

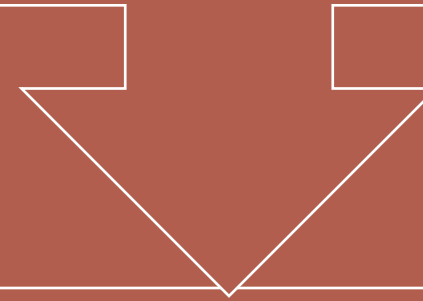
Project:- customer retention

Presenting a report on the project containing all important points to be briefly described in such a manner for the purpose of analyzing the project is done at the best



Summary on the problem statement and my understanding on it

Customer satisfaction is the important factor for the survival of ecommerce stores. customer satisfaction means the every single money used to purchase product should give the customer ,equal value compared to money spent.in other words that the customer needs and desires should be fully statisfied by spending the money on the product. customer satisfaction depicts that customer is willing to buy more product from the company ,this results repeat repurchase intention of the product. customer satisfaction contains major factors inside it which influence the customer satisfaction and customer intention to repurchase such as trust ,net benefit ,system quality ,information quality ,service quality , etc. these factor includes the information provided by the customers as the feedback on the product and the ecommerce store.



According to me ,the problem statement is the customer satisfaction on the product which is getting value for money spent.

Describing the Steps performed and assumptions used for EDA and to complete the project in dataset and thus visualizing of the EDA steps for better understanding. Page 1

- Importing the libraries required for the process of EDA.
- With the help of pandas library ,reading the excel (xlsx file) file in python.
- Knowing about the first five rows and last five rows in the dataset.
- Knowing about the number of rows and columns present which is 269 and 71
- Knowing the data types of the columns present which is object and int 64
- Knowing the every single information about the columns such as dtypes, non null, etc.
- Finding out the null values present in the dataset which is represented as False
- Getting the sum of the null values present if any which is zero

Page 2

- Visualizing the null values present if any by heatmap



Page 3

- Getting the columns name.
- Using the label encoder technique to convert the object datatype columns present into the numerical datatype columns so that the accuracy can be maintained and making the machine learn the dataset.
- Seeing the statistical view of the dataset such as mean ,median, minimum ,maximum, etc.
- Checking the correlation between the columns with each other and with target columns and visualizing it with the help of heatmap ,see in the next slide



Page 5

- Doing the visualization of the dataset by plotting bar graph, histogram, kdeplot, boxplot. As the graphs are many so I am not able to show .
- Finding out the outliers present in the dataset with the help of boxplot. there are outliers present in the columns so after removing the outliers from the dataset with the help of zscore ,I came to know that the data loss percentage is near to 7% so therefore I didn't remove outliers as if I remove outliers then the data is becoming biased as there can be possibilities of losing important information and also the data is categorical in nature.
- As the dataset is categorical so I am not finding out the skewness present and also not removing the skewness because the skewness operation is not done on the categorical datatype.
- Now I dividing the dataset in the feature columns as x and target column as y variable. this is done for the purpose making the data balanced ,scaling the data and sending the data for its futhur processes to be done.
- The target column has different classes present so therefore we are checking the value counts of it , I came to know that the classes present in the target column is imbalanced so we need to balance the dataset.to balance the dataset I am using the smote technique from imbalanced learn.
- Now the dataset should be scaled before sending it to the splitting and for machine learning so to scale the data , I am using the standard scaler.

Analysis and conclusion of the project

- According to me ,after studying the data ,I came to know that there are so many factors that influence the customer behaviour regarding the purchase of the product after buying it once .there are factors which influence the success of ecommerce stores too such as their services ,accurate information provided ,website used to sale ,price offered for the products ,discount given on the products, etc. The factors of success of ecommerce stores depends on the customer response. Customer response can be in positive or negative .the positive or negative response is mostly depends on the customer satisfaction after the purchase of the product .customer satisfaction not only depends on the amount spend for getting the desires and wants to be satisfied but also on the after sales services too. At last the conclusion is that the customer needs and desires to be fulfilled or satisfied is the important factor in the success of ecommerce stores. customer intention to repurchase the product is the result of customer satisfaction for the product purchased. Without buyers , the ecommerce retailers can't sell their products and the buyers can be attracted only by providing the good quality products in cheap prices and also providing after sale services.