

- For Clustering I have merge all the three table
- Then I have make 3 new data to check transaction count , total spending , avg spending.
- Then I have selected relevant features from the data.
- Then encoded categorical data and featured it.
- Then I have normalize the data then used Elbow method to find optimal number of clusters.
- Then applied K means clustering and checked DB index and silhouette score which are 1.4111 and 0.2094
- Although my DB index is good enough but i cant get good silhouette score
- But I have clustered my customer on the basic of there purchasing power. There are 5 cluster means customer segmentations which I have applied in python.