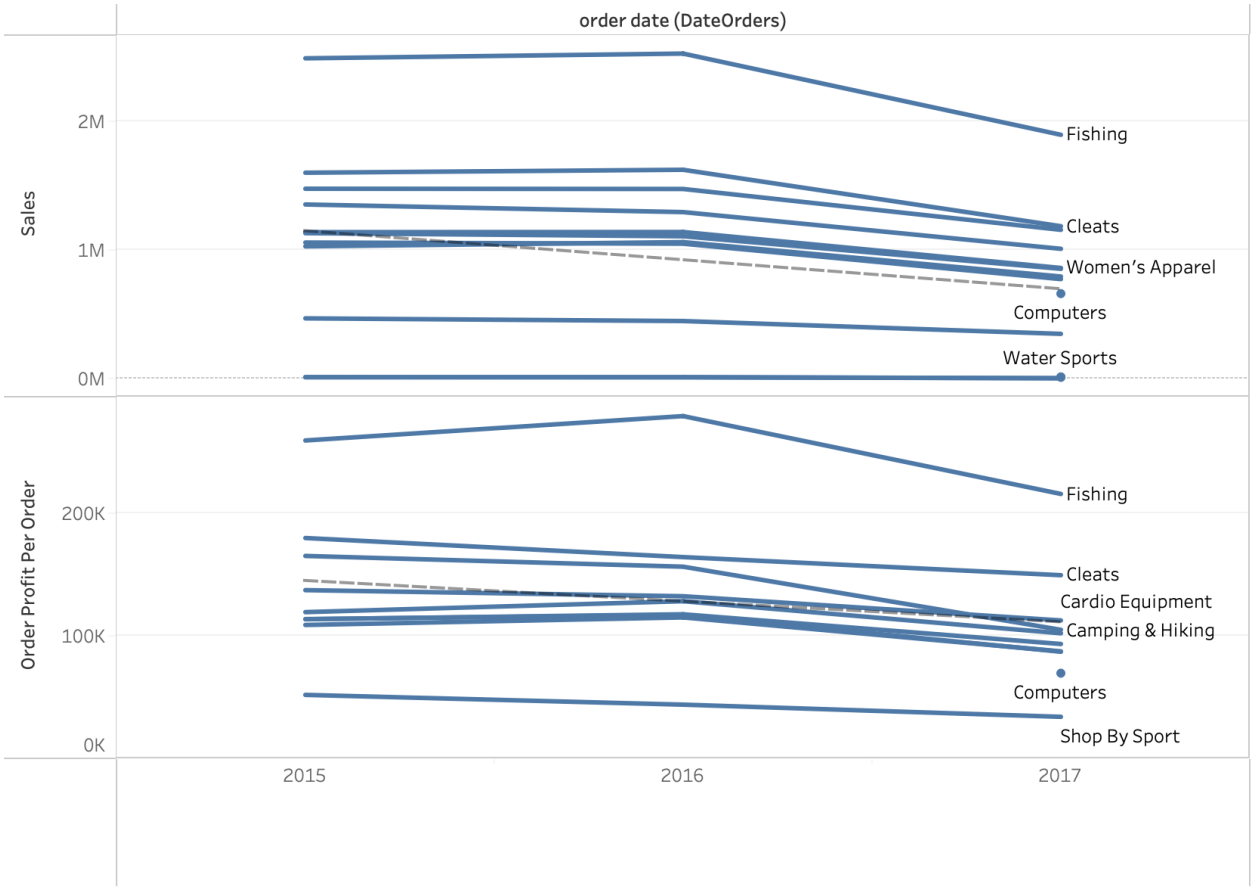


Figure 6: This line graph shows the sales over time and profit over time for various product categories.

After analyzing the DataCo. Global dataset we have found that the company has a declining sales and profit (on the basis of the trend line plotted above) and this has to do with orders arriving late and product stockout. This could also be a result of negative impact on the company as customers would choose to avoid this company.

Having analyzed the sale and profit sector of DataCo Global and all that entails, we will now investigate the supply chain logistics sector of DataCo Global. By doing this, we hope to get a better understanding of how their order processing to shipment to delivery process works. First, we will look at products which result in the most late orders.

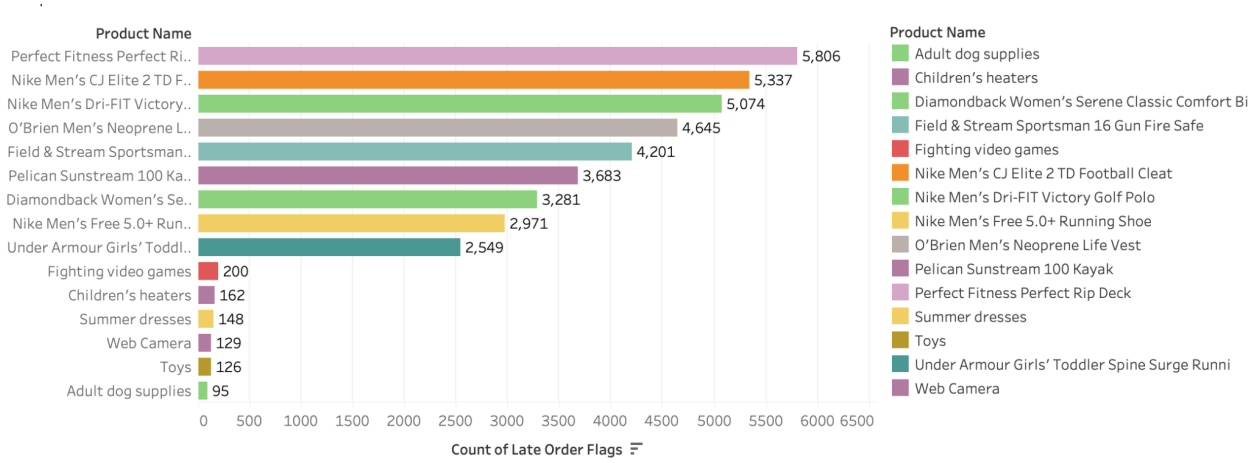


Figure 7: This bar chart shows the late orders per product for the top 15 products.

As we can see in the Tableau visualization that the top 9 products (when sorted by Descending order) comprised the majority of the late orders. This shows that roughly the 20% of the products account for the 80% of the late orders (Pareto principle). Now this could be region specific issue or product specific for which more analysis is done below.

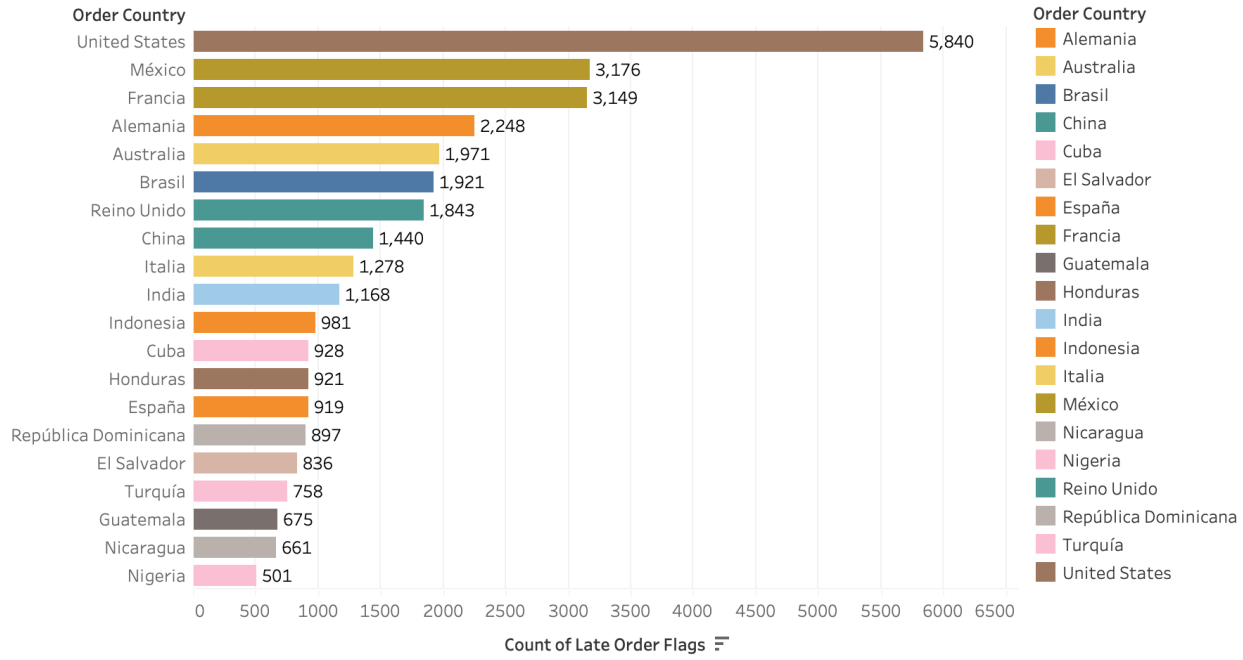


Figure 8: This bar chart shows late orders per region.

As we can see in the above visualization the United States of America comprises the majority of the late orders. Further investigation needs to be done to find the cause for that. Whether it is because relatively more products are being shipped within the United States or is it a supply chain related issue.

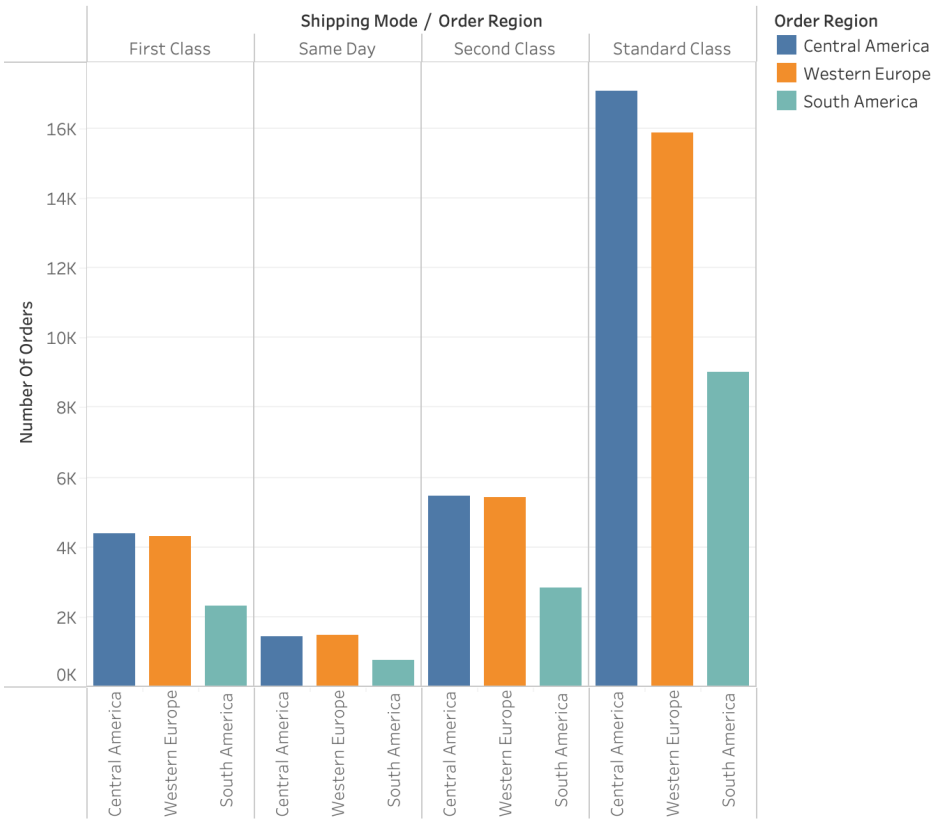


Figure 9: This bar chart shows the type of shipping mode per region.

We see that the majority shipping is of standard class and that is mainly because of the type of product and the customer segment (mostly consumers) which the DataCo Global company pertains to.

3.1 Inventory Stockout

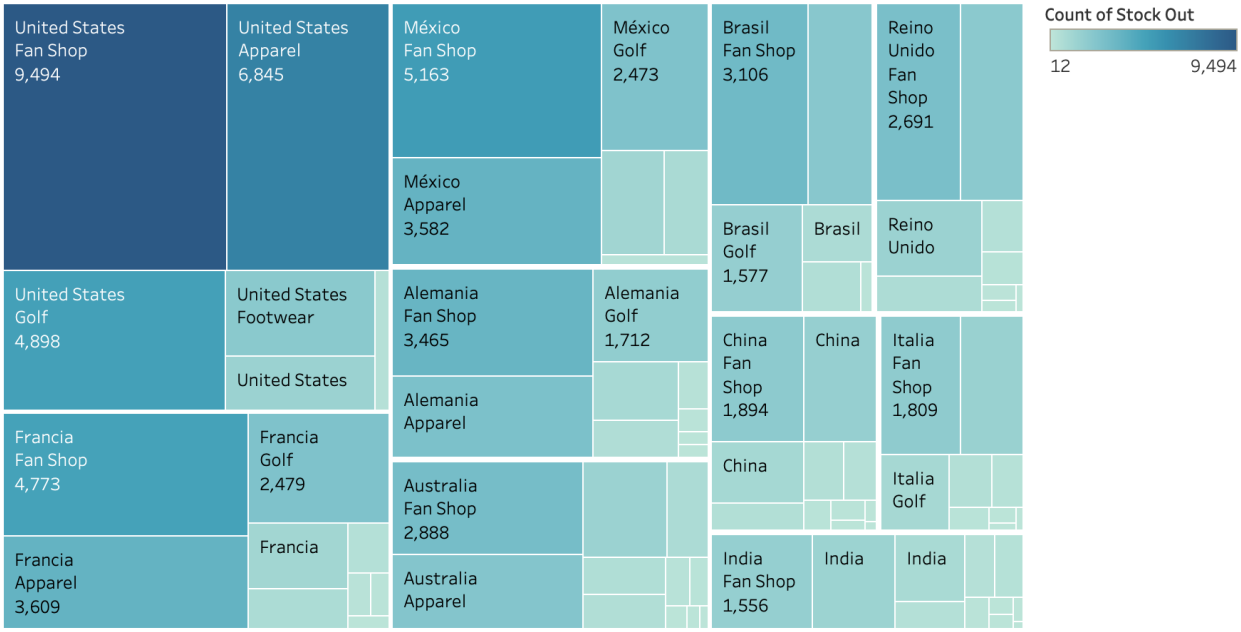


Figure 10: This tree map shows the number of stockouts per region by department.

Stockout cost is the lost income and expense associated with a shortage of inventory. We come to the conclusion through the visualization made in Tableau that Fans in all countries have the highest stockout. This could be avoided by recalculating Economic Order Quantity based on customer orders in a given region and keeping buffer stock in those regions where demand is more and unpredictable. The stockout of Fans could also be seasonal and that could be taken care by increasing the buffer stock.

Summary/Conclusions

Upon completion of the data analysis and the thorough investigation of our topic, we realized just how influential decisions about supply chain logistics can have on the global supply chain. Using the data from DataCo Global, we were able to sift through various components of data and make inferences and conclusions about not just their sales sector, but their supply chain logistics sector as well.

Looking into the sales and profit sector of DataCo Global, we were able to not only see which products were being purchased, but also where they were purchased and the profit that the company made on that purchase. Along with many other components given in the dataset, we were able to make inferences based on certain trends we noticed and using this information, make predictions as to what the company's next plan of action should be.

Now looking into the supply chain logistics sector of DataCo Global, we were able to analyze the order punctualness, type of shipping method, and stockout data. From this, we were able to

show the relation between various components of the data. These correlations may prove useful to the company when it comes to turning the highest profit on their products. This is because not only does the sale of a product determine the profit made, but also the timeliness of the order.

Finally, we would like to stress the quintessential role that data analytics plays in regards to the global supply chain. Through thorough investigation and research, more informed decisions can be made in the realm of supply chain management, benefiting all aspects of the supply chain.