Global Superstore Business Analytics:

Shipping Inefficiency and Driving Profits

Steven Fernandez, MD Miah, Mohammed Hussain

Python for Business Analytics

ECO 32500

14 December 2023

Introduction:

The retail and consumer goods industry is an interesting and dynamic sector, offering valuable insights into the intricacies of consumer behavior and market dynamics. Its main purpose is to satisfy the consumer's desire by facilitating the exchange of goods, from manufacturers to end-users. Retailers generate revenue through the sale of products to their clientele, keeping them afloat from the costs that they incur by sourcing the products that they sell on their shelves and the cost of operating the business. This is the business model by which retailers abide by and along with it involves navigating numerous challenges throughout their tenure.

From customer retention strategies to efficiently managing logistics to provide the products that consumer's need and want, retailers face a multitude of hurdles in their quest for sustained success of their business. In the context of this landscape, our focus turns to a similar case study, a global superstore facing performance challenges as it approaches the upcoming year of 2018. This report delves into the intricacies of the retail giant's challenges, aiming to identify opportunities and solutions that will contribute to its revitalization and long-term success. As we navigate through the data and find insights of the company operations from 2015 - 2018, our objective is to reveal the complexities this global superstore is confronting and propose strategic insights for navigating the competitive landscape and enhancing overall performance for the upcoming year.

We did the following steps in order to arrive to our insights and data-driven solutions:

The following was the question that was asked by our regional director:

- I am looking to improve this year's performance, what suggestions do you have for success?

We reframed the question to the following:

- What opportunities are there in order to increase profits and have better customer retention ?

What data did we need, we needed any data that was relevant to the following:

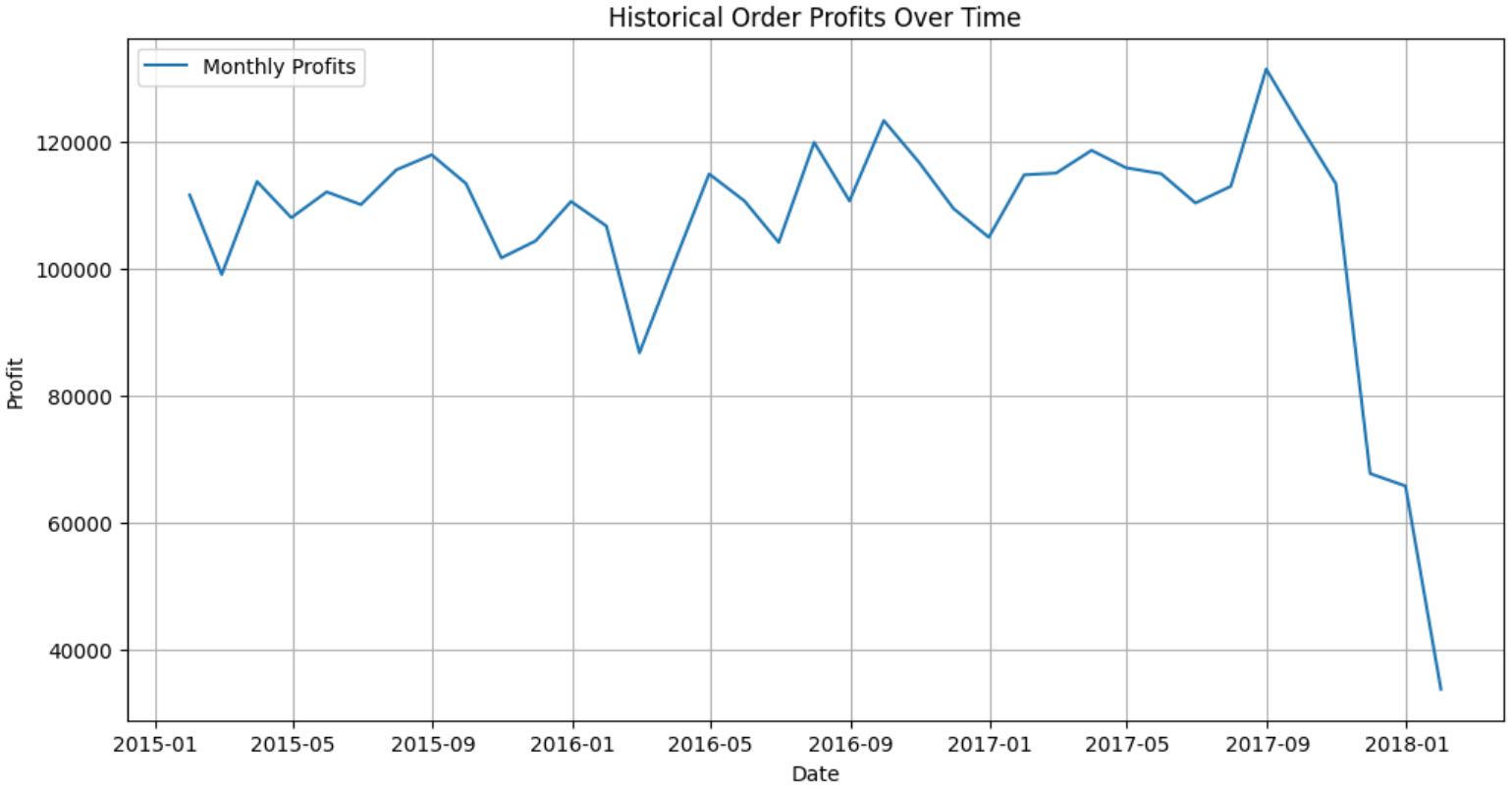
- We needed sales data
- Revenue data
- Order details data
- Product details data
- Customer demographic details data
- Shipping details data

Data Cleaning:

- Change order_dates to datetime
- Manipulated expected shipping days and real shipping days data to get the amount of days that deliveries where late
- Converted customer_IDs and category_names, in certain cases, to gain count of those attributes and later manipulate them to derive analysis on orders per markets and orders per category
- Transformed the Product_Name using libraries and for loops to create a combination of the most frequently pair of products that were sold together

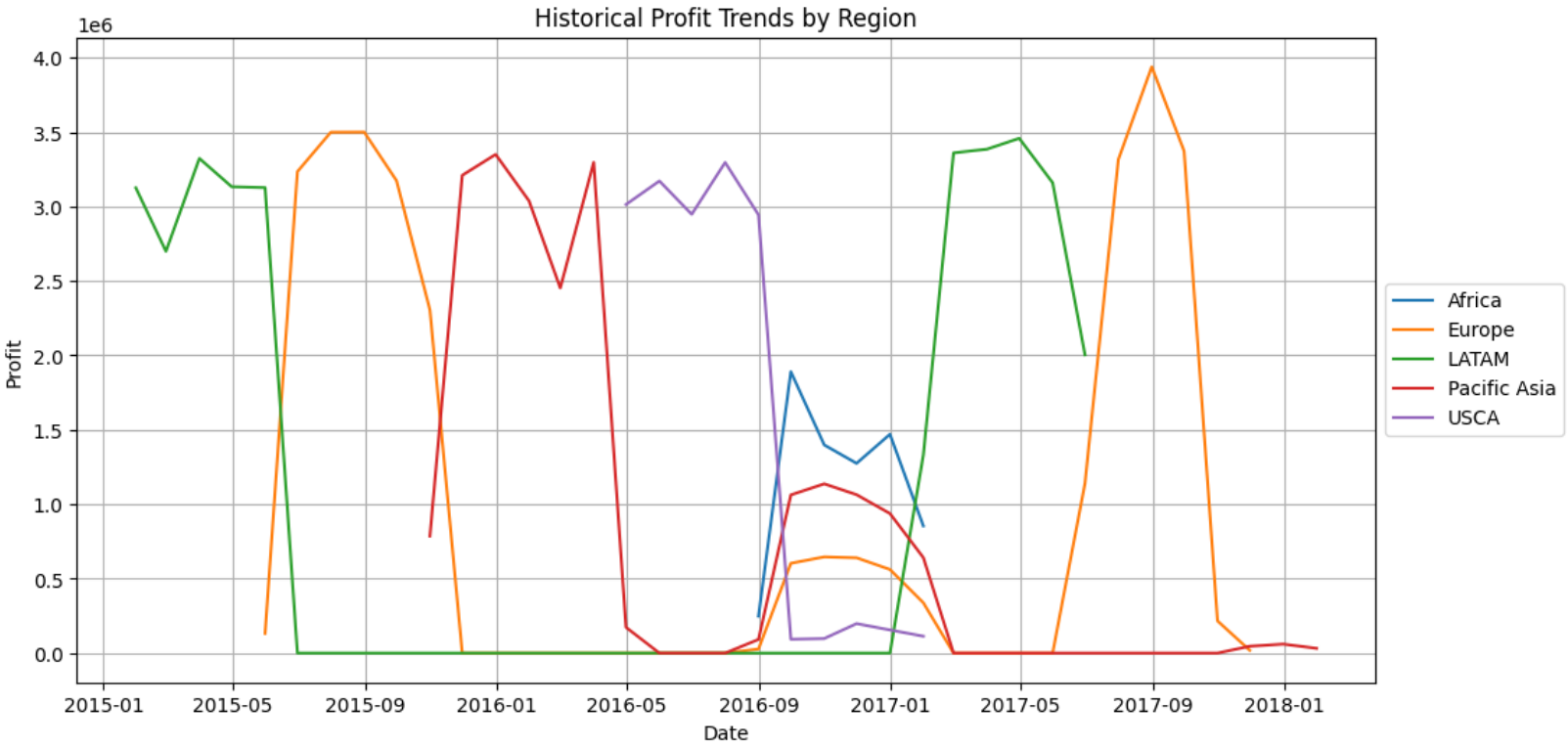
Analysis:

Overall profit trend:



We can see that we have some normal profit trends from 2015 - 2017. However, our performance plummeted at the end of the last few months of 2017 and the same trend is leading into our 2018 year. This led us to become curious of what happened and ask why has our performance dropped immensely in the last few months of 2017, leading into 2018. So we decided to go a level deeper and explore the profit trends by the region distribution.

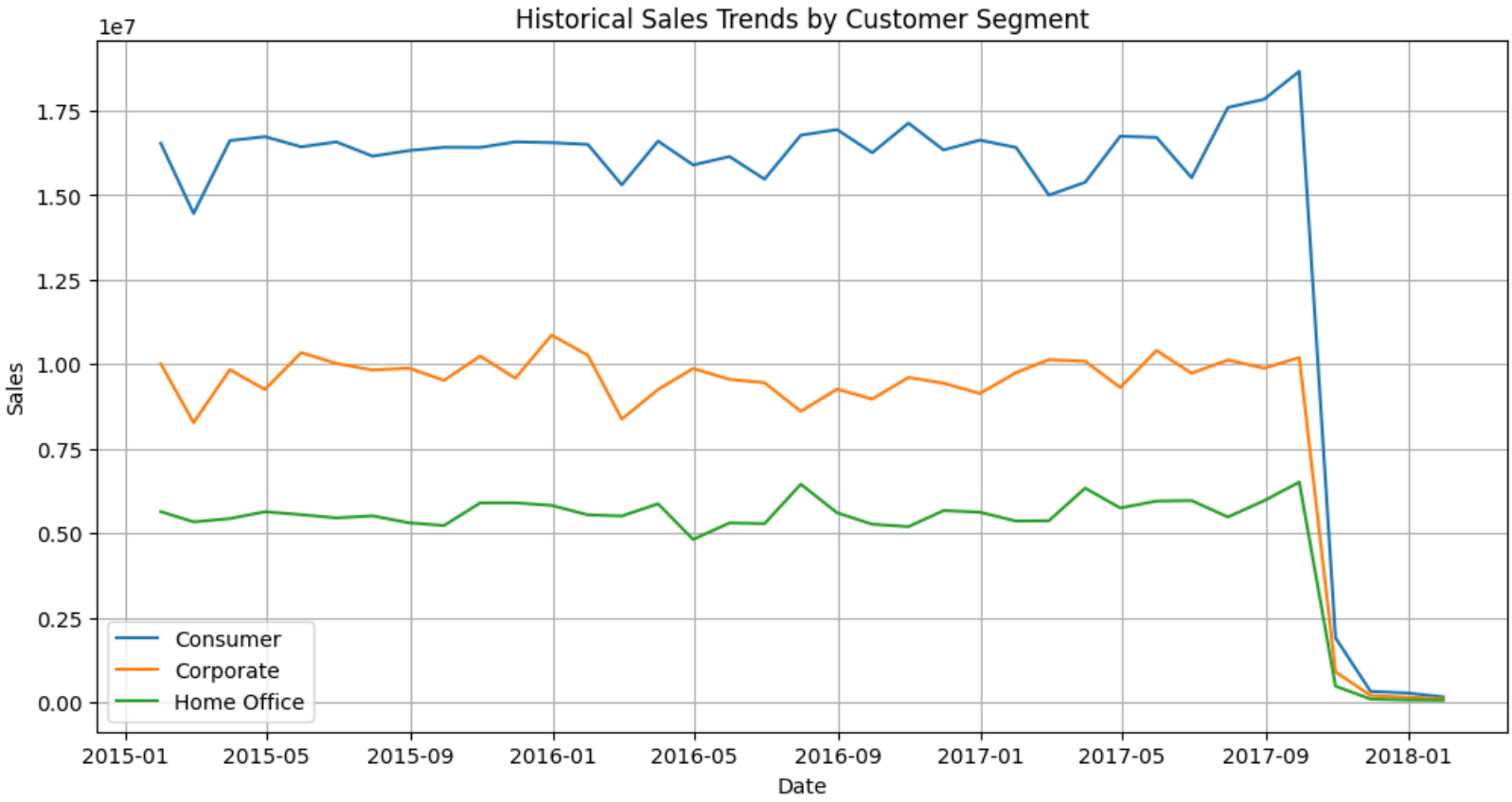
Profit trend by region channel analysis:



From this graph we can see what our profit trends looks like over the three years (from 2015 - 2018) of data that have available by the superstore. We can see here that our profit streams are not consistent for each region, indicating that the superstore is relying on streams of profit from only a few regions each year. This is what has solely kept the superstore profit streams fairly stable from 2015 through mid-2017. Until eventually it lost its profit stream when getting close to the 2018 year. We can see that as we got closer to the pivotal profit trend, the superstore was relying on making profits from the Europe and Pacific Asia regions only. As a result, the superstore profit's plummeted at the end of year 2017 because it only had two streams of profit channels and it was unable to maintain it to last the business when entering 2018. We now wanted to know what factors has contributed to this plummet in profits and what is causing these channels to be profitable for only a short amount of time. In order to answer this, we reframed our questions as the following:

- What type of customers are we dealing with and what is there profit trends per channel ?
- How is business performing in terms of retaining our customers ?

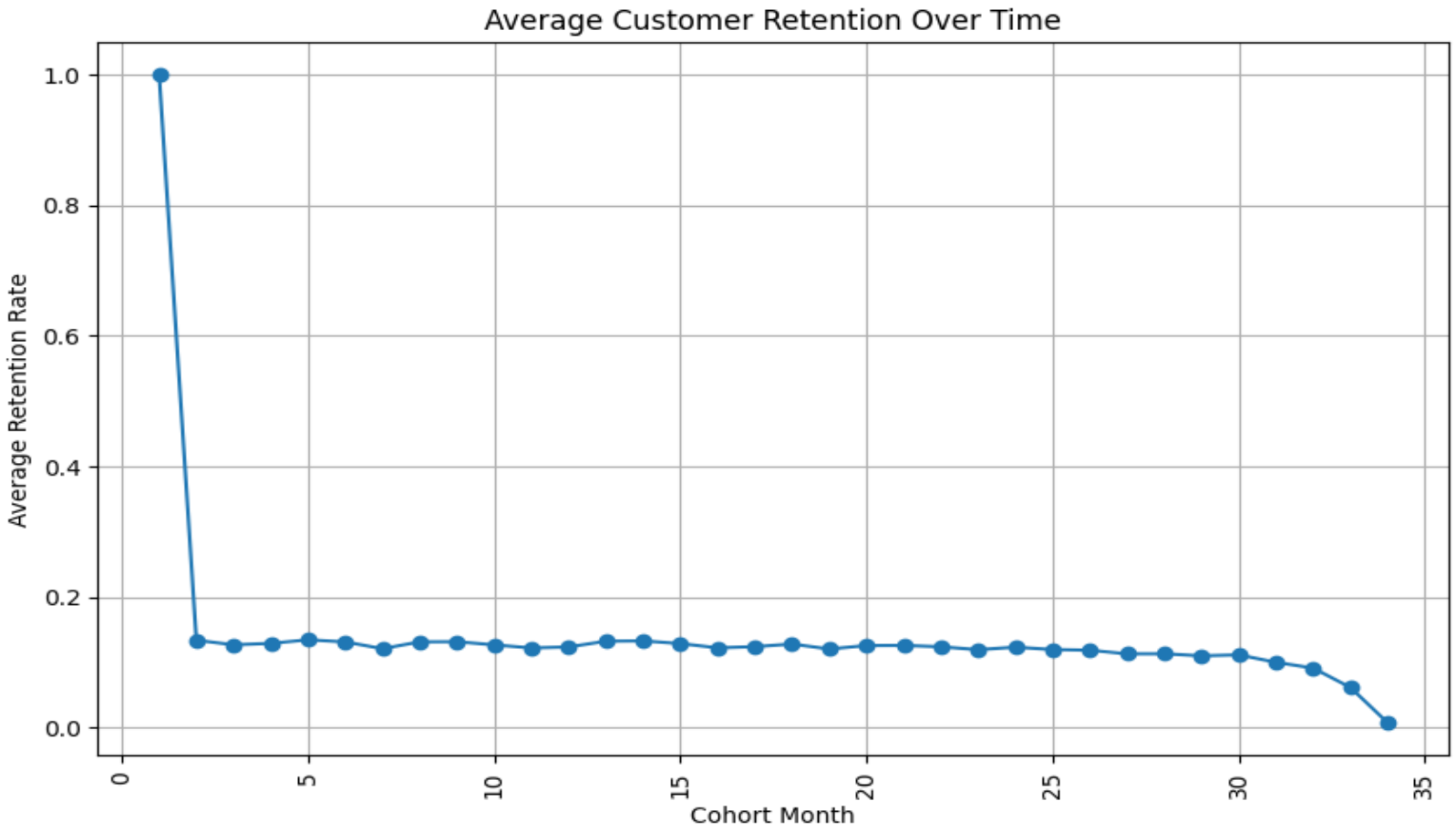
Customer segment profit trend analysis



From this graph we can see that our customers are divided into three segments; consumer, corporate, and home office. The graph tells us that our main customers are consumers with corporate being second and home office being the third. We can say that most of our profits are derived from our consumer customers and we can also see that before the superstore plummeted our consumer customer segment was increasing much more rapidly when compared to the other customer segments. This piece data tells that we should mostly focus on consumer customers as they are the ones that are driving most of the store's profit and delivering growth to our profits in the long-term, making them our most valuable customers. Opportunities that we can derive from this is to focus our inventory around what our

consumer segment is buying and providing them with benefits, such as a customer loyalty membership that gives them perks (such as discounts and special events) or we can provide them special offers.

Customer Retention:



We wanted to see how the superstore was retaining its customer over the course of the three year of data that we were provided. After the first month of acquisition of the customer we can see a huge drop of about 80 - 85% in our customer retention rate. This rate remains fluctuating between 15% and 18% until the final few months where our retention rate drops to 0% along with our plummet in profits mid 2017. This entails that the superstore is not doing a good job of being able to attract and maintain its customer in the long run. Hence, this leads to the sudden drop in overall profits in late 2017. This also explains the

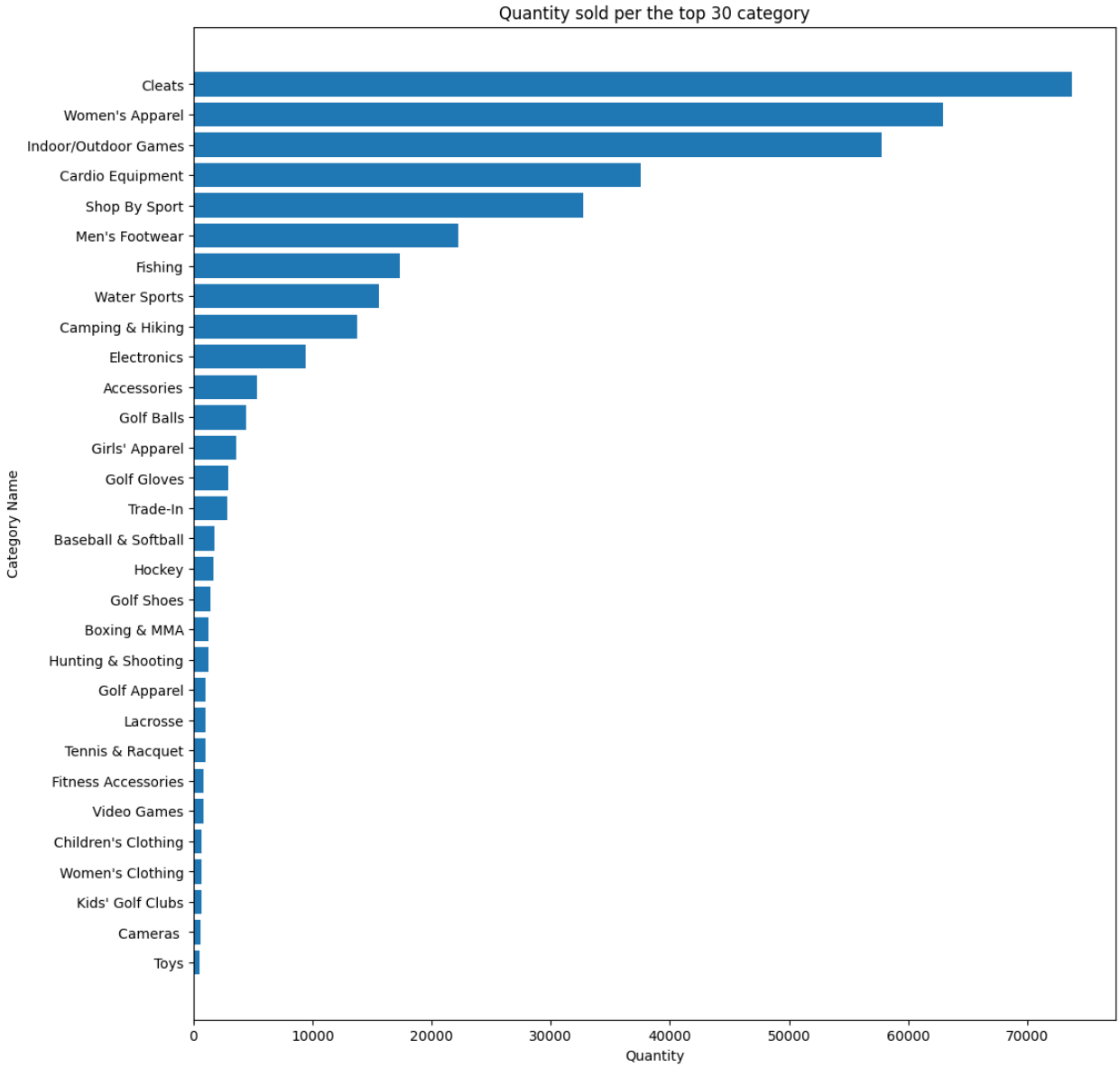
anomalies that's going on in the profit trends per region, the superstore is only profiting for a short amount of time for each region because it isn’t retaining customers year round.

So how can we make this a better situation for the superstore?

We reframed this question to the following:

- What were the distribution of quantity sold per category products ?
- Are there any attributes from our shipping affecting our business?

Quantity distribution sold per (the top 30) category product analysis:

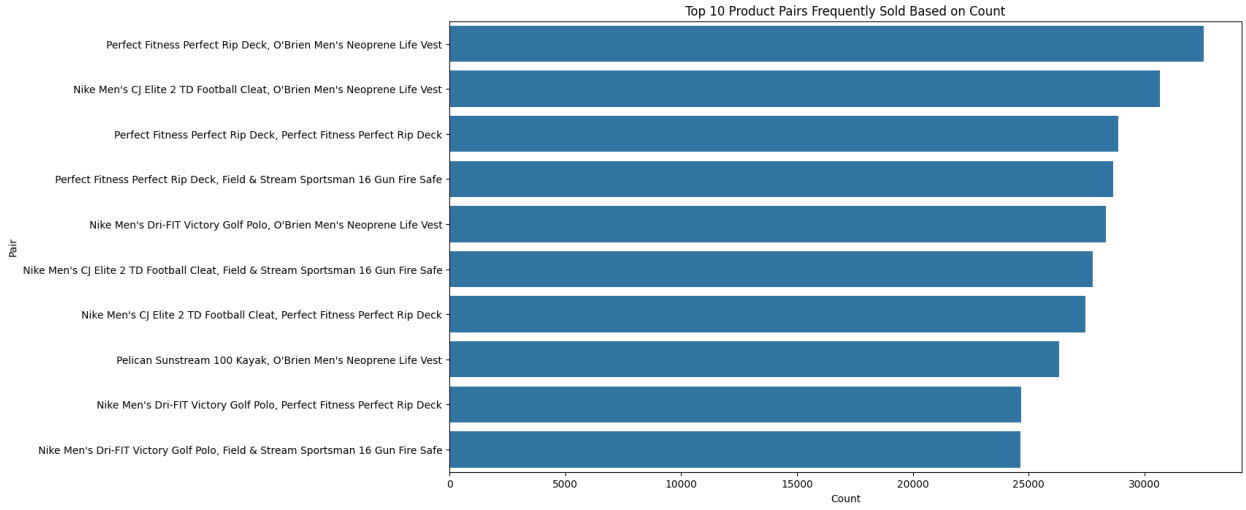


This graph tells us that we are only selling certain types of category products, as we can see that there are certain categories of products where there is no quantity of them being sold. Trends that we can see here is that the superstore is selling a particular type of categories of products, mostly all of them revolving around sporting goods. Hence, we can taylor our store to prioritize selling these categories of products to our customers as they are the ones that are most in demand in accordance to our customer base and what they are buying off our shelve.

Questions that came from this graph:

- How can we develop a package of products that would drive in sales ?
- When is the best time to market those products
- At what price should we sell those products
- Who and where should we target to sell these products/ test out the strategy ?

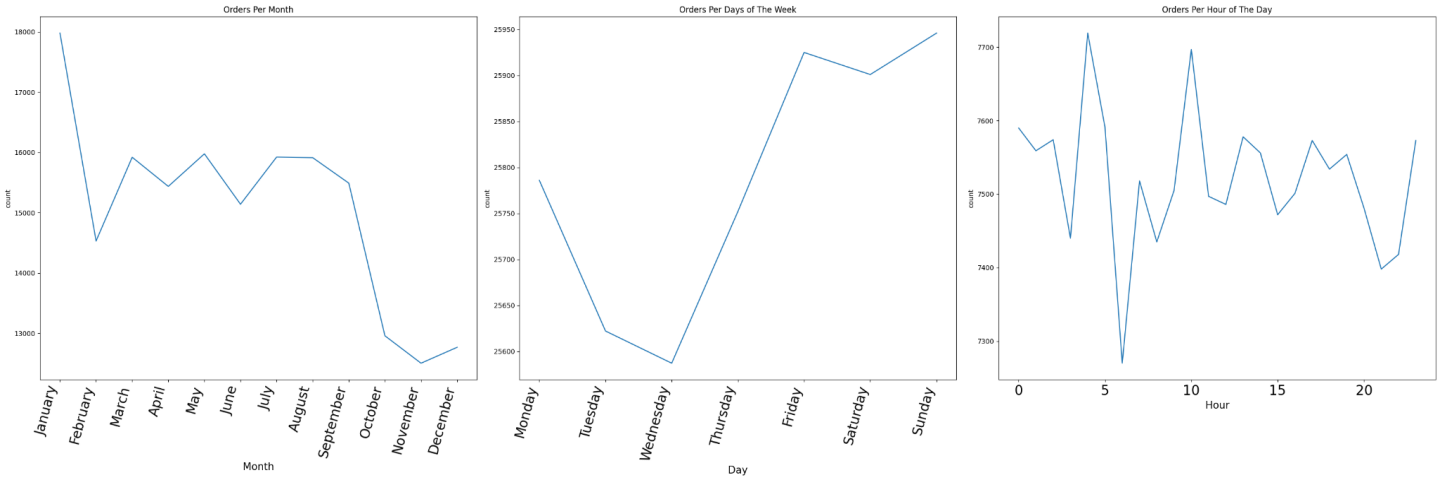
Top 10 Product Pairs Frequently Sold Based on Count:



From this graph we can see the top 10 pairs of products that are sold together most frequently. It can be seen that most of the pairs that are sold together are related to sporting goods so we should focus our business around those items. We should also have a good amount of inventory on those items in order to meet the demand for them. Opportunities that we can derive here is that these products can be sold together in a package and be priced at a fair value. We can also provide promotional discounts to promote the sales of these products together. We are going to explore those options further in the following, where we are going to determining the pricing and the hours to release advertising for the sale of these product packages.

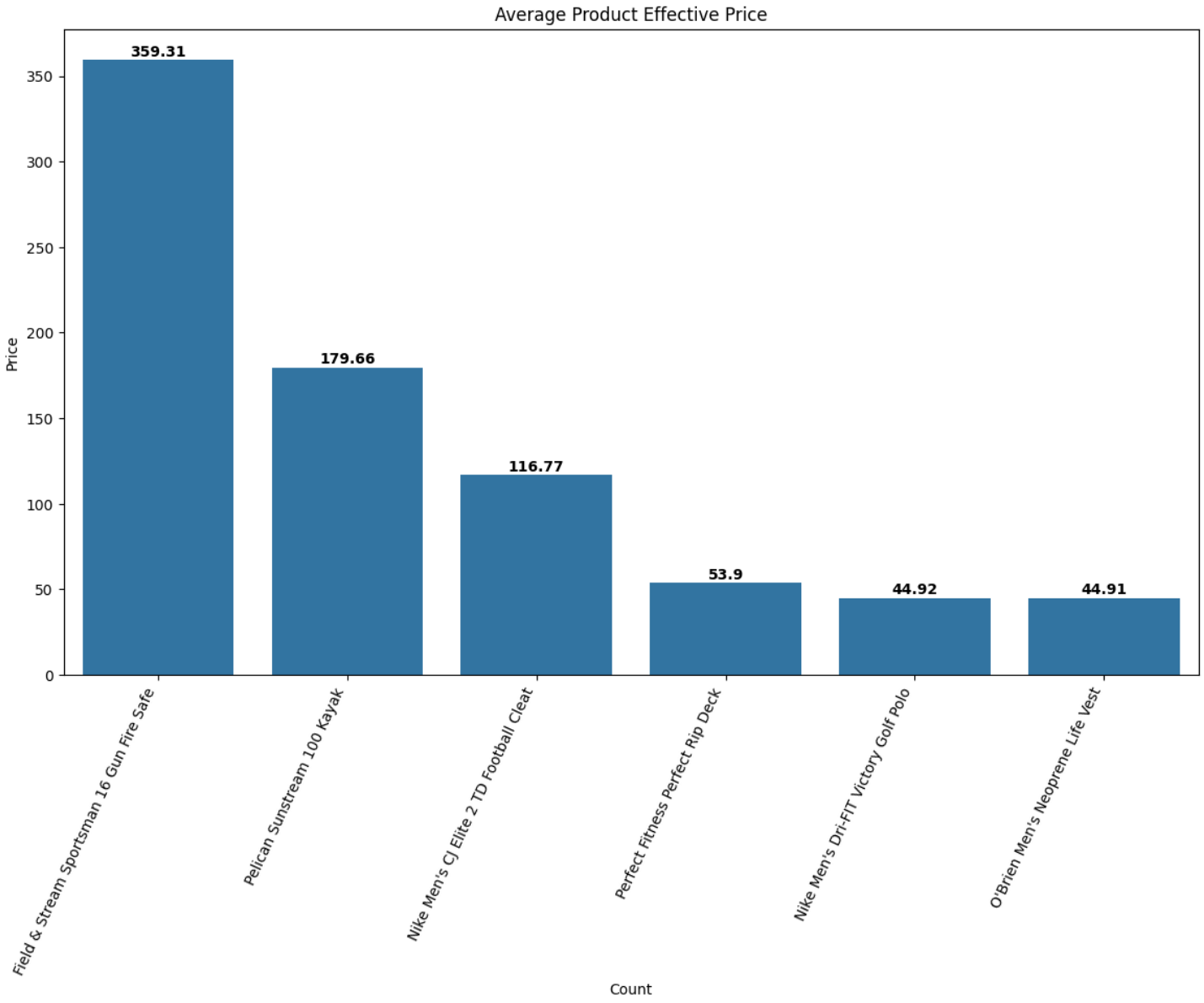
Marketing Analysis:

Order Market Trend Analysis:



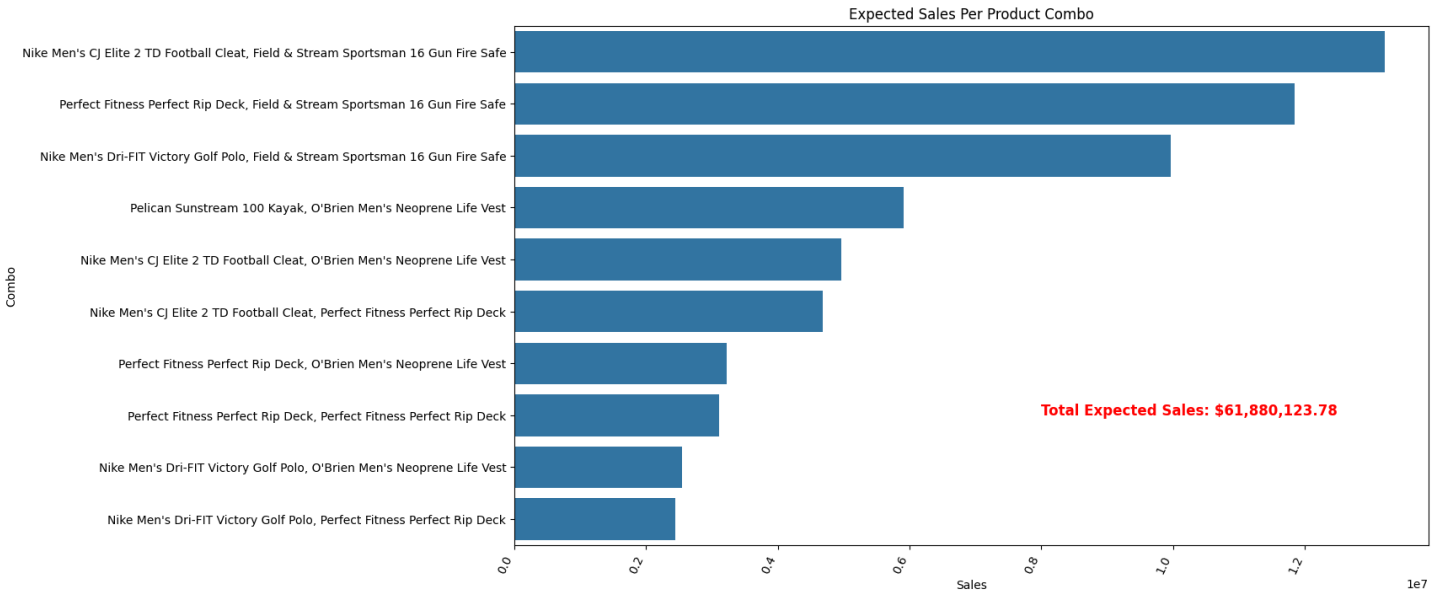
From these graphs we can see the best time to advertise the products. In terms of months, it is best to advertise the products in January, March - May and in the month of July. It is not recommended to advertise as much from October through December as those months we have very low orders of our products. In terms of what days of the week we should advertise, we should aim to advertise more towards the weekend (Wednesday - Sunday) as our order trend tends to be at its peak moments during those days of the week. In terms of what hours to advertise, we should aim to have advertising slots around 3am, 8-10am, and 5-7pm.

Pricing Analysis:



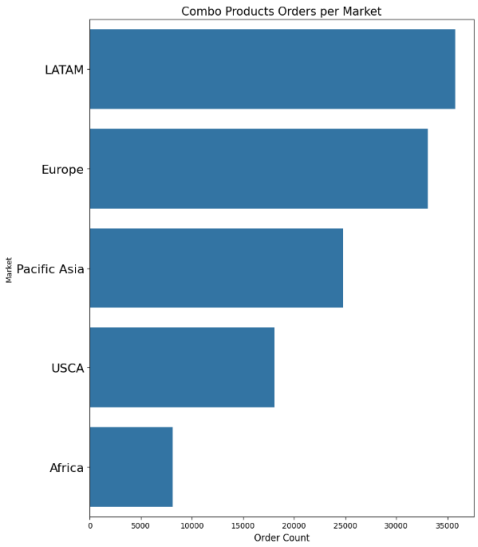
According to our pricing analysis we should aim to price each of the products as the stated in the graph above.

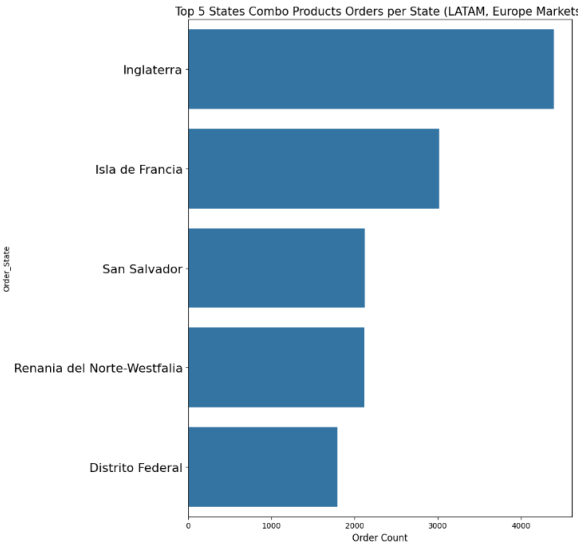
Combo Pricing and Expected Sales Analysis:

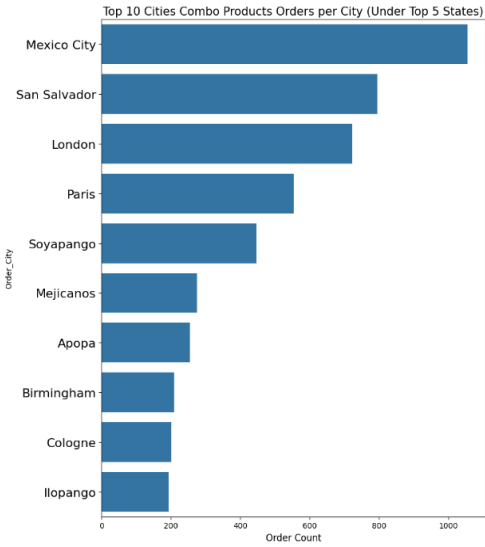


According to our pricing analysis of the combo product we should aim to price each of the products as stated in the graph above. In addition, according to our historical data, if our demand is the same as our historical demand then we should expect sales of around $61,880,123.78.

Best Places to Try Out This Product Package Strategy:



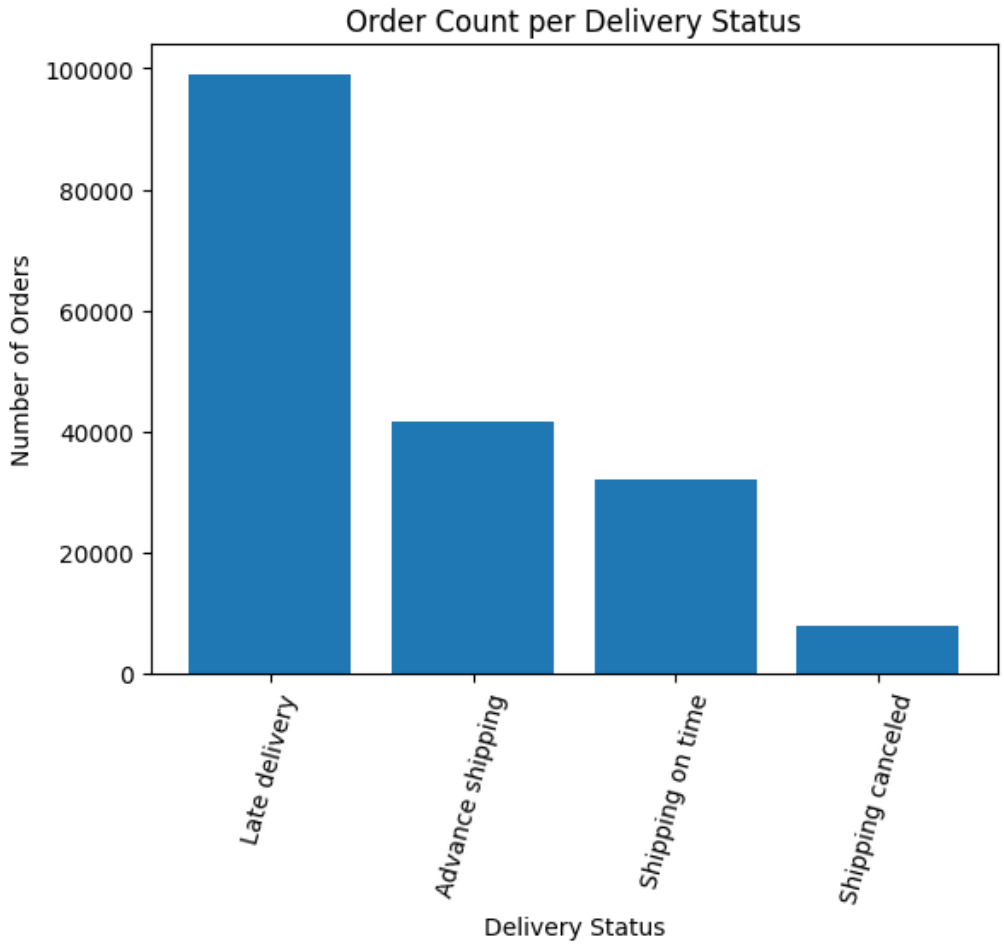




According to this graph the best places to implement this marketing strategy is in Latin America and Europe regions. Specific states that we can explore under those region are England, Island of France, San Salvador, North Rhine-Westphalia, and The Federal District of Mexico. The best cities that we can explore under those states are Mexico City, San Salvador, London, Paris, Soyapango, Mejicanos (Salvador), Apopa, Birmingham, and Cologne and Ropango. This will ensure that we are testing our combo product where the demand is most formidable.

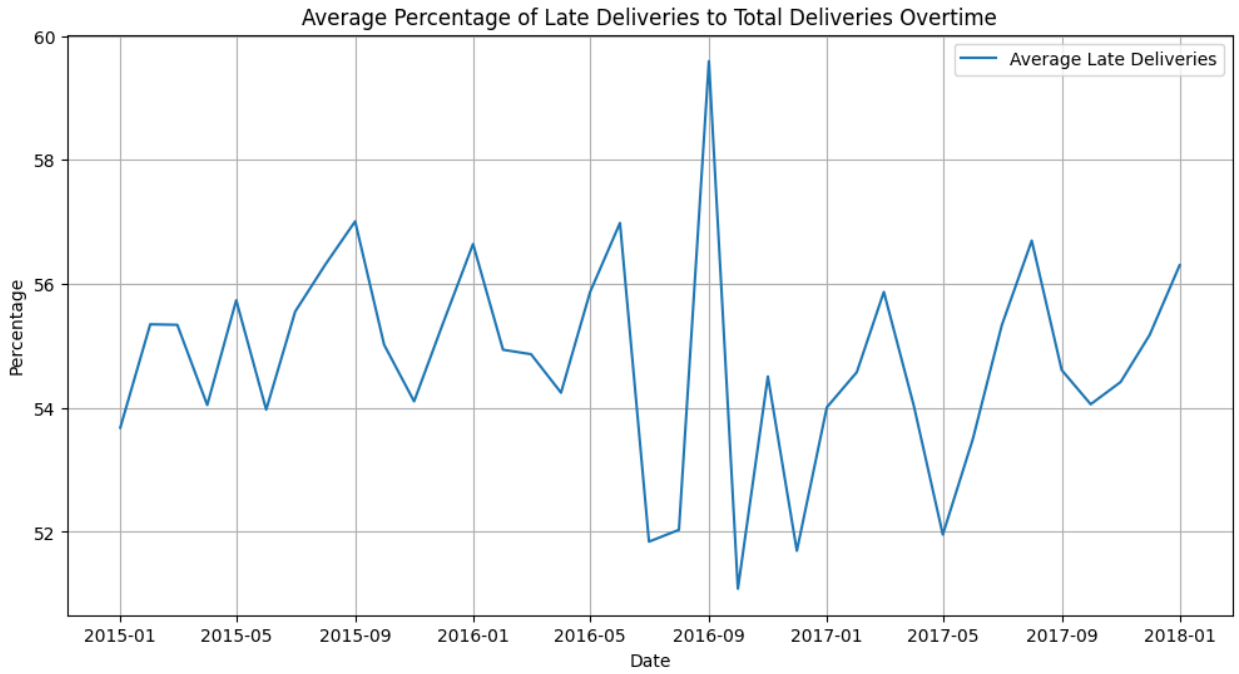
Shipping analysis:

Shipping order status analysis:



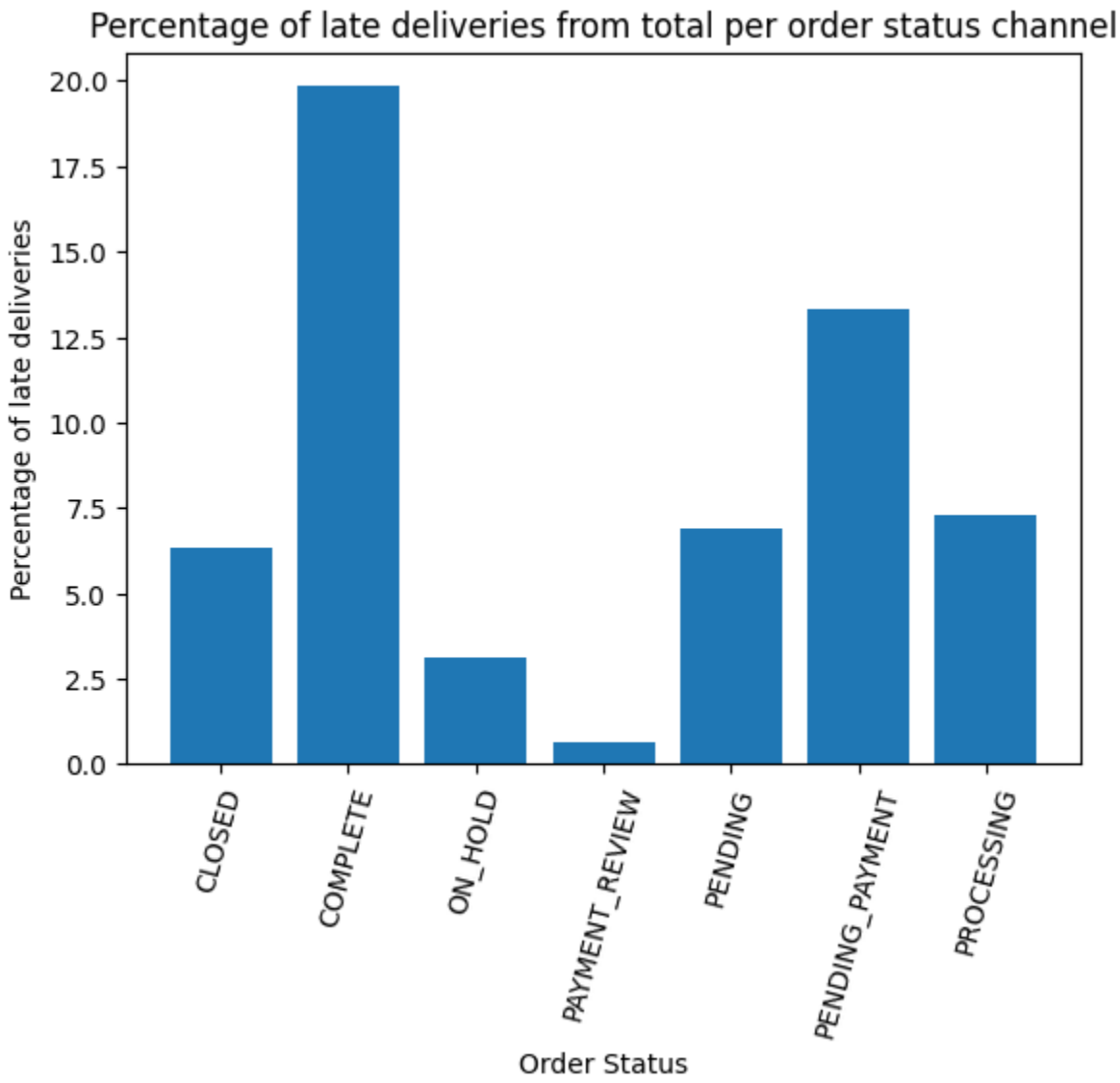
We can see that out of all of our deliveries, most of our deliveries are late. We now wanted to know what was the average percentage of late deliveries to total deliveries per each month that passed by.

Average percentage late deliveries trend:



We can see that the average percentage of late deliveries to total deliveries is around 54%. This is a big deal when we are talking about fulfilling the promise that we will deliver our products to our customer on the expected delivery time that the customer chose based on shipping priority. If we aren't able to meet this at least 90% of our deliveries, then we will be losing our customer's trust and eventually lose them as potential returning customers. In order to combat this we need to know what is the source as to why we are having such a high percentage of late deliveries. So we went to look at different types of attributes, the main one being the different types of statuses of the orders.

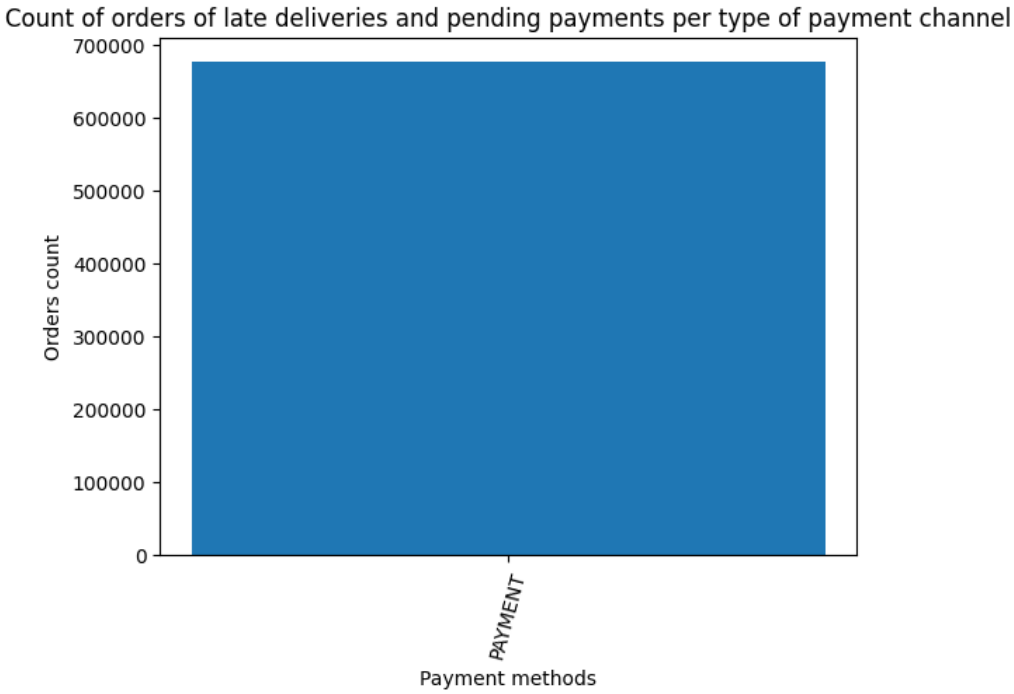
Percentage of Late deliveries from total per order status channel analysis:



Here we can see that most of late deliveries come from complete, pending, pending_payment, and processing order status. This tells us that we can make delivery improvements based upon these order statuses. We have the following opportunities for improvements based on the order status:

Pending_payment: we need a more efficent system of processing our payments from the different payment methods that our customers have. This could be outsourcing third-party tools that can probably process these

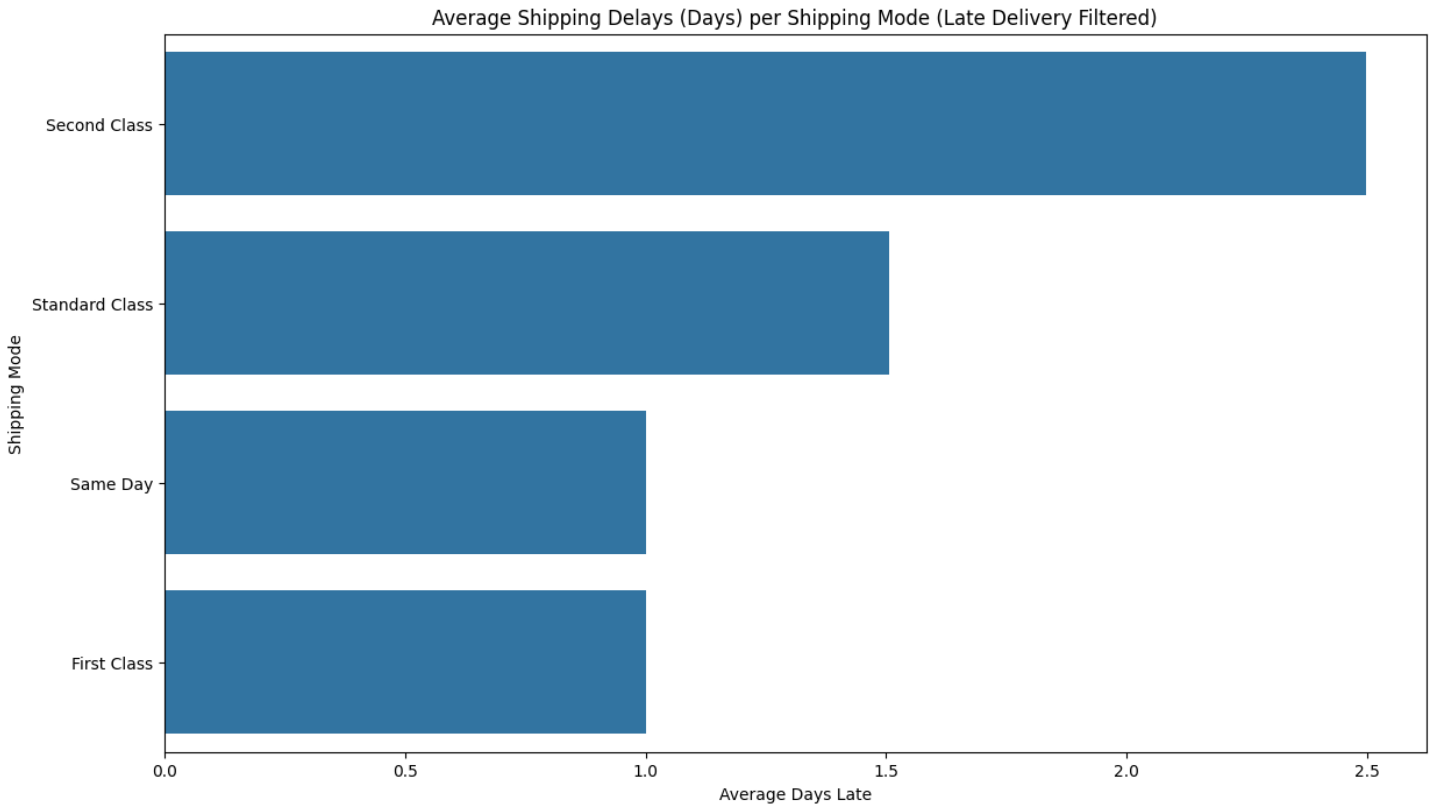
much faster. We can also introduce different varieties of payments methods, as we can see below we only see one type of payment method for those orders that have a status of Pending_Payment:



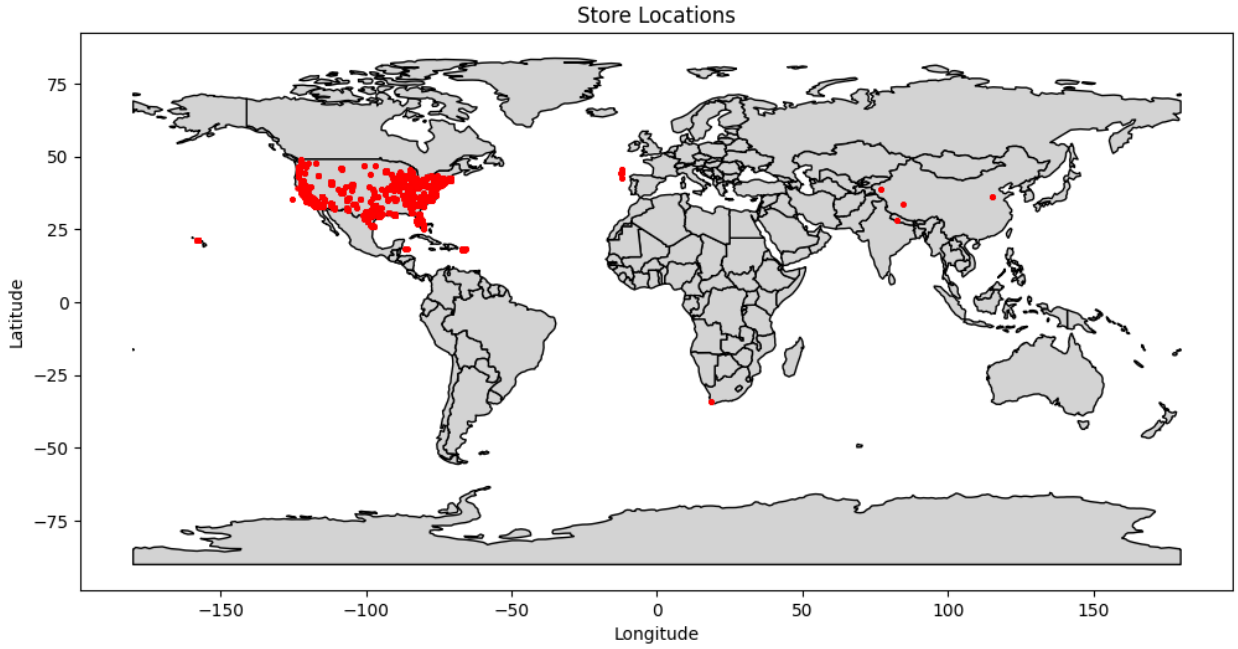
Processing: we can better our fulfillment centers and optimize them for faster fulfillment of the products that were ordered. If needed we can go ahead and do a fufillment product analysis to check which products are stuck in this status and we can focus on optimizing the process for those products. However, we need to collect that data in order to decide on what type of improvements to do under those processes.

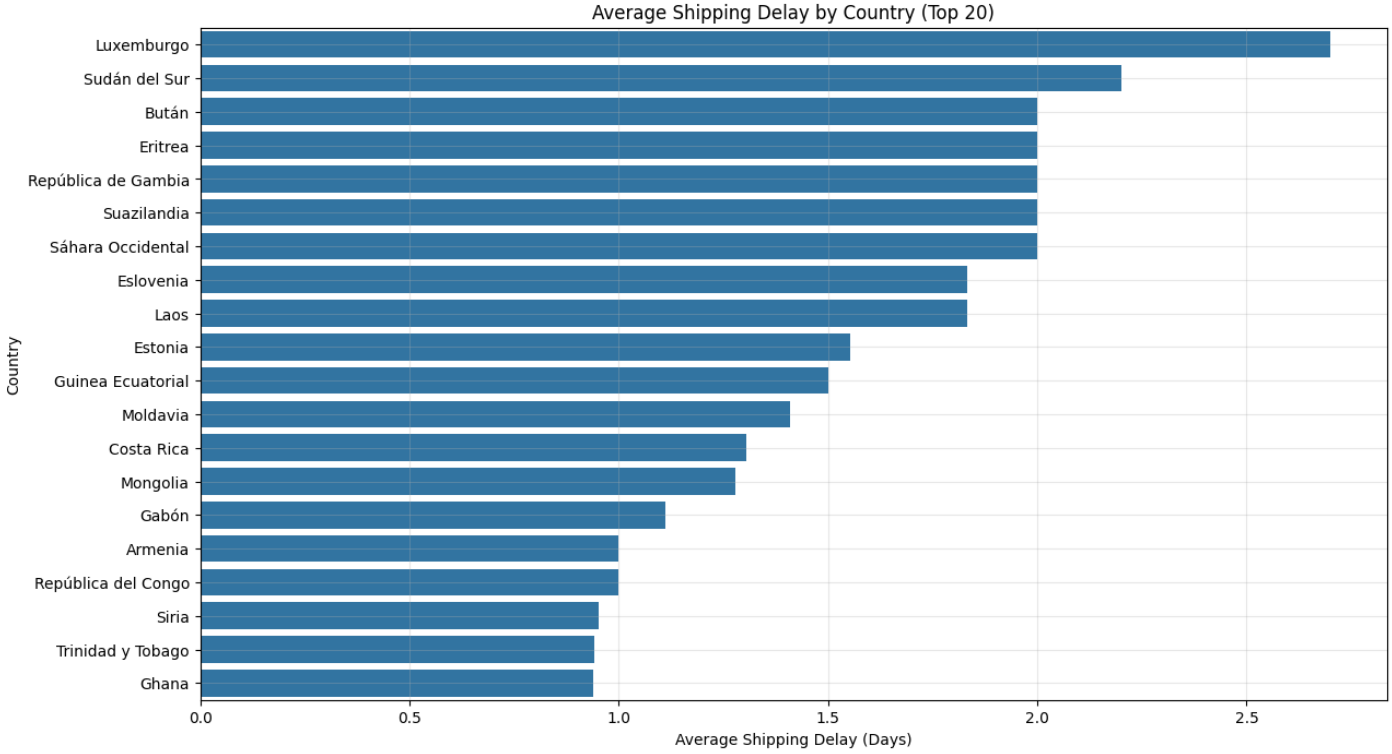
Pending: We need to better our processing for the clients details. This could be third-party tools in order to improve our data processing or we can look for ways to optimize our processes to be more efficient at processing our orders.

Complete/Closed Orders: we can do a product analysis and look at the top products with the highest numbers of days late and tailor our shipping dates to those numbers of days late (shipping method + days late = new expected shipping arrival). As seen the graph below, we aren't meeting the expectations of our expected delivery days to our customer per shipping mode channel. We can also see that the second class and standard class shipping modes are the one's with the highest variance of days from the expected delivery days. Opportunities that we can derive from this is to either search for a new shipping provider or change the expected delivery days for each shipping mode by adding the days below to each shipping mode's expected delivery days (shipping method + days late = new expected shipping arrival):



Store location and countries with top average delay days analysis:





We can see that most of our stores are located in the United States and the countries with the highest average shipping delay days are far in distance from those store. An opportunity that's available here is to consider

adjusting the expected shipping days for those countries or we can talk to the shipping provider to come to negotiations on shipping and logistics terms. Furthermore, we might reconsider doing business in these remote regions if it doesn't meet our profit margin quotas.

Conclusion:

After conducting extensive research and performing thorough data analysis, several potential solutions have emerged as we were gaining insights from companies data of its day-to-day operations. These solutions are based on a data-driven approach, ensuring that they are grounded in actionable insights derived from the analysis of customer behavior and the company's shipping logistics.

It is important to emphasize that the recommendations that are going to be made aim to improve the company's customer retention and the ultimate long-term growth of the company. By leveraging the powerful insights that data provides, we are able to explore potential solutions to the corporation's current operations and offer strategies to better the business model.

The following is a summary of potential solutions that could be implemented to improve the corporation's performance for the upcoming year of 2018:

- Consider creating a bundle of products. This could be the most frequently pair of product that are bought together. Then price the products at the fair price that optimizes the profits of the product. Then market the products from March to May, from wednesday to Sunday from 8-10AM
    - Consider on testing the products on product's popular regions, state, and cities first before making it an official launch.

- Consider on improving late delivery by any of the following:
    - Improving your processes for pending and processing payments. This could be third-party tools or finding improvements within our systems of how we process data and how we fulfill our orders. If we go this route we are going to have to survey our internal payments processes in order to collect data and do further indepth analysis.
    - We can increase our expected delivery days for each shipping mode by the average days late in order to make sure that we are being transparent and realistic with our customer. This should drive more retention of customers as we are meeting up to our expectations. We need to also gather more data regarding our shipping to see what areas we can improve upon. For example, we can improve shipping costs and delivery time. We would require a data set with different shipping providers, shipping costs of each shipping provider, the varieties of shipping priorities and delivery times.
    - We can include more payment options since we only have one type of payment option under pending_payment order status.

- We can also consider talking to our shipping providers about shipping and logistic terms or consider surveying our internal processing in order to optimize our shipping logistics.