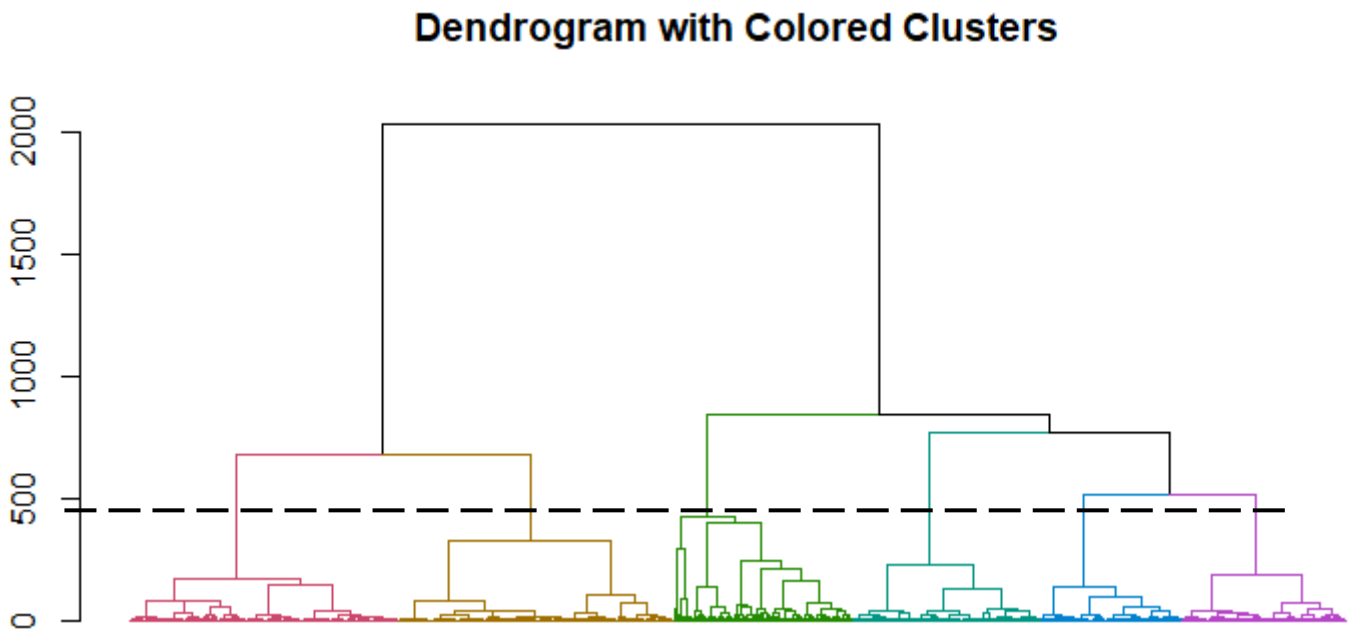
Salwa Niaz Preet
Sxp240086

Business Analytics with R
Assignment 4

a) At varying distances, we can find different number of clusters. Close to a distance of 400, 6 major clusters can be seen.



b) Raw distance measures are highly influenced by scale of measurements. Some variables have a high range and can go upto large numeric values, like incomes and revenues. But there might be equally important variables that only take binary values or a smaller range, for instance ranking of products from 1 to 5. Using distance formula on data that is not normalized means that the distances will also vary the same way (will be skewed), variable with wider range will have larger distances. The clustering may be biased towards the higher range variables. Additionally, the clusters in the dendogram will become difficult to interpret.

c) The clusters based on centroids are as follows:

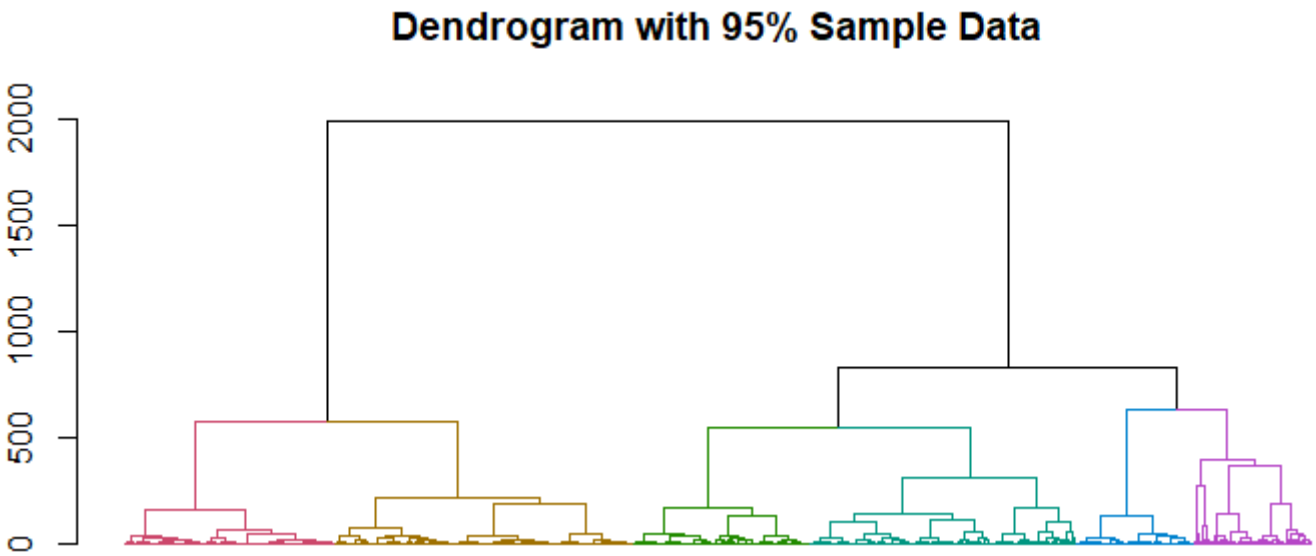
Cluster	Balance	Qual miles	cc1 miles	cc2 miles	cc3 miles	Bonus miles	Bonus trans	Flight miles 12mo	Flight trans12	Days since enroll	Award.
1	46,828	5.7	1.03	1.00	1.00	2,480	5.38	145.5	0.46	5,245	-
2	161,110	934.3	2.54	1.10	1.08	34,913	20.60	2,279.1	6.61	4,462	0.678
3	108,831	6.5	3.39	1.00	1.00	28,680	15.80	94.6	0.28	4,939	-
4	50,784	9.1	1.59	1.00	1.00	9,547	9.68	255.1	0.83	3,956	1.000
5	82,774	6.8	3.98	1.00	1.00	43,390	18.02	168.8	0.56	5,342	1.000
6	33,108	11.6	1.30	1.00	1.00	4,842	7.38	95.2	0.31	1,802	-

Cluster 1: Low Balance, Very Low Miles, Low Activity

Significant features that can differentiate this cluster from others are that it has low balance, very low miles earned from various transactions, no reward flight earned and very low flight miles and flight transactions in the past 12 months.

Cluster 2: High Balance, Very High Miles, High Activity
This cluster represents very high balance, high accrued and earned miles (high qual miles, bonus miles and cc1 miles), high flight transactions and miles flown in past 12 months and also that 67.8% won a reward flight. This cluster shows highly active and frequent flyers.
Cluster 3: High Balance, Low Activity, Bonus Focused
This cluster represents high balance, moderate miles earned through various transactions, high bonus miles and low flight transactions and miles flown in past 12. This shows that these customers have low activity and likely participate in bonus programs to earn miles.
Cluster 4: Moderate Balance, Moderate Activity
This cluster has low to moderate balance, moderate qual miles, low bonus miles, low qual miles, moderate flying activity and low transactions. Nearly all members receive reward flights.
Cluster 5: High Balance, High Miles, Moderate Activity
These customers have high balances, high cc1 and bonus miles but moderate recent activity. Almost all seem to have received a reward flight.
Cluster 6: Low Balance, Low Miles, Inactive
This cluster represents customers with very low balances, low accrued miles very low activity, and no awards. This cluster also shows flyers that are relatively new in the flyer program (1802 days since enrolled). They are inactive.

d)

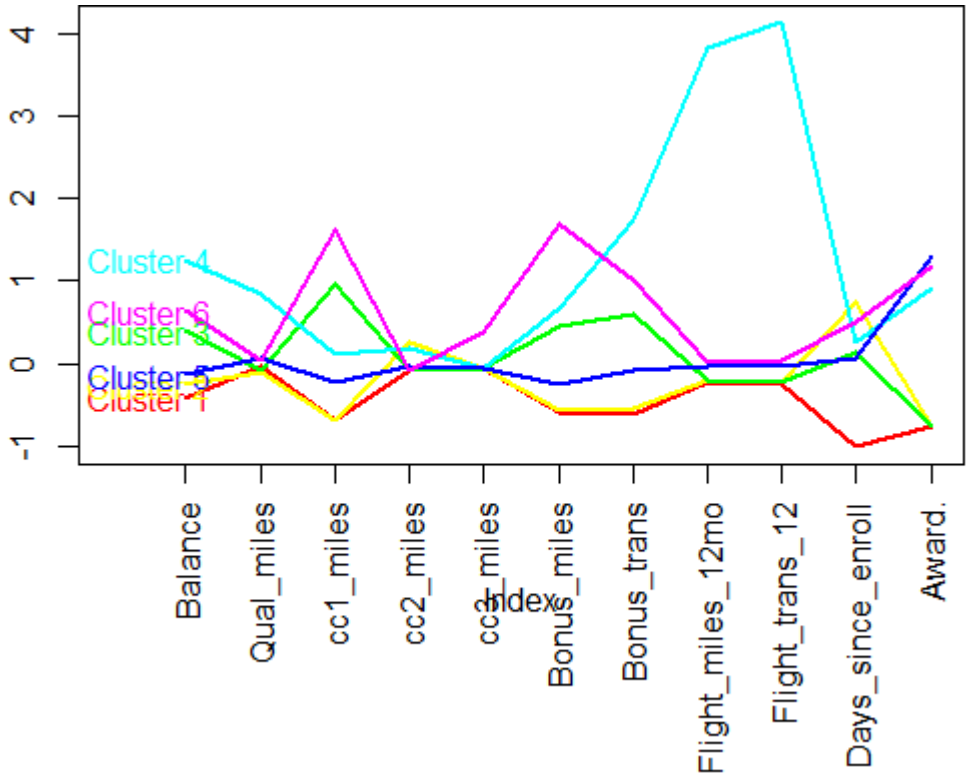


Comparing this dendogram with the previous one we can say that the general clustering structure in both is the same. In the dendrogram, created with a 95% sample of the data, the overall grouping pattern is visually consistent with the original dendrogram. There are some minor differences in cluster boundaries and the teal-colored branch but that is probably due to the random 5% data exclusion. The main clusters and their remaining branching pattern and distances remain similar, suggesting that the clustering is stable. This is a positive indication that this clustering not overly sensitive to small changes in the data, affirming the validity of the clusters identified in the analysis.

e)

Salwa Niaz Preet
Sxp240086

Cluster	Balance	Qual miles	cc1 miles	cc2 miles	cc3 miles	Bonus miles	Bonus trans	Flight miles 12mo	Flight trans 12	Days since enroll	Award.
1	-0.419	-0.042	-0.700	-0.092	-0.063	-0.588	-0.626	-0.223	-0.247	-1.012	-0.767
2	-0.255	-0.108	-0.695	0.245	-0.056	-0.571	-0.548	-0.196	-0.220	0.735	-0.754
3	0.391	-0.091	0.947	-0.087	-0.055	0.444	0.599	-0.196	-0.219	0.145	-0.767
4	1.256	0.842	0.107	0.192	-0.063	0.678	1.735	3.810	4.137	0.257	0.919
5	-0.144	0.061	-0.226	-0.035	-0.063	-0.247	-0.082	-0.036	-0.035	0.062	1.304
6	0.636	0.030	1.617	-0.086	0.382	1.681	1.008	0.014	0.037	0.508	1.181



Cluster 1: Low Balance, Inactive/ Low activity (table and red line for cluster 1 shows low values for each variable)
Cluster 2: Moderate Balance, Low Overall Miles, Low Activity (moderate cc2 miles, overall low miles, positive enrollment period)
Cluster 3: Moderate Balance, High CC1 Miles, High bonus miles, Low Activity (bonus programs participation)
Cluster 4: High Balance, Very High Activity (significantly above average balance, bonus miles, transactions, and flight miles – cyan line shows high values for variables)
Cluster 5: Moderate Balance, Moderate Miles, Moderate Activity (slightly below average in most features)
Cluster 6: High Miles and Bonus Miles, Moderate Activity (very high cc1 and bonus miles)

Both clustering methods give similar groups but with different emphases on balance, miles, and activity levels. Hierarchical clustering is more absolute, while k-means gives a clearer picture of deviations from the average.

f) Each cluster can be approached in a different way. All customers are important and in order to prevent churn, increase loyalty and sales, each cluster needs to be targeted with a tailored strategy:

High Balance, High Activity (Hierarchical Cluster 2 / k-Means Cluster 4): Loyalty rewards, premium benefits and upgrades can be given like VIP lounge access, faster check-ins and boarding on priority. This exclusive treatment will increase loyalty, spending and decrease chance of customer churn as well.

Moderate Balance, Moderate Activity (Hierarchical Cluster 4 / k-Means Cluster 5 and 3): Seasonal promotions on specific occasions can be given, for instance extra bonus miles on holidays, referring to friend incentive, some discount on hotels.

Low Balance, Low Activity, Inactive (Hierarchical Cluster 1 / k-Means Cluster 1): To attract these customers return, reactivation promotions can be given. For instance, any small non-flight transaction can help earn miles, a purchase in a specific range can activate flight and hotel discounts.

Recently Enrolled: 6-month and 1-year membership anniversary bonus – maybe extra miles, refer to a friend incentive, extra luggage on free or minimal cost