

Assessment – II

1. Load the data from the following link:
<https://archive.ics.uci.edu/ml/datasets/Facebook+Live+Sellers+in+Thailand>
2. Display the number of rows and columns of loaded data set.
3. Display summary of the loaded data set.
4. Count the number of missing values are there for each column and drop the columns which have more number of null values approximately equal to the number of rows of data.
5. List out number of categorical variables and numerical variables are available in the data set.
6. Display statistical summary of numerical attributes.
7. Count the number of unique values are existed in categorical attributes and drop the all columns which have number of unique values approximately equal to number of rows of data.
8. Convert the remaining categorical variable into equivalent numerical form and perform Min-Max Scaling for all attributes.
9. Develop a python program to perform K-Means Clustering with number of clusters as $K = 2, 3, 4, 5, 6$ and display inter-cluster distances and sum of all inter-cluster distances. Identify at what K value the inter-cluster distance is minimal.