

Amazon

1. Define

- **KPI:** An e-commerce website always aimed to entice clients with promotions and discounts. A three-day annual campaign is suggested to increase the company's sales. This study aims to identify three days throughout the year when the advertising would be most effective. In order for clients to benefit the most from the campaign and receive the greatest discounts, the research's goal is to identify the three days that will be the most successful in the future year.

CTQ: The results of this study will be helpful to the company and its customers. If the campaign is launched on productive days, both customers and the company would benefit from the discount offered by the campaign.

- **KPI:** The second analysis looks at the most and least popular products in each state wise. This research is required to optimize the company's warehousing because it will reveal which products are most and least in demand in each state, allowing the company to expand or decrease their warehousing of those products in those states.

CTQ: The Company would benefit from the study's findings since they could lower and better manage its warehousing costs. Customers would gain from this study as well because of the quicker and more precise delivery.

- **KPI:** The third analysis looks at the most frequently purchased items together in an effort to improve customer recommendation accuracy. Customers will receive recommendations when they choose a particular product and frequently purchase additional products along with it.

CTQ: Customers will gain the most from this analysis as they will receive better recommendations for which things to buy and how they might use the products they have chosen. Sales would undoubtedly increase, which would be beneficial for the company as well.

2. Measure

Amazon Sales Data 2019 is the source of the data used in this analysis. The academic institution "Asian Academy of Film and Television" has offered this data. The data include the following characteristics:

- **Order ID:** The order ID is the unique ID assigned to each placed order. (Note: If numerous products are ordered at once, their order IDs will all be the same.)
- **Product:** Name of the product ordered.
- **Quantity Ordered:** This tells the quantity ordered of that particular product in that particular order.
- **Price Each:** Price of each quantity ordered.
- **Order Date:** Date and time of the order placed.
- **Purchase Address:** The address of the buyer where the order would be delivered.

Upon inspection, it was discovered that the data contained missing and inconsistent values. In certain rows, the columns names or "NaN" were written in place of the missing data. When we looked more closely, we saw that when the data was missing, the issue affected the entire row. Therefore, deleting the entire missing rows is the best course of action when dealing with this kind of missing data.

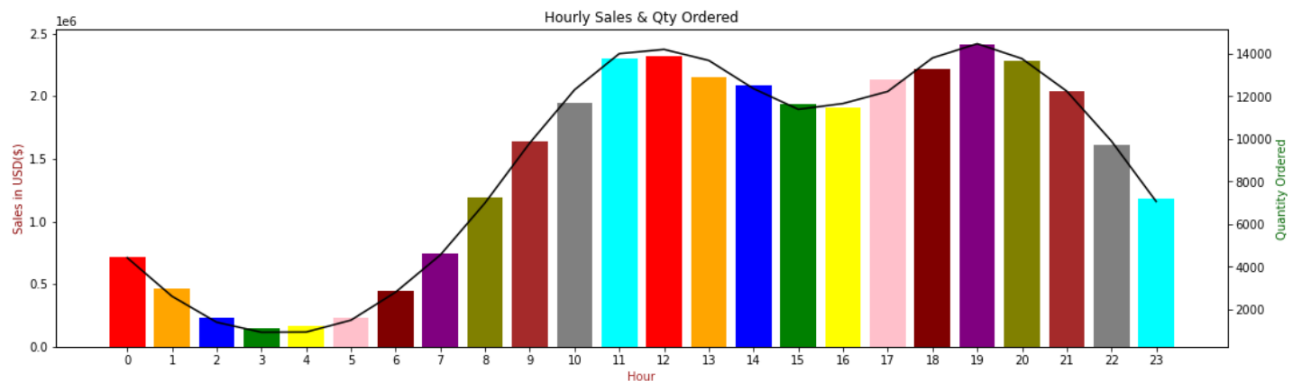
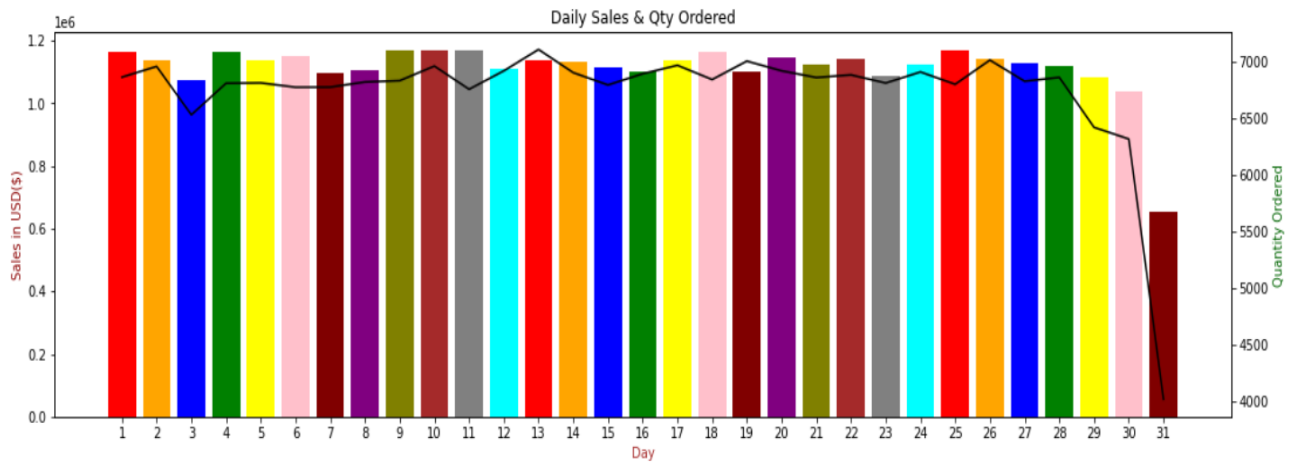
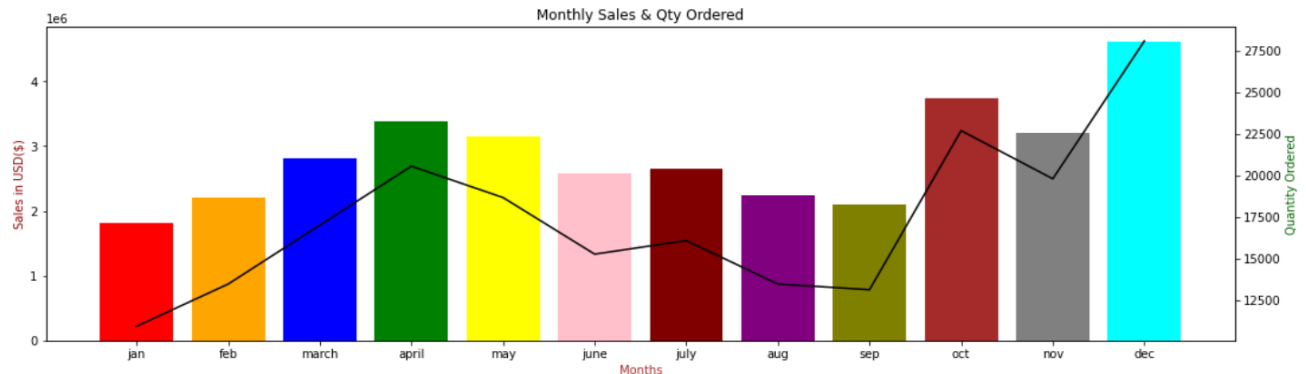
Following that, a few columns that were required for the analysis were added. These columns are listed below:

- **Sales:** This details the exact amount of revenue that particular order generated. It was created by multiplying the "Price Each" and "Quantity Ordered" columns.
- **Month:** Month of the ordered place. It was extracted from the "Order Date" column.

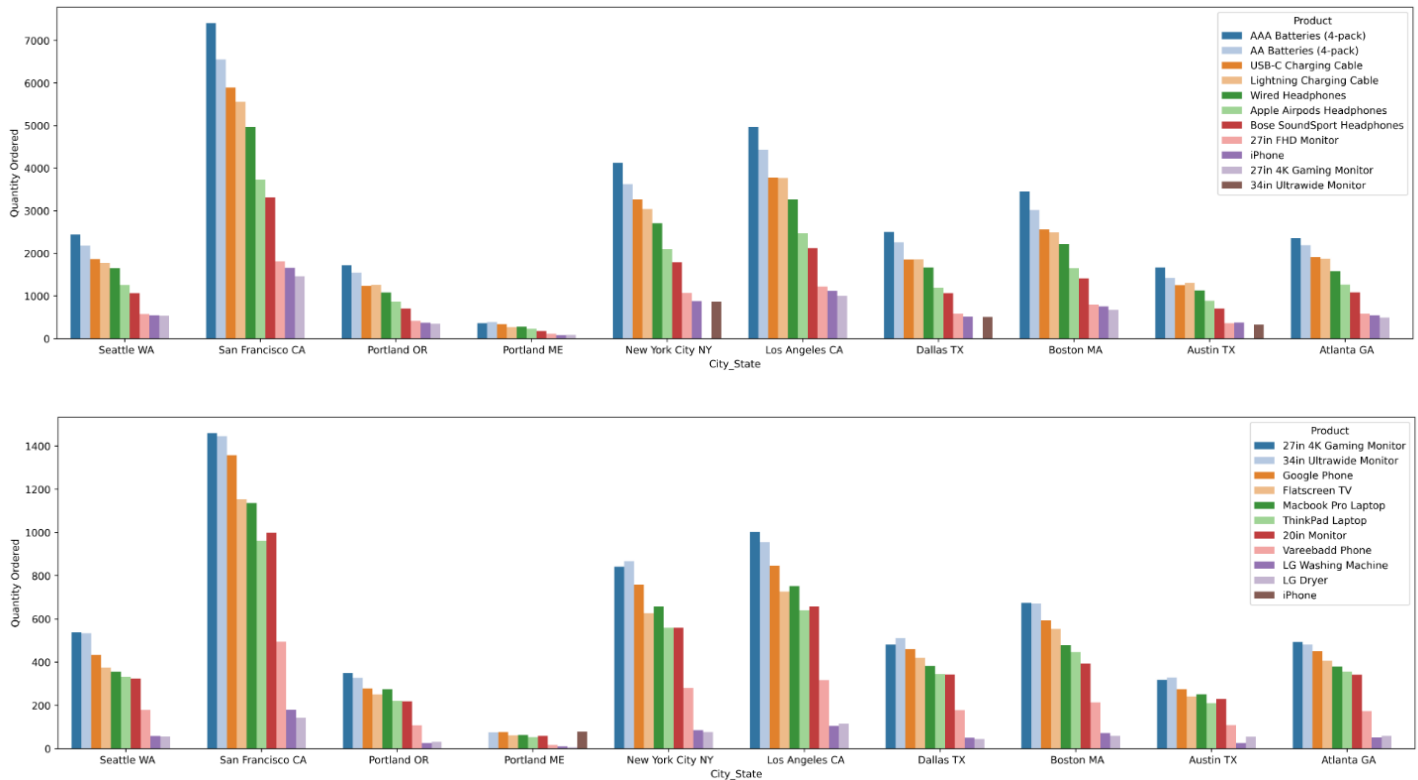
- Day: The day of the month. Again it was extracted from the “Order Date” column.
 - Hour: The particular hour at which order was place. “Order Date” column was used.
- Now, our data is ready to perform analysis to get our solutions for the mentioned problems.

3. Analyze

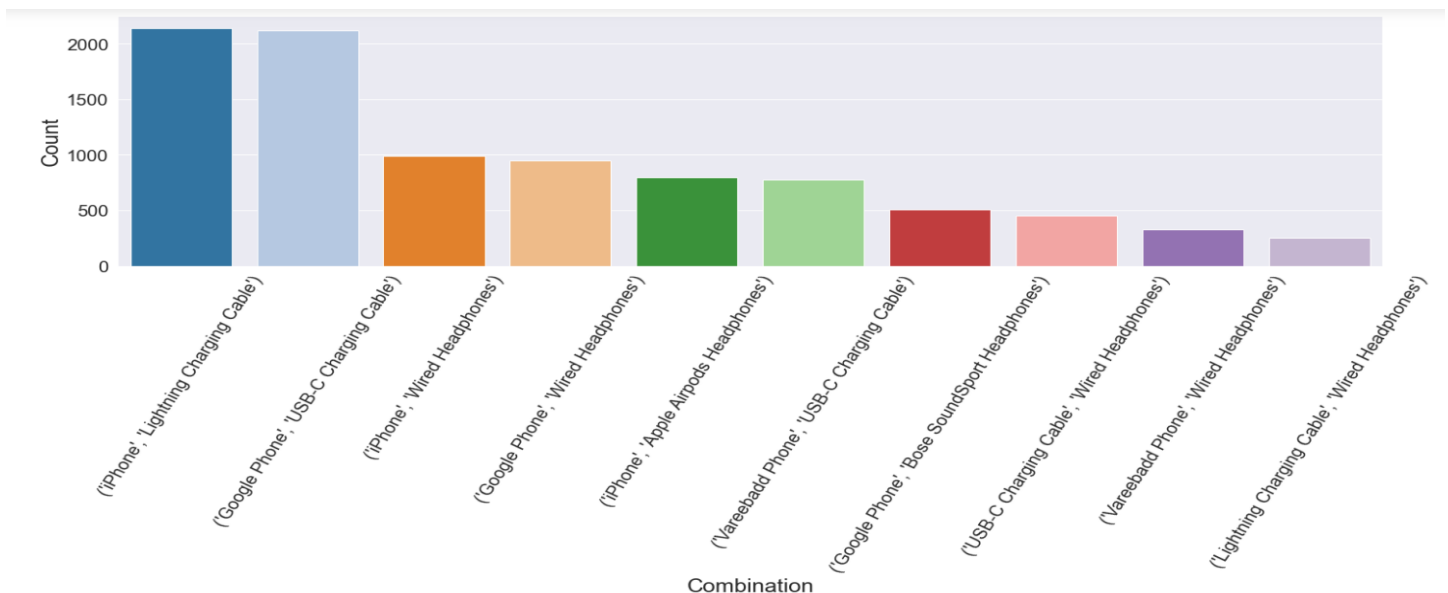
- For a three-day annual campaign, finding the most relevant month, day, and time was our first challenge. To check the month, day, and time with the most sales and quantity ordered, we have generated three graphs. The graphs show the sales (bars) and quantity ordered (line) on the y-axis and the relevant feature (Month, day or hour) on the x-axis. Here are those three graphs in order:

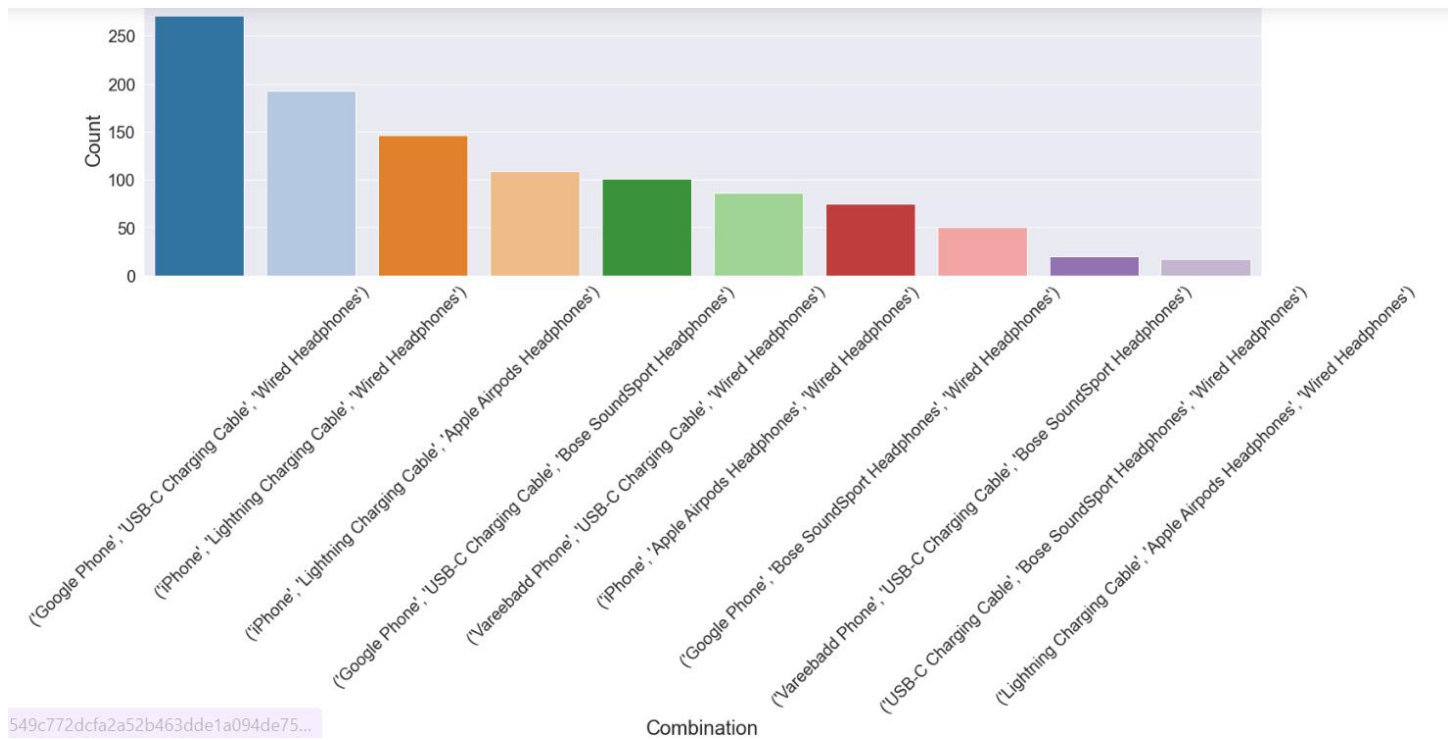


- The objective of the second problem statement was to address the management issue with warehousing. One may accomplish this by looking at the quantity of each product that was ordered individually in each state. For this, we have developed two graphs, one showing the most quantity of products ordered in each state, and the other showing the least quantity ordered in each state. The x-axis shows each state, and the y-axis shows the quantity ordered. These next two graphs show:



- The improvement of the customer recommendation system was the focus of our most recent problem statement. By comparing the same order id for many orders, we were able to compile a list of the most popular products. Two graphs have been created; the first shows the two products that are purchased together, and the second shows the three products that are purchased together. These next two graphs show:





4. Improve

- According to the first analysis graphs, December is the month, 25th is the date, and 19th is the hour with the highest volume of sales and orders. However, the 25th is the last of the month and customers' purchasing power is very low, so it is not a suitable day to start a campaign. Additionally, the sales decline immediately after the 25th, and the same is true for the hour, since the 19th hour is late at night and customers are more fatigued. Because sales have been increasing since the beginning of the month, the ninth day would be ideal for the campaign, and the eleventh hour would be ideal because sales have been increasing from the beginning of the day. So, the ninth of December at 11 a.m. would be the suggested day to start a successful three-day campaign.
- According to the graphs from the second analysis, eleven products are identified as being ordered most and least in each state. Given that San Francisco has a very high demand, the company can establish a large warehouse there with a high volume of the product with higher demand (as seen in the first graph) and a relatively low volume of the product with lower demand (as shown in the second graph). Furthermore, given that Portland has a relatively low demand, the company can establish a combined warehouse for several states based on with their locations. The demand graphs for each state can be used to manage the stock in those warehouses, keeping large volumes of goods for higher demand and low volumes of stock for lower demand.
- An intriguing finding from the third analysis graph is that customers who buy a "iPhone" or "Google Phone" also buy a "USB-C Charging Cable", "Lightning Charging Cable", "Wired Headphones", or "Apple AirPods Headphones". So, whenever a customer buys an iPhone or a Google phone, we may recommend all of these things to them.