**SUPER MARKET SALES ANALYSIS**

A picture containing text, gear, metalware

Description automatically generated

**COURSE TITLE:**

PROBABILITY AND STATISTICS (MT-0205)

**DUE DATE:**

29th APRIL, 2022

**MEMBERS:**

*ISMAIL AHMED ANSARI (20K-0228)*

*INSHA SAMNANI (20K-0247)*

*YUSRA ADAM (20K-0207)*

*ABDULLAH MEMON (20K-0283)*

**HOW DIFFERENT FACTORS CONTRIBUTES TO HIGHER SALES IN A PARTICULAR CITY**

**READING AND UPLOADING DATA:**

A screenshot of a computer

Description automatically generated

**FACTOR 1: CLIENTS**

1. VISUALIZATION OF CLIENT PER CITY: The graph shows that how much MEMBERS and NORMAL CUSTOMERS are there in a particular city out of three. Normal customers being inspired by services, referral codes, discount offers, promotes expertise’s to others that further attract new clients. On the other hand, member customers shops on daily basis, increasing the profit of a particular branch.

Chart, bar chart

Description automatically generated

1. TYPES OF CLIENTS PER CITY: It is very important to know that what is the ratio of NORMAL CUSTOMERS and MEMBERS. It must be noticed that NORMAL CUSTOMERS in each city are more as compared to MEMBERS. However, their ratio is almost equal. As when Normal customers gets introduced to a market so the services, referral codes, and discount offers introduce them as a member time by time that increases the sales of that market. This means that Yangon, in future will have more sales as compared to other two cities.

Graphical user interface, text, application, email

Description automatically generated

1. GENDER WISE CLASSIFICATION: It must be noted that MALE and FEMALE approach shopping with different motives, perspectives, rationales, and considerations. Research have found that females shops using visual imagery and interpreting context. While males use logical computations and processing facts. Females shopping motives tend to be hedonic and it is utilitarian in case of males. This means that females shops more as they buy essential things along with those things that are not necessary, but they love them. Overall, in three cities, the ratio of females and males are almost equal resulting in balanced increase of sales in all the cities.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

1. GENDER VS PAYMENT METHODS: Different payment methods increases conversions and develops market trust and loyalty. To reach the largest portion of the market possible, local payment methods that are convenient to customers must be introduced. If not provided, so here is the highest possibility that they may never return. As upper-class people usually prefer to use e-wallets and credit cards and on the other hand middle and lower-class people prefer to pay through cash. And all three payment methods are used equally by males and females, so availability of required payment methods make it easy for the clients to shop that further increases sales.

Chart, bar chart

Description automatically generated

OTHER VISUALIZATION: As described that how NORMAL CUSTOMERS and MEMBERS are important for increasing the sales in a particular market so here is the visualization of overall cities that their ratio appears almost same as both types of customers are shopping and interchanging their roles simultnaneoulsy.

Graphical user interface, text, application, email

Description automatically generated

**FACTOR 2: MODES OF PAYMENT**

1. BEST MODE OF PAYMENT PER CITY: The receiving channel allows the customers to make the payments without any difficulty and for ease and their comfort, modes of payments are introduced, it also helps supermarket owners to finance the payments. The values show the evenly equal distributed modes of payments for each branch across 3 different cities.

Table

Description automatically generated

1. VISUALIZATION OF BEST MODE OF PAYMENT PER CITY: The statistical model, also elaborated that the customers pay the amount in equal fashion of modes of payments in all the branches.

Chart, bar chart

Description automatically generated

1. PAYMENT TYPE CLASSIFICATION: This pie chart statistical model is built upon the modes of payments in the different branches. It is based on the activity and relationship the customers have with the Cash, Credit card and E-wallets.

Graphical user interface, text, application

Description automatically generated

**FACTOR 3: ENVIRONMENT OF CITIES AND BRANCHES INSIDE IT**

1. BRANCH WISE CLASSIFICATION: The below graph shows the Invoice VS Branches. It must be noted that BRANCH A have more invoice number as compared to BRANCH B and C. This means that there are a greater number of customers in branch A. The environment of branches causes great impact on lowering or enhancing the number of customers. The factors may include some SUPPLY LINES (weather, cost of fuels, etc.), ECONOMIC FACTORS (local and national economies, good and bad time, rates at which necessities are provided), LAWS AND GOVERNMENT (Increase in tax on products, restrictions on purchases, etc.), etc. Positive factors will enhance the number of customers whereas negative factors will downturn them.

Graphical user interface, application, Word

Description automatically generated

1. BRANCH WISE INCOME/REVENUE: As described regarding the number of customers, we don’t know that which customer shops for how much. Here a question arises that how the BRANCH C shown have greater income over all three branches, even if it had a smaller number of customers? It depends upon the need of customers that impacts on how many products they are buying.

Graphical user interface, application, Word

Description automatically generated

Graphical user interface, application, Word

Description automatically generated

1. PRODUCT SALES OF DIFFERENT BRANCHES: Needs of customers are the basic factor behind the sales of a market. Different customers have different needs and so desire to buy different categories of products. The well-known categories must be provided by the market to increase their sales.

BRANCH A: The 1st branch buys more home and lifestyle products. Fashion accessories are their exceptional option for them. Rest of the categories appears as their common need. The markets in this city must be well-aware regarding the needs of their customers to attract them to promote their sales.

Graphical user interface, application

Description automatically generated

BRANCH B: The 2nd branch buys more sports, travel, and fashion accessories products. Home and lifestyle products are their exceptional option for them. Rest of the categories appears as their common need. The markets in this city must be well-aware regarding the needs of their customers to attract them to promote their sales.

Graphical user interface, text, application, email

Description automatically generated

BRANCH C: The 3rd branch buys more food and beverages products. Sports, travel, home and lifestyle products are their exceptional option for them. Rest of the categories appears as their common need. The markets in this city must be well-aware regarding the needs of their customers to attract them to promote their sales.

Graphical user interface, text, application

Description automatically generated

1. TOTAL SALES BY CITIES PER DAY (only visualization because data visualization each cell represents that): In mixture of all above mentioned factors and needs affecting the customer grocery and shopping’s, below is the total sales by cities per day.

Teams

Description automatically generated

**FACTOR 4: TYPES OF PRODUCTS AND THEIR COSTS**

1. PRODUCT WISE CLASSIFICATION: Customers will not have the same attitude and behavior towards all the products category as the products differ based on degree of product involvement and the degree of differences in brand, which allows the consumers to select from a wide variety of diversified products and allowing the supermarket owners to focus on the areas to expand where customer flux is high.

Graphical user interface, text, application, email

Description automatically generated

1. AVERAGE RATING FOR EACH PRODUCT: Average rating is highly volatile early on but tends to stabilize after more reviews are collected. It basically monitors your average ratings, but also considers how many reviews the product has received, which gives a meaningful perspective on consumer sentiment. It also helps the sellers to provide a content on a winning product-even if the rating has stabilized.

Text

Description automatically generated

Chart, bar chart

Description automatically generated

1. UNIT PRICE OF DIFFERENT KINDS OF PRODUCTS (unit prices fluctuation as well as ranges): It allows you to compare the goods. Customers and owners both can easily compare which product is more convenient to them, allowing them to explore the same products in different packages, sizes, forms, or configurations.

Chart, histogram

Description automatically generated

1. COST OF GOODS SOLD: Cost of goods sold is the carrying value of goods sold during a particular goods during the period, helping the supermarket for restock and to estimate the expiration period of the items stocked.

Chart, histogram

Description automatically generated

1. RATINGS VS UNIT PRICE: Determining the price quoted in terms so much per agrees or estimating the standard product service fee ranging in the customer’s view. Basically, it is the customer’s views on pricing of the products which help the supermarket to raise the products of that pricing for better valuations and sales.

Chart, bar chart

Description automatically generated

**REGRESSION AND CORRELATION:**

**CONVERTING AND TRAINING DATA:**

To train our model for regression, we convert the nominal data into categorical data for order or sequence. We create a new variable, against each value either in binary, dummy for different disciplines that attempts to determine the strength and character of the relationship. After building this, we just printed the new converted data.

Text

Description automatically generated

Graphical user interface, table, Excel

Description automatically generated

**KNEIGHBORS CLASSIFIER - REGRESSION:**

The K-Nearest Neighbors Algorithm is a supervised learning algorithm that can be used to solve both the classification and regression problems. It assumes that similar things exist in proximity or maybe like similar things are near to each other. Here for instance the algorithm is working on the Genders coming to the supermarket to hinges on the assumption being true enough to satisfy all the correct outcomes like the rating given by some specific gender will be this, these gender will might possibly behave in this fashion on a random based variable given. The idea basically is to capture the concept of similarity with some distancing relationships and Euclidean formulae. Precision to the estimation is recorded in general.

**Text

Description automatically generated**

**PREDICTING GROSS INCOME VS RATING (Linear Regression Model):**

When rating is plotted against the gross income to know the linear regression model i.e., knowing the prediction of the gross income coming from different products according to their ratings, where the linear line is representing the model fit.

**Chart, scatter chart

Description automatically generated**

**PREDICTING COGS VS RATING (Linear Regression Model):**

The COGS vs rating predicted ratio in lmplot showcases the linear regression model of best rated products revenue estimation (up to the precision according to the values) that is used for expenditures, gross income and sales unit balancing that vary directly with the sales of your business.

Chart, scatter chart

Description automatically generated