

### **E-retail factors for customer activation and retention: A case study from Indian e-commerce customers.**

We need to understand and analyze five qualities (service quality, system quality, information quality, trust and net benefit) that creates customer utility and desire to buy products from a specific online retailer. To analyze these qualities first we need to calculate utilitarian value and hedonistic values which can be used to calculate customer retention points. This analysis can bring several aspects on customer satisfaction and new customer integration.

### **Data Cleaning and Data Manipulation**

1. Identified data type of the columns whether they are Int or Object.
2. Identified whether the column contains continuous or categorical data.
3. Every categorical data had Object data type so changed the Object data type to Int for analysis purpose.  
Below are the data points which are converted to Int based on their points.  
Agree (4) = 4  
Strongly agree (5) = 5  
indifferent (3) = 3  
Dis-agree (2) = 2  
Strongly disagree (1) = 1  
Indifferent (3) = 3  
Disagree (2) = 2  
indifferent (3) = 3
4. Created hedonic\_value column by adding gratification, certain\_roles, discounts, social\_status and sense\_adventure column values to analyse hedonic value the customer is getting from mentioned columns.
5. Created utilitarian\_value column by adding product\_variety, convenience, complete\_product\_information and monetary\_savings column values to analyse utilitarian value the customer is getting from mentioned columns.
6. Created customer\_retention\_points column by adding hedonic\_value and utilitarian\_value column values to analyse overall customer satisfaction and retention of a customer on the basis of hedonic and utilitarian values.
7. Created separate table for those columns which has URLs and analysed those differently.
8. Created separate table for those questions which are related to customer experience and analysed those separately.
9. Created separate table for those columns which have Int values and analysed those separately. Also picking those columns for model building.
10. Plotting scatter plot to analyse the correlation of selected features with Label i.e Customer Retention Points.
11. Plotting distribution plot to analyse the distribution of each feature.
12. Used corrwith technique to analyse the correlation of label with features.
13. Have used the heatmap to show the Multicollinearity between features.
14. Have used boxplot to show the outliers.