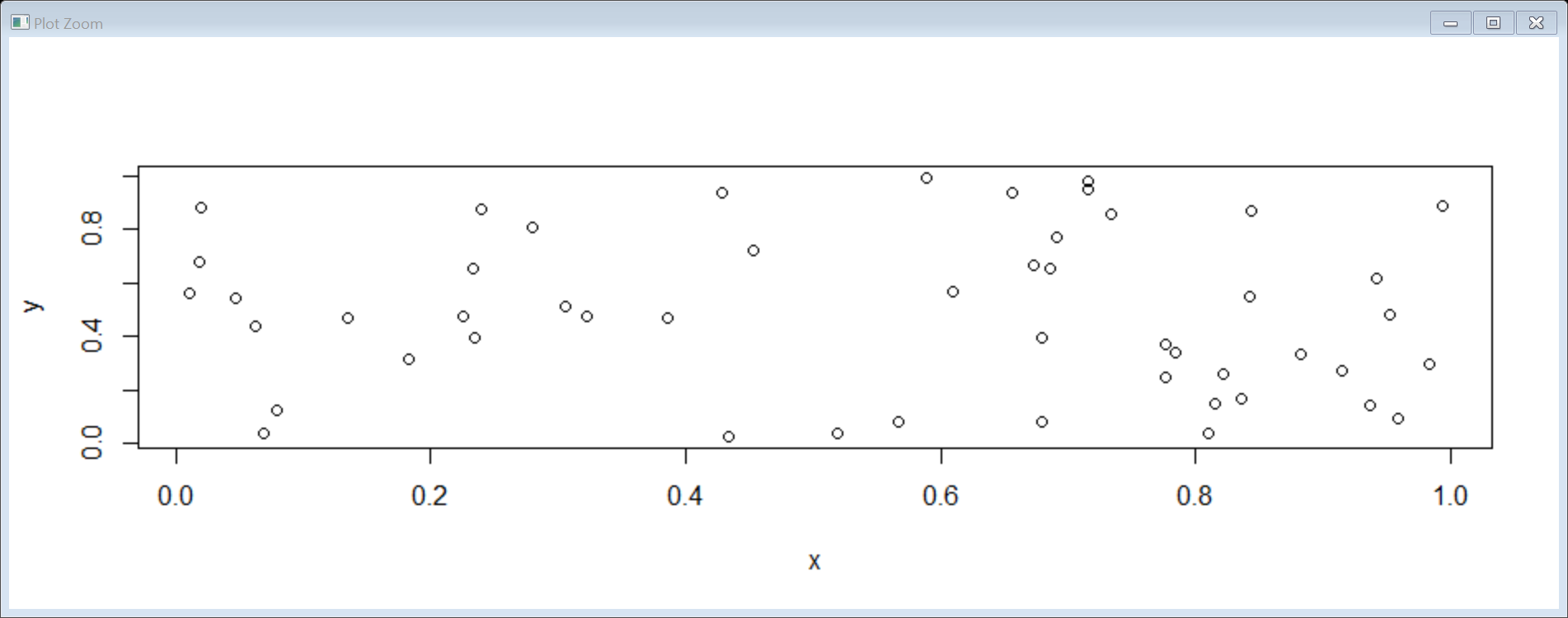
# K Means Clustering

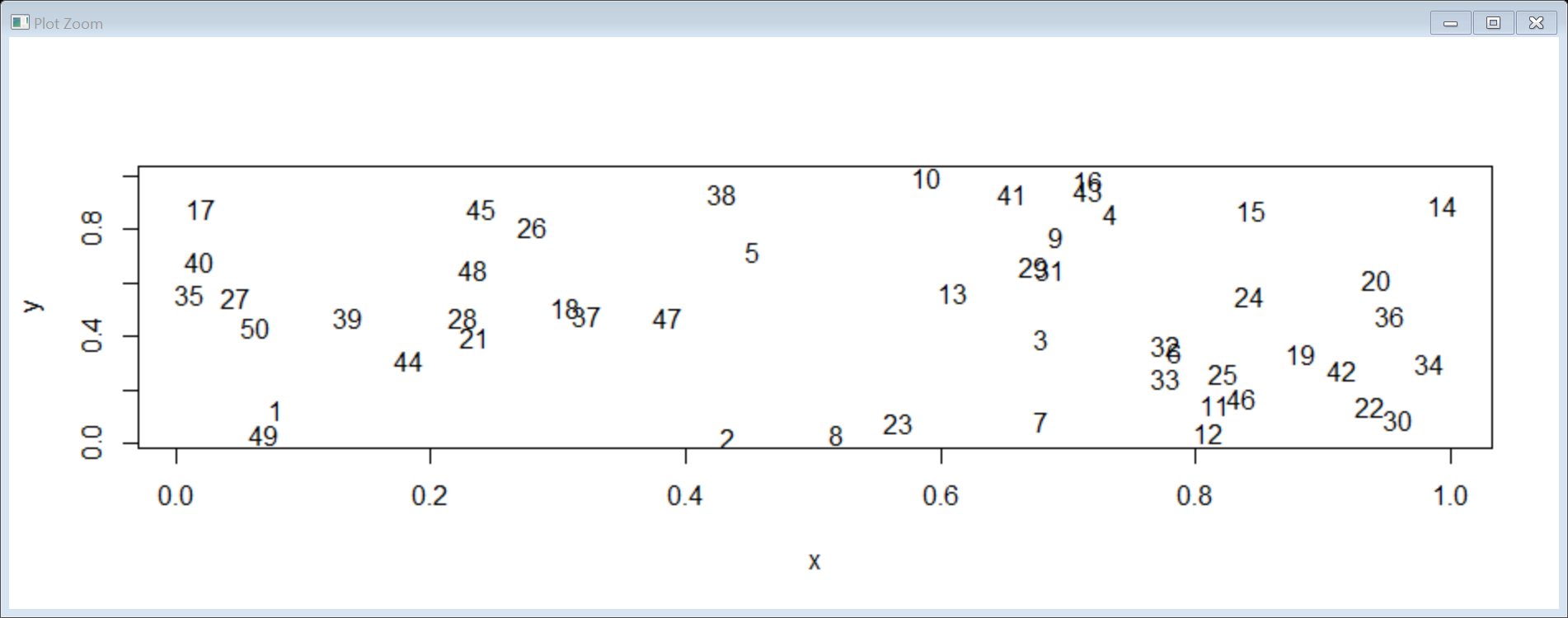
**Business Objective: Clustering for the different types of mileage offers**

Step1: generating 50 random numbers of x and y and combining them

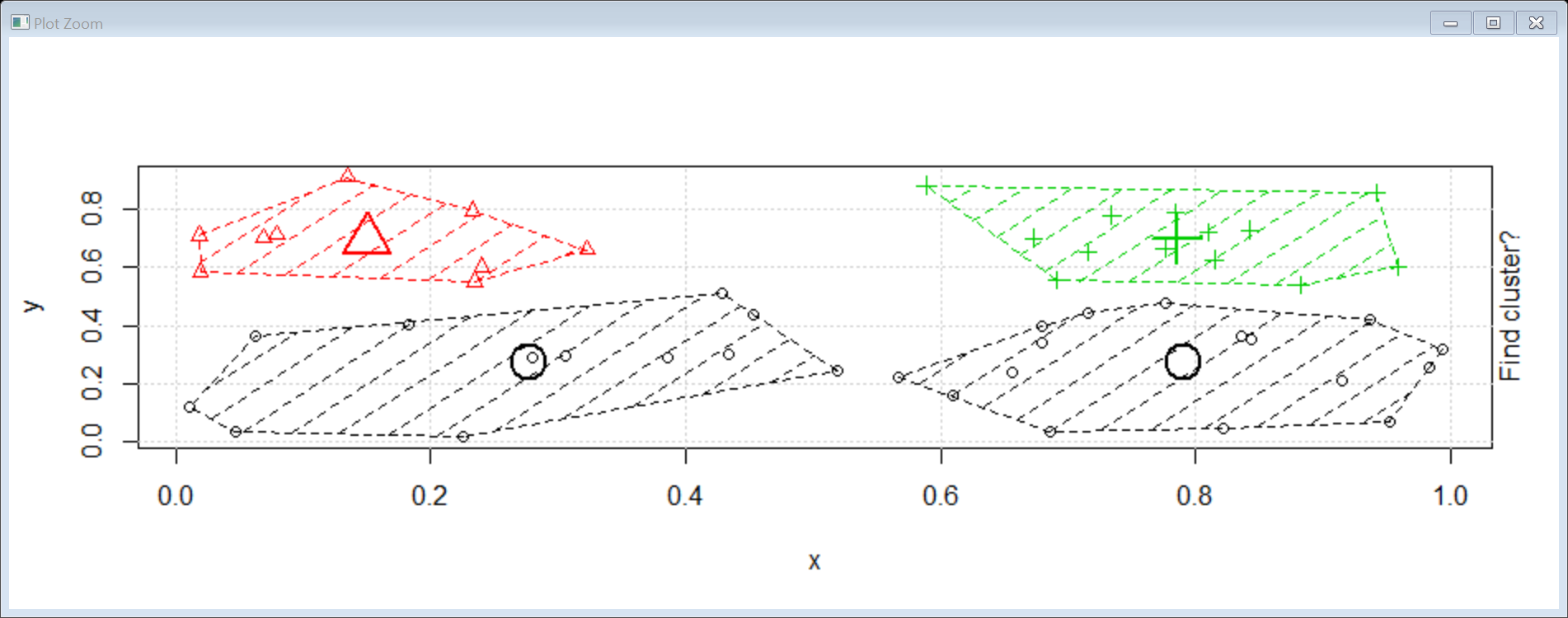
Step 2: Scatter plot.

****

Step 3: Naming the data point with numbers.

****

Step 4: Animated representation of the data which forms the centroid and checks each data point with that centroid and keeps calculating until the most appropriate clusters are formed since 4 clusters are mentioned so it calculated with 4 centroids accordingly.



Step 5 : Normalize the data and calculate the K means with 4 clusters

Step 6 : Aggregate value for each cluster is calculated.

Group.1 Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles

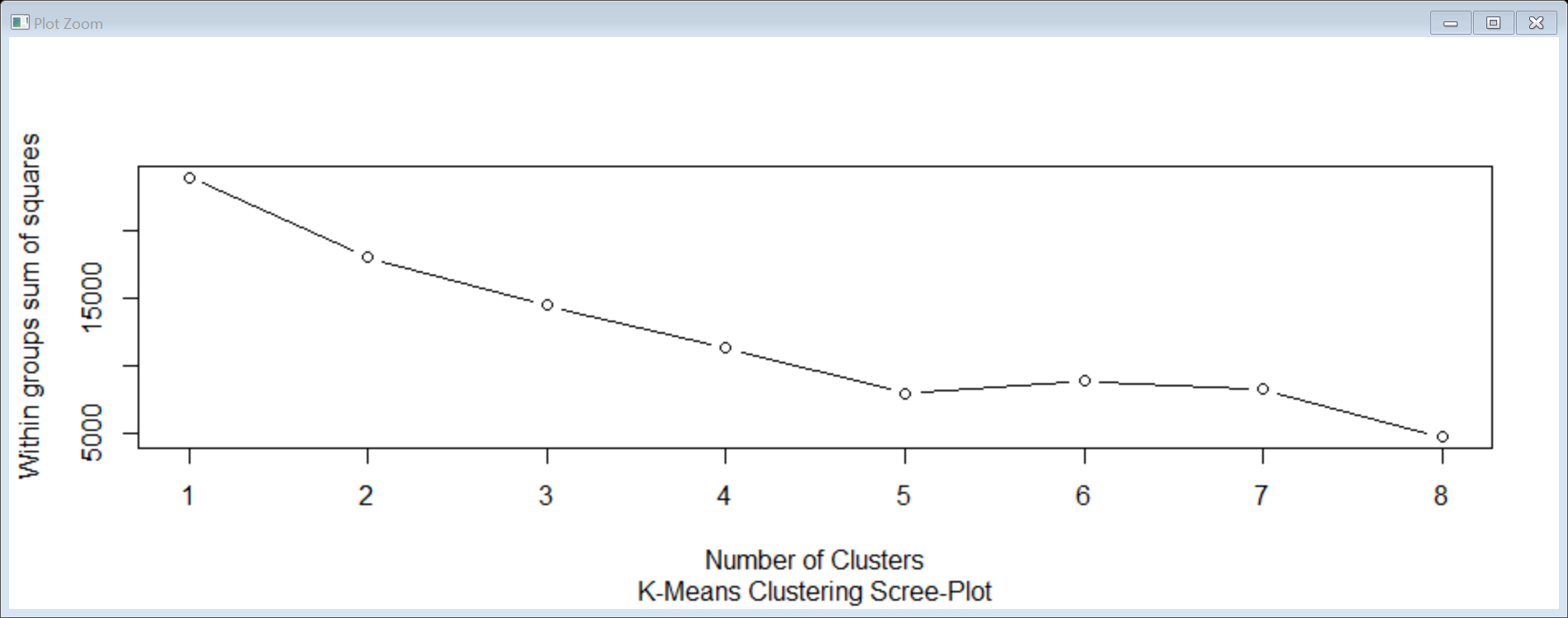
1 1 145394.4 78.4609 4.084463 1.003128 1.002086 48802.725

2 2 128537.1 4333.1042 1.864583 1.000000 1.000000 16510.625

3 3 138061.4 78.8000 3.466667 1.000000 4.066667 93927.867

4 4 47964.5 28.6480 1.395698 1.018778 1.000341 6407.132

Step 7: Determine number of clusters by scree-plot and determining the elbow curve.



Conclusion: 4 clusters are formed in the airlines data

Cluster 3 - Highest milege offers,

Cluster 1 – Higer Milege offers

Cluster 2 - High Milege offers

Cluster 4 - Medium Milege offers.