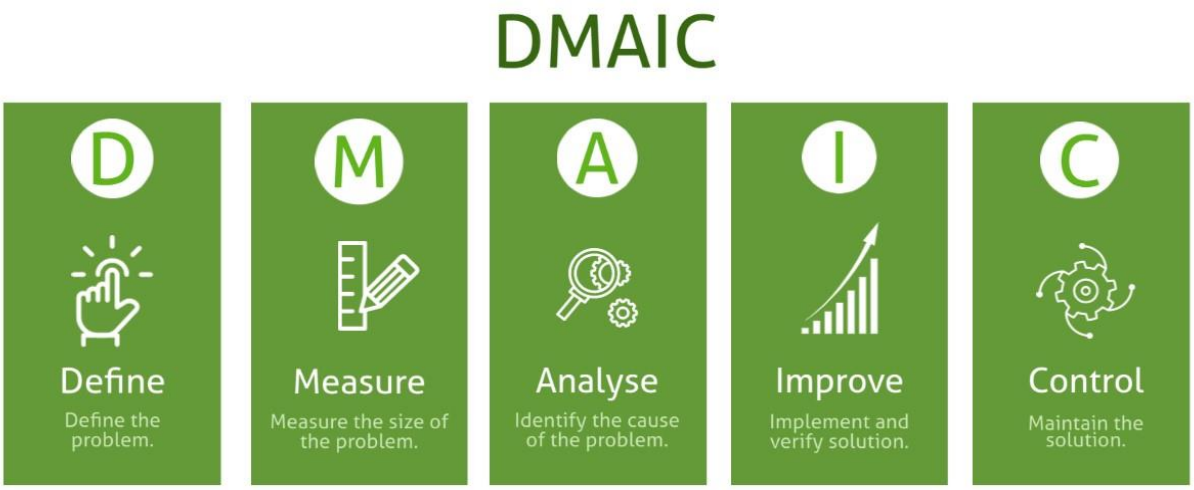
DMAIC Of Amazon Sale Prediction

What is DMAIC?

DMAIC is an acronym for Define, Measure, Analyze, Improve, and Control. It is a guideline for delivering sustainable and quantifiable results



Define

As per DMAIC, the first and the most important step is the Define phase which is important for knowing areas of improvement. This project was created for resolving our three mains problem statements:

1. Release data of our sale: - Selecting the right launch date for a sale is crucial for maximizing customer lifetime value (CLV). By considering factors like past customer behavior, peak buying periods, and competitor activities, businesses can strategically plan sales to attract more customers and drive higher profits. Market trends, upcoming events, and holidays should also be considered to leverage customer preferences and enhance engagement. Additionally, segmenting customers and aligning launch dates with their preferences enables targeted promotions and increased sales effectiveness. Remember, while launch dates are important, an overall CLV strategy should also

prioritize exceptional customer experiences, loyalty-building efforts, and nurturing long-term customer relationships.

KPI: - The Key Performance Indicators (KPI) are the critical success factors that are used to overcome the problem statement, so the KPI in this case are Price each, Sales, Quantity ordered, Month name, Day, Time.

CTQ (Critical to Quality): - The results of the study will be advantageous to the organization and its customers. If the promotion started on a day when business was brisk, both customers and the company would benefit from the discount it offered.

2. Supply Chain Optimization: - The goal is to determine the sales of the most and least demanded products on a state-wise basis for supply chain optimization. This information will provide insights into the need for warehousing adjustments in specific regions. By analyzing product sales, we can identify whether to increase or decrease warehousing capacity in a particular state. This optimization strategy enables cost savings by avoiding the construction of redundant warehouses in regions with concurrent demand. By accurately assessing product demand, we can ensure that a single warehouse is sufficient to meet the needs of multiple regions, effectively streamlining the supply chain.

   KPI: - The KPI in this case are City States, Product, Quantity Ordered.

   CTQ: - This study's results would be advantageous to the company because it would give them an idea to reduce and manage its warehousing costs. This study would be beneficial to customers as well because it would result in faster and more accurate delivery.

3. Recommendation and Association: - Building a recommendation system involves analyzing customer behavior to identify frequently co-purchased products. By collecting transactional data and applying market basket analysis techniques, businesses can generate association rules that represent product relationships. Using the association rules, personalized recommendations can be generated for customers, suggesting items that are likely to be purchased together.

KPI: - The KPI in this case are Order Id, Product, Quantity Ordered.

CTQ: - Customers stand to benefit most from this analysis since they will get better advice on what to buy and how to use the products they have selected. Sales would surely rise, which would be good for the business as well.

Measure

Amazon is one of the world's biggest E-commerce websites which deals with a humongous amount of data. The amazon sales 2019 dataset on which we are doing our analysis is a mere part of that big data. The dataset was provided to us by our institution Skill Circle. The dataset contains these following

columns: -

• Order Id: - The unique Id provided to the purchase a customer made from the website. The order id is assigned based on purchase, if a customer purchases multiple items at a single time their order id would be same.

• Product: - The name of the product which had been purchased by the customers.

• Quantity ordered: - This column tell us the number of units of the product which had been purchased

• Price each: - Price of each product

• Order date: - date on which product was ordered

• Purchase address: - address of the customer who purchased a product from the website

These were the default columns available in our dataset. During our analysis we created and added a few more columns to the dataset which were essential for our project. The columns that we created were: -

• Sales: - This column tells us the amount of revenue a particular order has given. It was calculated by multiplying the values of quantity ordered and price each.

• Month: - The name of the month on which the order was placed. It was extracted from order date column

• Day: - The day of the month on which the order was placed. It was also extracted from order date column

• Hour: - The time of the day at which the orders are placed. It was also extracted from the order date column.

• City & State: - The City and the States from where the ordered the product.

On the basis of these following columns, we are going to do our analysis.

Analyze

1. We have created three graphs to determine the best time for a three-day annual sale. The first graph analyzes monthly sales and quantity ordered, highlighting peak months. The second graph focuses on sales and quantity ordered by day of the week, helping identify the most favorable days for the sale. The third graph examines hourly sales, revealing the most profitable time slots. The order of these three graphs is as follows:

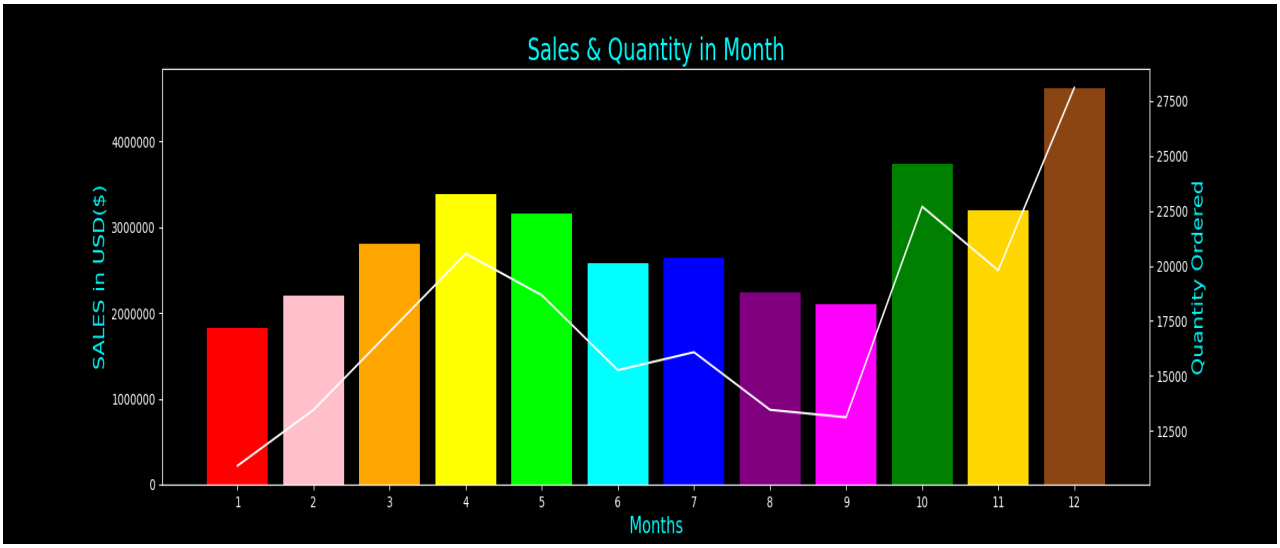


FIG. 1.1

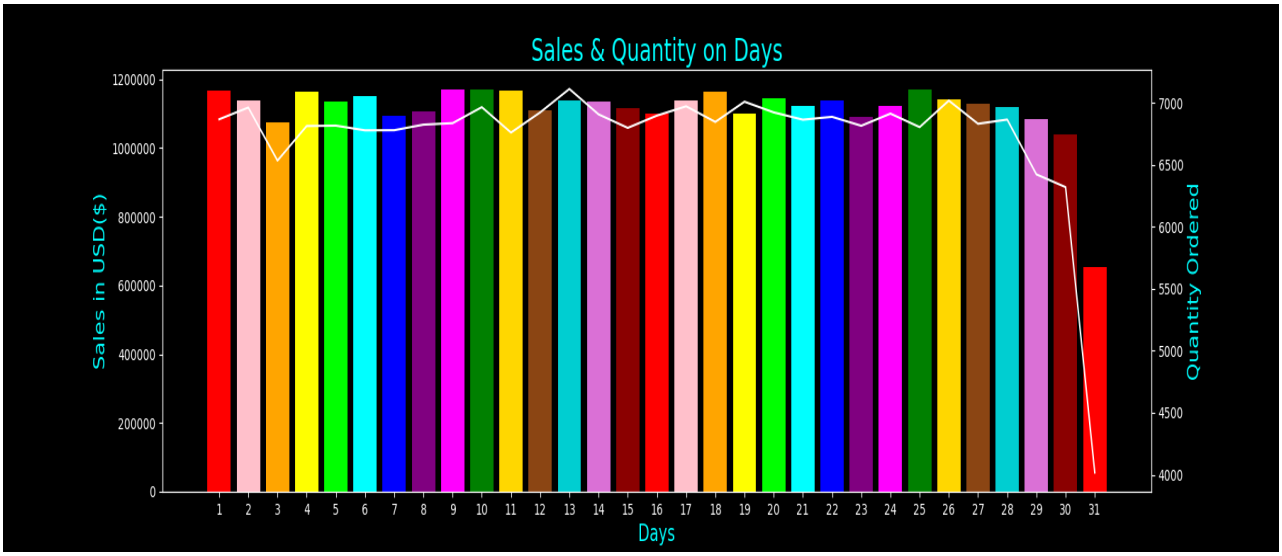


FIG. 1.2

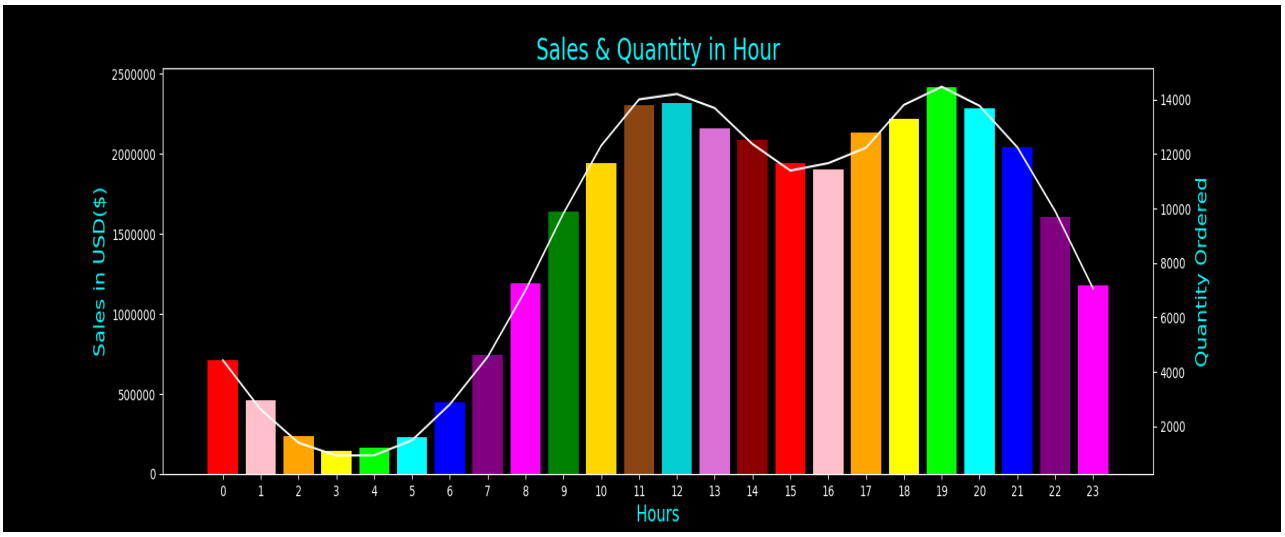


FIG. 1.3

2. The second problem statement focused on managing warehousing difficulties. To tackle this, we analyzed the quantity of individually ordered products in each state. Two graphs were created: one displaying the states with the highest quantity of products ordered, and the other highlighting the states with the lowest quantity. The x-axis represents the states, while the y-axis represents the quantity requested. These graphs provide valuable insights into state-specific product demands, aiding in effective inventory allocation and

distribution management. The following two graphs display: **No table of contents entries found.**

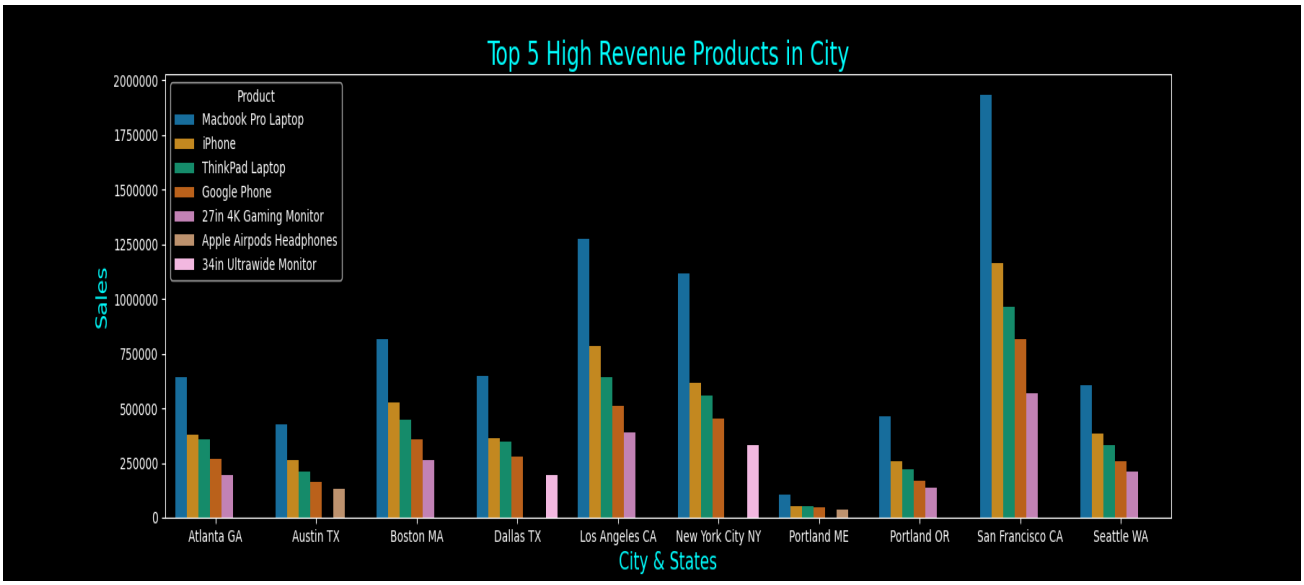


FIG. 2.1

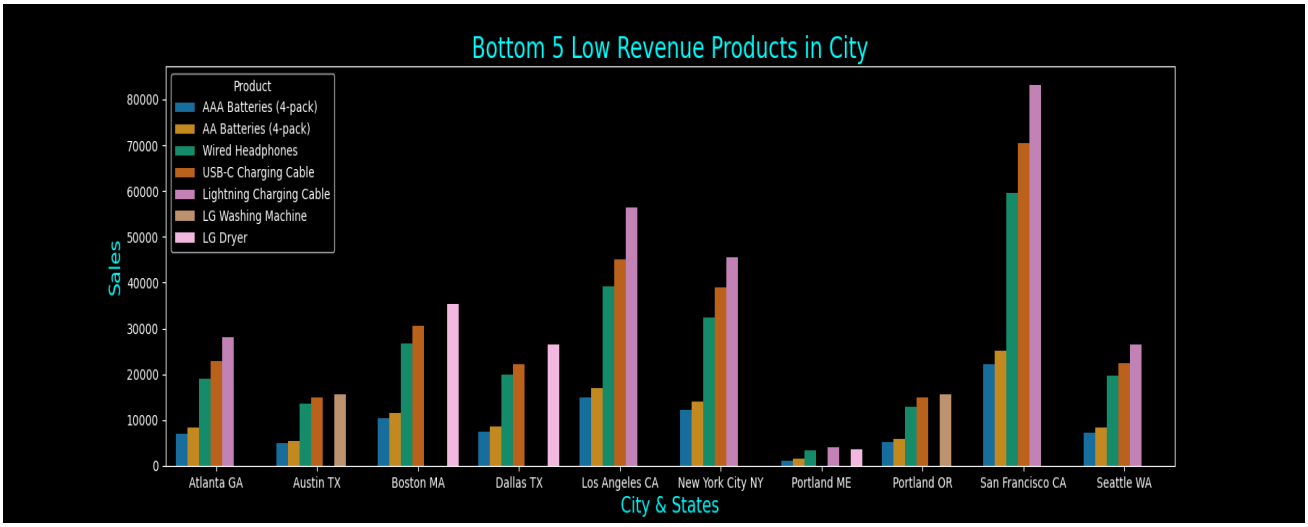


FIG. 2.2

3. In our latest problem statement, we aimed to improve the consumer recommendation system. By comparing order IDs from multiple orders, we compiled a list of the most popular products. To visualize this information, two graphs were created. The first graph illustrates the purchases of two specific items together, while the second graph showcases the purchases of three products together. These graphs provide insights into frequently co-purchased products, enabling the recommendation system to offer more accurate and relevant product suggestions. The following two graphs display:

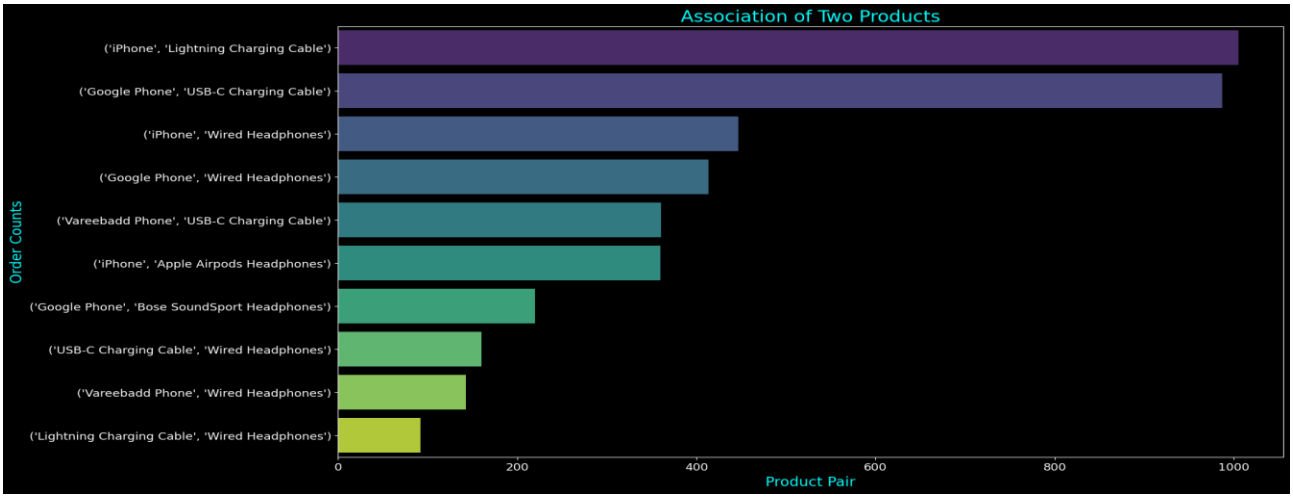

Association of Two Products
('iPhone', 'Lightning Charging Cable')
('Google Phone', 'USB-C Charging Cable')
('iPhone', 'Wired Headphones')
('Google Phone', 'Wired Headphones')
('Vareebadd Phone', 'USB-C Charging Cable')
('iPhone', 'Apple Airpods Headphones')
('Google Phone', 'Bose SoundSport Headphones')
('USB-C Charging Cable', 'Wired Headphones')
('Vareebadd Phone', 'Wired Headphones')
('Lightning Charging Cable', 'Wired Headphones')
Order Counts
0
200
400
600
800
1000
Product Pair

FIG. 3.1

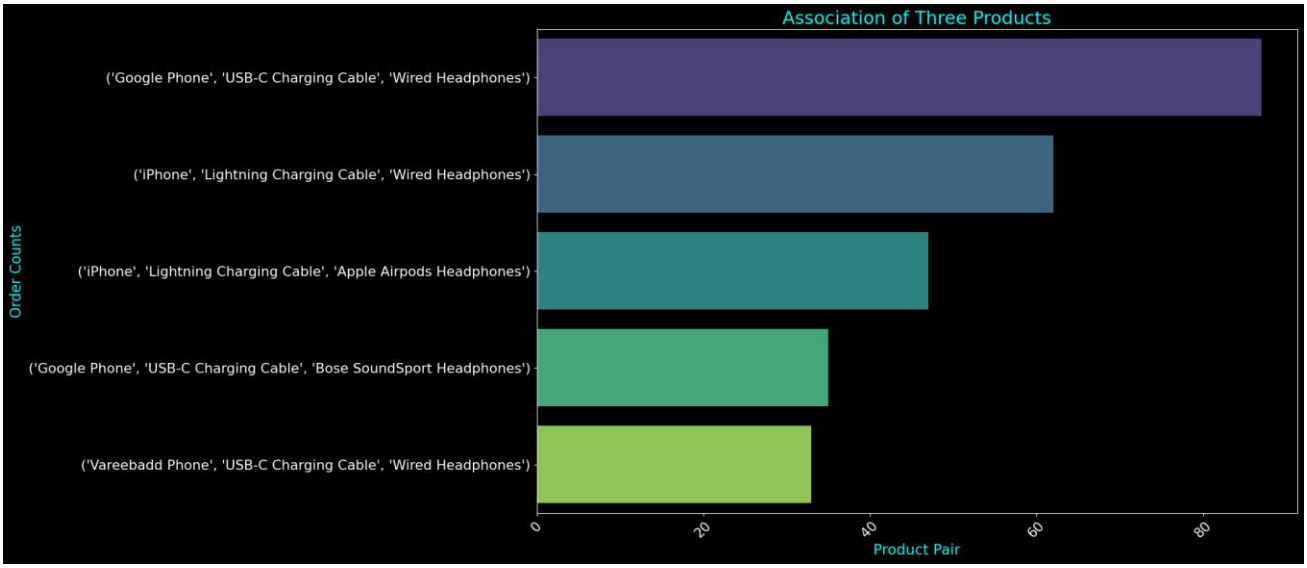

Association of Three Products
('Google Phone', 'USB-C Charging Cable', 'Wired Headphones')
('iPhone', 'Lightning Charging Cable', 'Wired Headphones')
('iPhone', 'Lightning Charging Cable', 'Apple Airpods Headphones')
('Google Phone', 'USB-C Charging Cable', 'Bose SoundSport Headphones')
('Vareebadd Phone', 'USB-C Charging Cable', 'Wired Headphones')
Order Counts
0
20
40
60
80
Product Pair

FIG. 3.2

Improve

- Based on the above analysis, the following are the findings and decisions regarding the best time, day, and month for releasing the sale:

  - The instances of highest sales and orders were observed at 7 PM, on the 25th of December, in December overall.
  - Releasing the sale on the 25th of December is not recommended as it precedes Christmas, leading to lower sales after that date.
  - To capitalize on the increased orders and sales before Christmas, the sales will be released before Christmas, specifically on the 23rd of December.
  - The sale will span three days, ending on the 25th of December.
  - Releasing the sale at 7 PM is not optimal due to customers potentially being exhausted by that time.
  - Instead, the sale will be launched at 10 AM, providing a fresh start to the sale and attracting more customer engagement.

  In summary, the revised plan for the sale release is as follows:

  - Date: December 23rd to December 25th
  - Time: Sale launch at 10 AM to maximize customer engagement and provide a strong start to the sale.

- Based on the analysis of the second set of graphs, the following observations and decisions can be made regarding warehousing and product quantity management:

  - The graphs showcase the seven products that are most and least frequently ordered in each state.

- Strong demand is observed in San Francisco, indicating the need to establish a sizable warehouse in that location.
- The warehouse in San Francisco should stock a larger quantity of products with high demand (as seen in the FIG. 2.1 graph) to meet customer needs.
- Additionally, the warehouse should also store a sufficient quantity of products with lower demand (as shown in FIG. 2.2 graph) to cater to a broader range of customer preferences.
- To optimize warehouse distribution, a consolidated warehouse can be set up for multiple states based on their locations.
- Portland exhibits very low demand, suggesting a need for minimal inventory in that region.
- Managing product quantity based on demand is crucial. Higher demand requires increased supply and a larger stock of the corresponding products, while lower demand requires reducing the quantity of products to save costs.

In summary, the recommended actions based on the analysis are as follows:

- Establish a sizable warehouse in San Francisco, stocking a larger quantity of products with high demand and a sufficient quantity of products with lower demand.
- Consider setting up a consolidated warehouse for multiple states, while adjusting inventory based on demand.

- Based on the analysis of the third study graph, the following observations and recommendations can be made regarding customer purchasing patterns:

  Customers purchasing an "iPhone" or "Google Phone" often also purchase the following items:

  - "USB-C Charging Cable"
  - "Lightning Charging Cable"
  - "Wired Headphones"
  - "Apple Air Pods Headphones"

  These findings suggest a strong association and cross-selling potential between these products.

In summary, based on the analysis, the recommendation is to suggest items such as the "USB-C Charging Cable," "Lightning Charging Cable," "Wired Headphones," and "Apple Air Pods Headphones" to customers who purchase an "iPhone" or a "Google Phone" to drive additional sales and enhance the customer's shopping experience.