Analysis of Shipping Data

Behavior For E-Commerce Portal

**Guided By-Dr Ankur Jain**

Ajay Singh, Isha Kondurkar,Preet Sharan, Kunal Singh, Apoorva Kharya

***Abstract -*** Data analysis has a role to produce a decision. The problem to be investigated in this study is the analysis of the sales of the supermarket sales data set using data visualization, using various techniques through Tableau. The research method used is qualitative methods in the form of case studies. Based on research that has been conducted on supermarket sales data sets, information is obtained from several data visualization methods.   
***Keywords –*** *Tableau, Visualization, Analysis, E-Commerce, Shipping Data*

# I. INTRODUCTION

Data analytics reveals how customers interact with a website, what their preferences are and their favorite brands. New doors open when data is analyzed. Consumers’ needs and wishes are hidden behind numbers, so there is tangible proof of what they really want. When a company is equipped with the facts, transformation occurs in the form of new product launches and building a brand around market demands. Loyal clients are often the result of a well-thought-out data analytics strategy. Because data sets help a company get to know their customers better, they will be able to cater to their needs more efficiently.

No matter what the business or career, data visualization can help by delivering data within the most efficient way possible. Together with the essential steps in the business intelligence process, data visualization takes the data, models it, and delivers the info so that conclusions can be reached. In advanced analytics, data scientists are creating machine learning algorithms to compile essential data into visualizations that are easier to understand and interpret. Specifically, data visualization uses visual data to speak information in a manner that is universal, fast, and effective. This practice can help companies identify which areas have to be improved, which factors affect customer satisfaction and dissatisfaction, and what to try with specific products (where they should go and who they should be sold to). Visualized data gives stakeholders, business owners, and decision-makers a far better prediction of sales volumes and future growth. The significance of the data visualization:

**About Dataset –** We have taken the dataset of an International based e-commerce company. We want to discover key insights from their customer database. The company sells electronic products and they want to use some of the most advanced machine learning techniques to study their customers.

This is a fictional e-commerce dataset created for helping the data analysts to practice exploratory data analysis and data visualization. We have taken the dataset of an International based e-commerce company. We want to discover key insights from their customer database. The company sells electronic products and they want to use some of the most advanced machine learning techniques to study their customers.

The dataset used for model building contains 10999 observations and 12 variables. i.e.  
Number of rows: 10999  
Number of columns: 12

**Applications -**

a. Without data visualization, it's challenging to identify the correlations between the relationship of independent variables. By making sense of these independent variables, we will make better business decisions.

b. While this looks like an obvious use of data visualization, it's also one of the most valuable applications. It’s impossible to form predictions without having the necessary information from the past and present. Trends over time tell us where we were and where we will potentially go.

c. Closely associated with trends over time is frequency. By examining the speed , or how often, customers purchase and once they buy gives us a better feel for how potential new customers might act and react to different marketing and customer acquisition strategies.

d. Data visualization takes the knowledge from different markets to give you insights into which audiences to focus your attention on and which ones to stay away from. We get a clearer picture of the opportunities within those markets by displaying this data on various charts and graphs.

e. Watching value and risk metrics requires expertise because, without data visualization, we must interpret complicated spreadsheets and numbers. Once information is visualized, we will then pinpoint areas that may or may not require action.

f. The power to obtain information quickly and easily with data displayed clearly on a functional dashboard allows businesses to act and respond to findings swiftly and helps to avoid making mistakes.

**Focus and Purposes –** We aim to analyze and examine the dataset through the data visualization tool Tableau, without lengthy codes and much effort. Our purpose is to understand sales, relationships in between, trends, and much more.With the help of given data we are going to evaluate which gender has more contribution in :

a. Using facility of Customer care calls

b. Doing more Prior Purchases

c. Getting More Discount Offered.

d . Providing more customer ratings

**Tableau:** As the market-leading choice for modern business intelligence, the Tableau platform is known for taking data from almost any system and turning it into actionable insights with speed and ease. It’s as simple as dragging and dropping. Plus, our industry-leading enablement resources, training, and global data community offer unparalleled support for our customers and their analytics investments. And on our mission to help people see and understand data, we go beyond our technology to ensure customer success by assisting people to build a data culture.

# II. PROBLEM FORMULATION

# 1) With the help of given data we are going to evaluate which gender has more contribution in :

# a. Using facility of Customer care calls

# b. Doing more Prior Purchases

# c. Getting More Discount Offered

# d . Providing more customer ratings

# 2) Evaluate the indirect impact on customer rating based on the mode of shipment and warehouse block and show:

a. Which warehouse block is returning highest cost of the products

b. Mode of shipment return max rating points.

3. Find which mode of shipment is delivering the maximum weight from each warehouse.

4. With the help given data we are going to analyze the relation between product importance ,mode of Shipment and the weights of the products we are shipping.

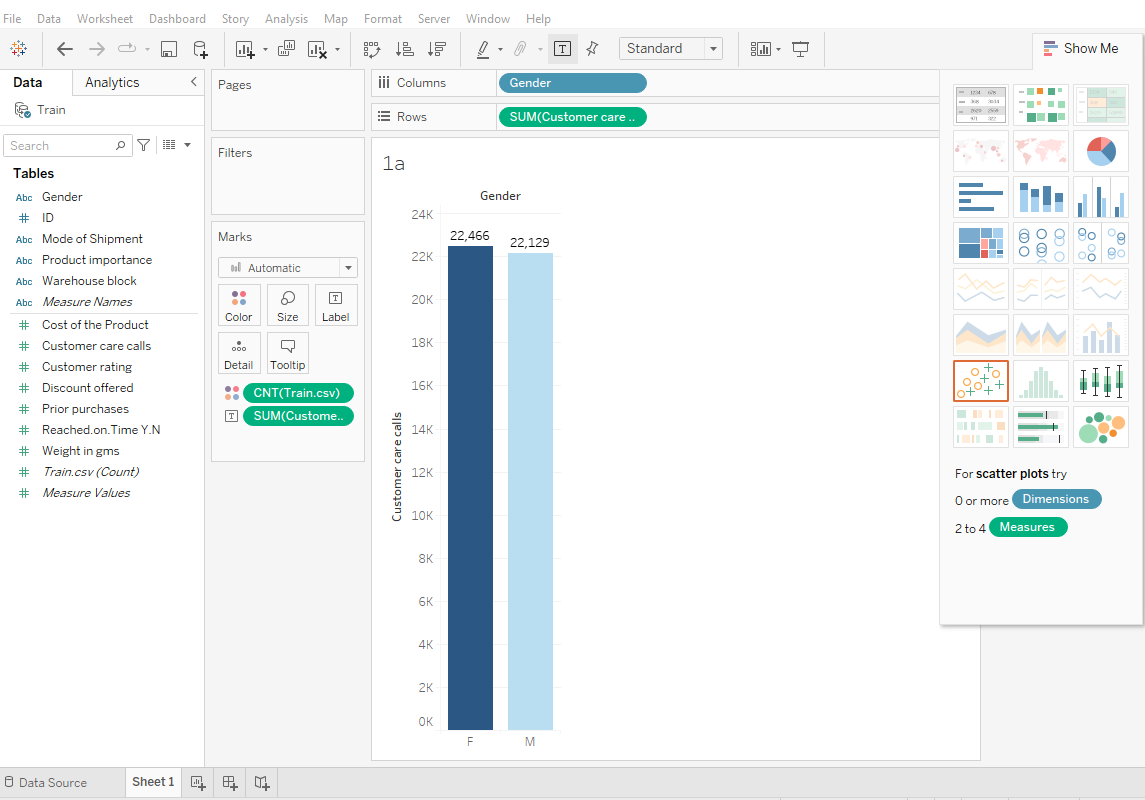
# III. METHODOLOGY

**A. Design -** By contrast, visualization is more intuitive and meaningful, and it is very important to use appropriate charts to visualize data. We are going to use different types of charts and graphs for the design purpose and visualization. Visual design aims to improve a design’s/product’s aesthetic appeal and [usability](https://www.interaction-design.org/literature/topics/usability) with suitable images, typography, space, layout and color. Visual design is about more than aesthetics.

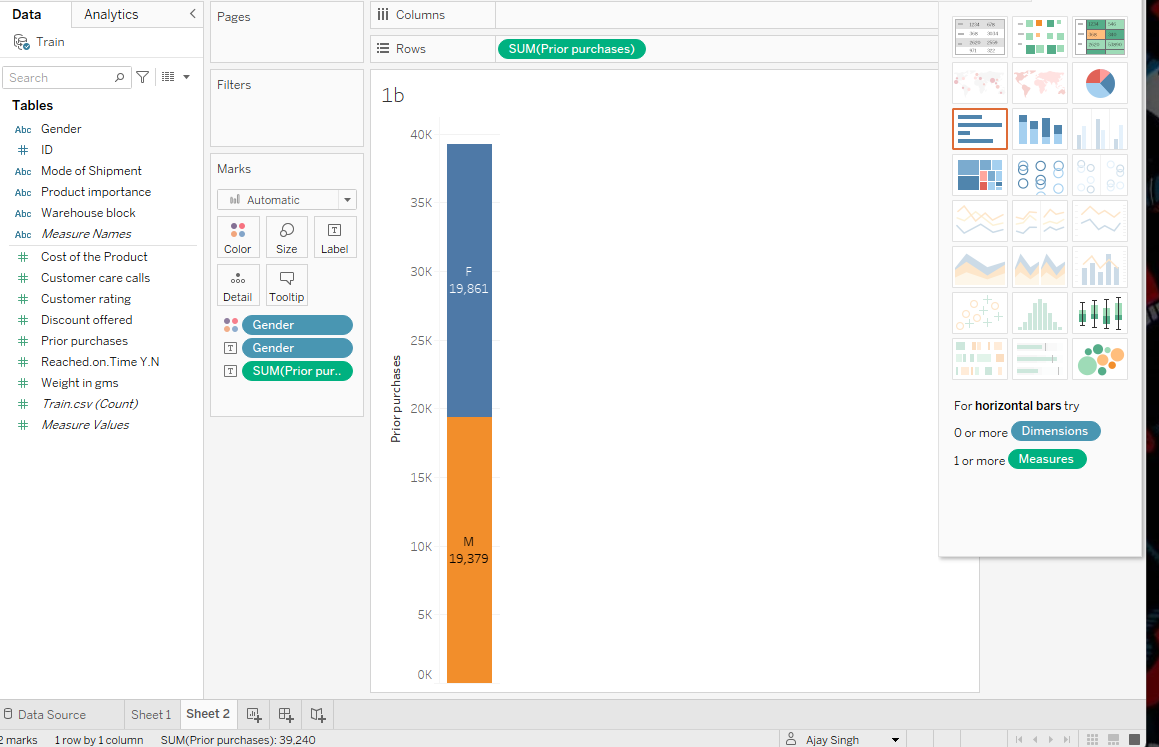
**B. Analysis -**

1. Parameters we consider for Analysis are -
   1. ID: ID Number of Customers.
   2. Warehouse block: The Company has a big Warehouse divided into blocks.
   3. Mode of shipment: Shipping in multiple ways such as Ship, Flight and Road, etc.
   4. Customer rating: It depicts the customer rating which shows 1 is the lowest (Worst), 5 is the highest (Best)
   5. Cost of the product: It depicts the Cost of the product in USD.
   6. Prior purchases: It shows the Number of Prior Purchases made by the customers regarding a particular product..
   7. Gender: Male and Female.
   8. Discount Offered: It depicts the Discount offered on a specific product.
   9. Weight in gms: It is the weight in grams.
   10. Reached on time: It is the target variable, where 1 Indicates that the product has NOT reached on time and 0 indicates it has reached on time.
2. Some observations that we made while analyzing the dataset through Tableau were -
   1. We see that the females use this facility slightly more than the males. They are also the ones making more prior purchases.
   2. Ship F mode of shipment returns maximum rating points and returns the highest cost of the products.
   3. Product importance, mode of shipment and the weights of the products have a very interdependent relationship.

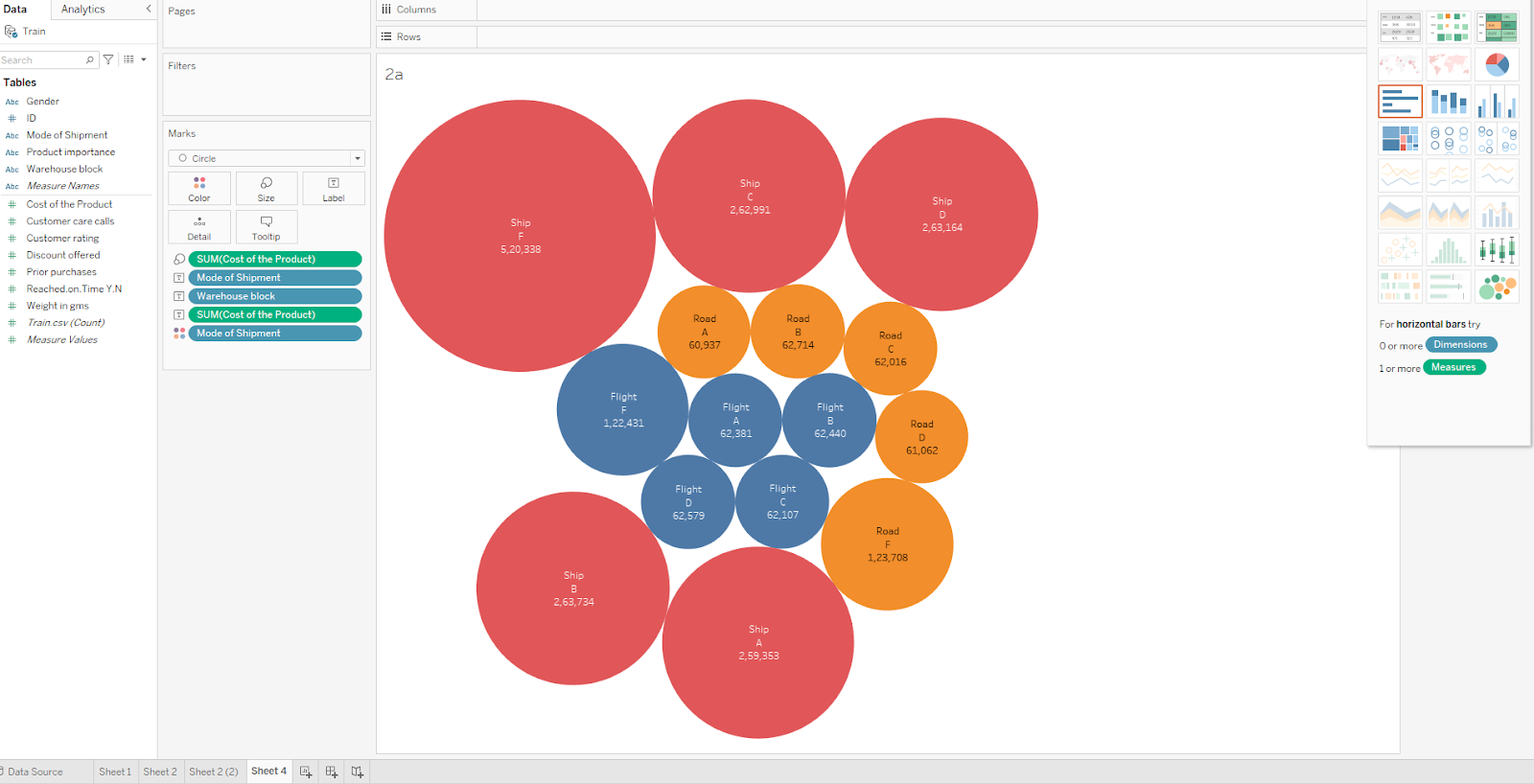
**C. Implementation:** The analysis was practically performed through Tableau.



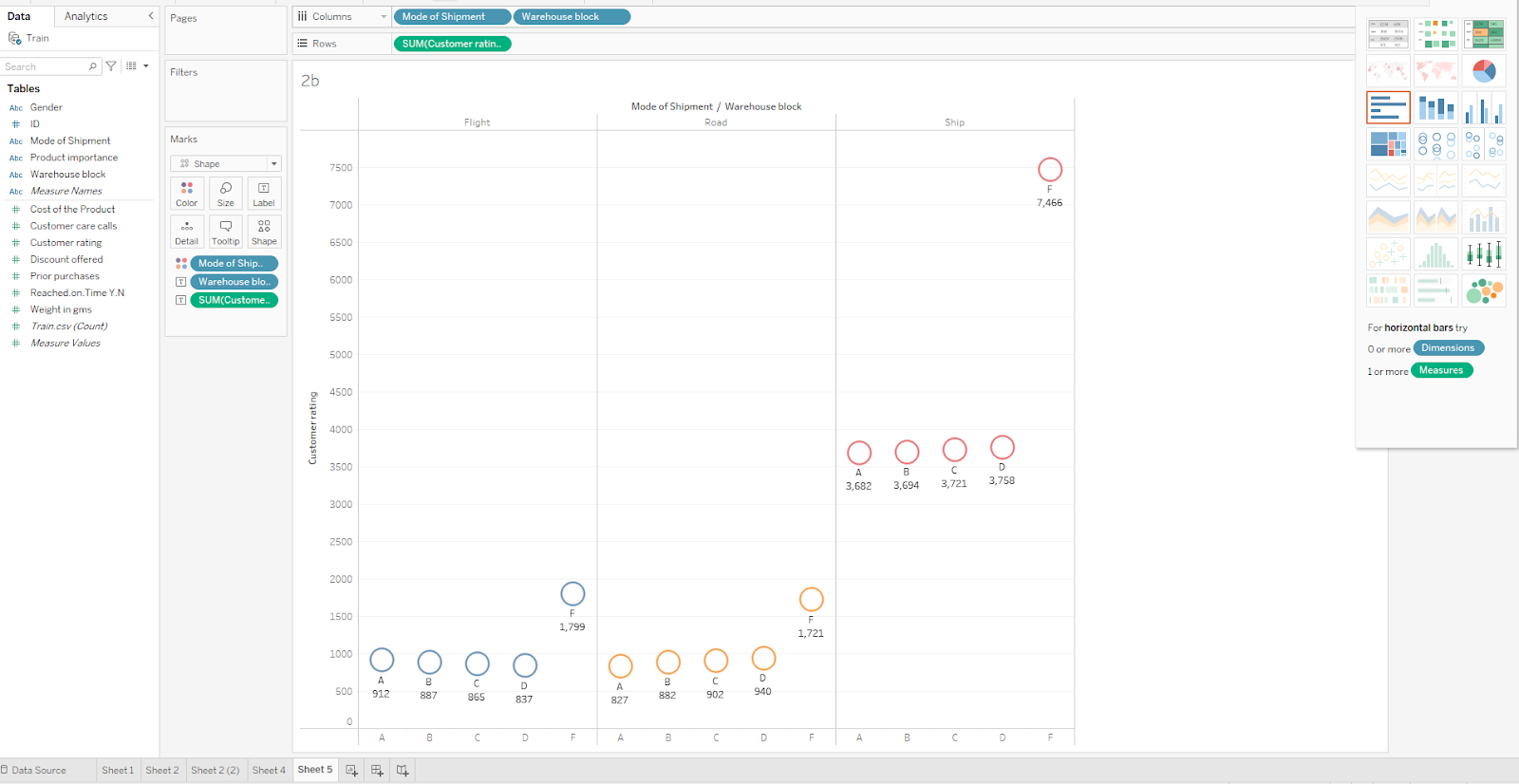
Here we evaluate which gender has more contribution in using the facility of Customer care calls. We see that the females use this facility slightly more than the males.



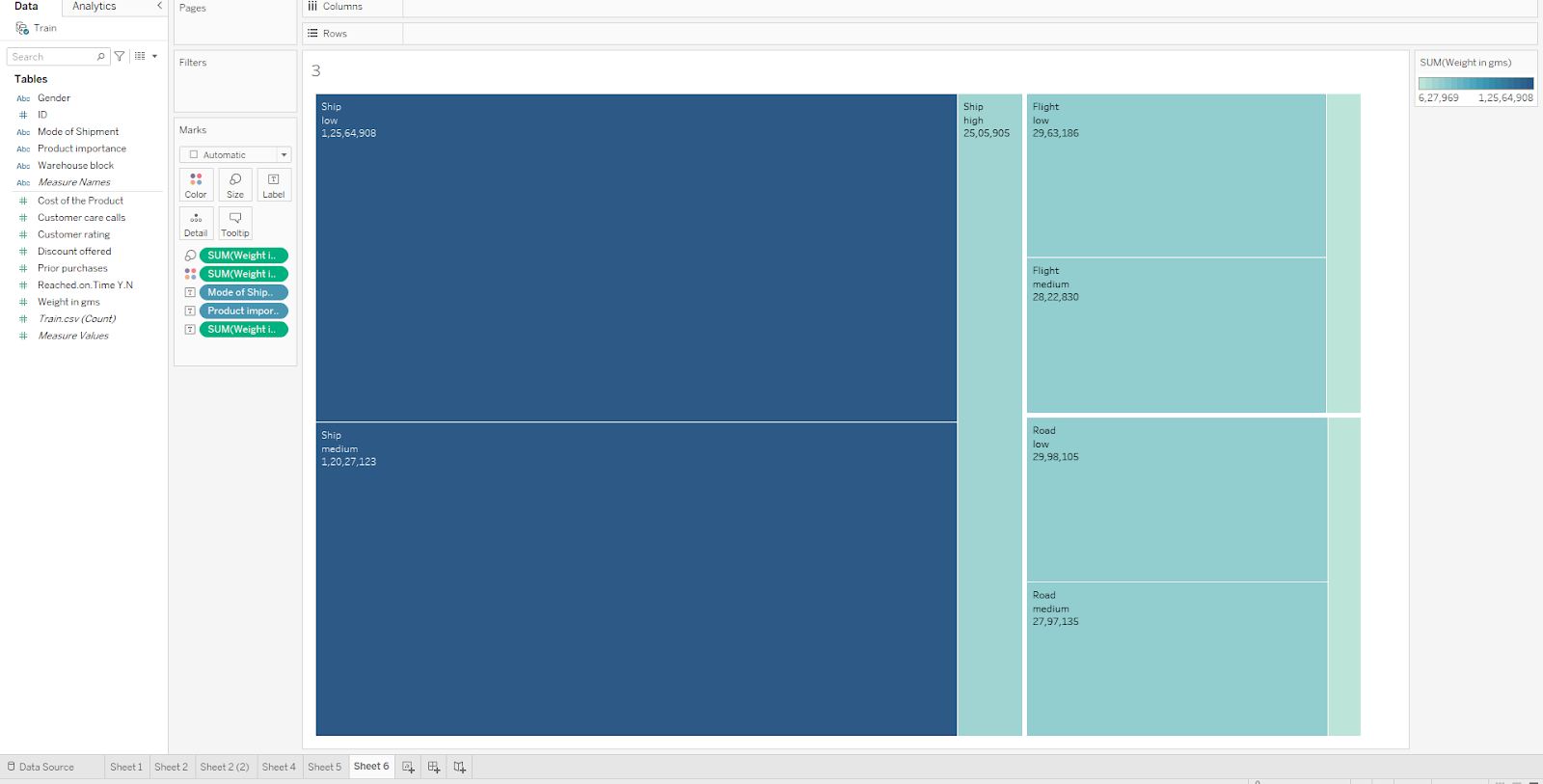
Here we evaluate which gender has more contribution in doing more Prior Purchases. We see that the females do more prior purchases as compared to the males.



Here we evaluate the indirect impact on customer rating based on the mode of shipment and warehouse block. Thus Ship F returns the highest cost of the products.



Through this graph we can see that Ship F mode of shipment returns maximum rating points.



Here the relation between product importance, mode of shipment and the weights of the products that are shipped is shown.

# V. RESULTS

# VI. CONCLUSION

VII. REFERENCE

<https://www.kaggle.com/datasets/prachi13/customer-analytics>

<https://www.rpubs.com/rudiharyanto/eda-ecommerce-shipping-data>