

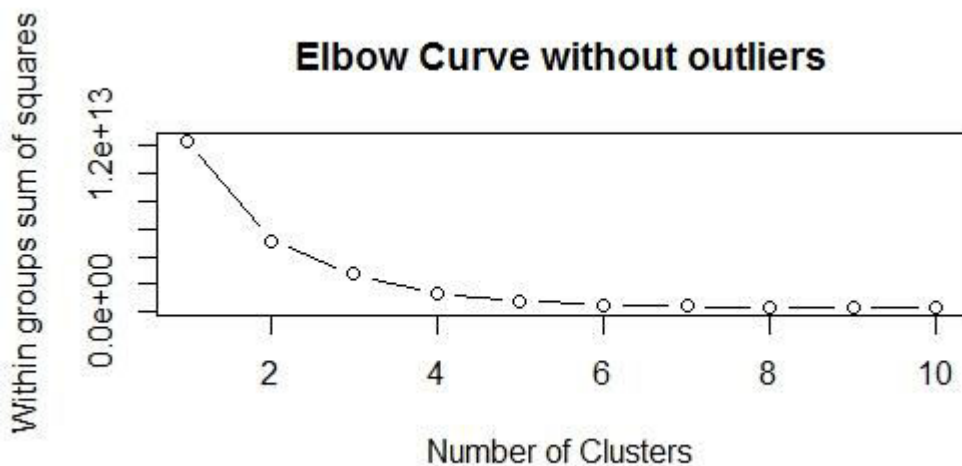
Loyalty clusters results

To find loyalty clusters we gave Score to each customer based on 3 factors.

1. Recency – how recently the customer visited.
2. Frequency – How frequently customer visited.
3. Monetary – What is the average spend of customer.

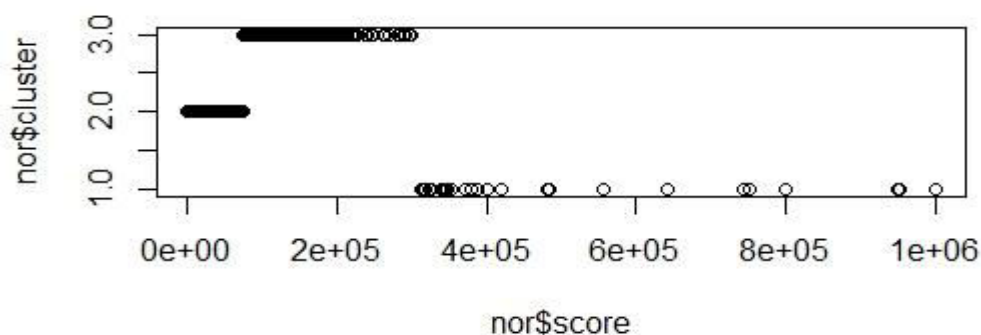
All these factors will be important to determine the loyalty of customers. For ex- A customer visiting every week spending \$5 won't be as good as a person visiting once a month and spending \$100. Similarly a person who hasn't recently visited in the last 6 months would be less likely to visit again.

Based on the RFM score and other attributes I decided to create clusters. From the elbow curve plotted I decided to select 3 clusters for clustering. That is least loyal, somewhat loyal and most loyal customers.



Inference

`plot(nor$score, nor$cluster)`



Based on recency, Frequency and Monetary scores customers were divided into three cluster. Above graph illustrates the same.

Average Score in each cluster

aggregate(score ~ cluster, data = nor, mean)

```
cluster  score
1      1 471442.72
2      2  13543.89
3      3 136996.83
```

Average M value in each cluster

aggregate(M ~ cluster, data = nor, mean)

```
cluster  M
1      1 48.294444
2      2  9.641662
3      3 28.185255
```

Average Spend in each cluster

aggregate(SPEND ~ cluster, data = nor, mean)

```
cluster  SPEND
1      1 2.357134
2      2 1.782737
3      3 2.121758
```

Average R value in each cluster

aggregate(R ~ cluster, data = nor, mean)

```
cluster  R
1      1 95.31250
2      2 45.49943
3      3 83.56275
```

Average F value in each cluster

aggregate(F ~ cluster, data = nor, mean)

```
cluster  F
1      1 69.37500
2      2 17.62392
3      3 45.41264
```

It can be inferred that cluster 1 is most loyal followed by cluster 3 and cluster 2 is least loyal.

cluster	R	F	M	Score
1 (32 Entries)	1st	1st	1st	1st
2 (2182 Entries)	3 rd	3 rd	3 rd	3 rd
3 (247 Entries)	2nd	2nd	2nd	2nd

The aim of using Kmeans clustering instead of directly separating observations in three segments is to find out similar characteristic of each cluster.

We are treating Outliers of spend as separate cluster as their spend is way more than rest of data.

Since they are spending a lot more than average spend of customers, they will also be considered loyal irrespective of their score.