

Task-3

The script her is for clustering customer into groups. Since there is no information what are the features the customers would be grouped on, clustering on the basis of RMF(Recency, Monetary, Frequency) features was done.

I used Within-Cluster Sum of Squares(WCSS) to determine the optimal number of clusters from this data, which I found to be 3 but I increased the number of clusters by 1 more for the ease of understanding w.r.t business prospects

So the clusters I created are -

Gr 1 - Customers that spend a lot and much frequently

Gr 2 - Customers that spend less and shop rarely

Gr 3 - Customers that spend a lot but shop rarely

Gr 4 - Customers that shop frequently but do not spend much

The Davies-Bouldin Index(DBI) for this is 0.874 which means that clusters are distinctively seperated

The Silhoutte score is 0.371 which means the clusters are spread apart a little and not compact with the centroid