MSCI 433 Assignment 4

Q1:

Q1.a.i ) The reason it will take a large amount of time to build a hierarchal clustering model is due to the fact that

Q1.a.ii) I would recommend a cluster size of 6 because you would like to not overwhelm the reader too many category recommendations.

Chart

Description automatically generated

Q1.a.iii)

Each cluster has the following number of observations

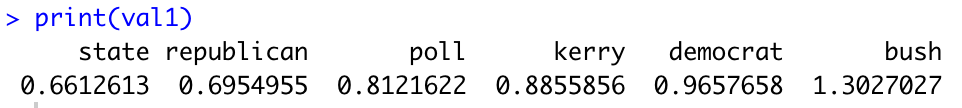
|  |  |
| --- | --- |
| Cluster | # of observations |
| 1 | 2220 |
| 2 | 436 |
| 3 | 324 |
| 4 | 166 |
| 5 | 228 |
| 6 | 56 |

Table

Description automatically generated

Q1.a.iv) Description of each cluster

Cluster 1:



Cluster 2:

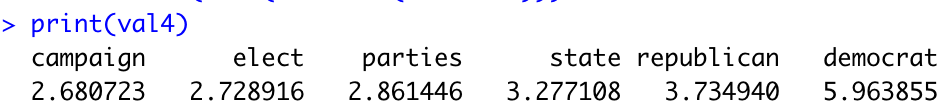
Text, letter

Description automatically generated

Cluster 3:



Cluster 4:

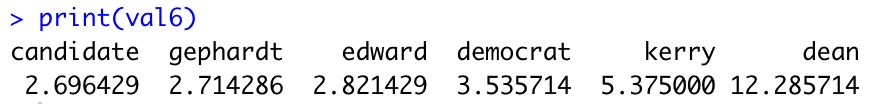


Cluster 5:

Text

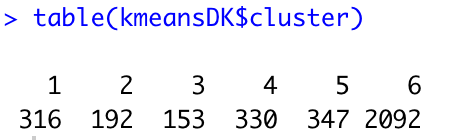
Description automatically generated

Cluster6:



Q1.b)

Q1.b.i) The answer is not exactly the same as in hierarchal clustering.



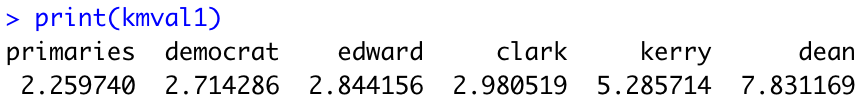
Q1.b.ii)

The clusters that are the same across the 2 models are:

Kmcluster 1:

Text, letter

Description automatically generated

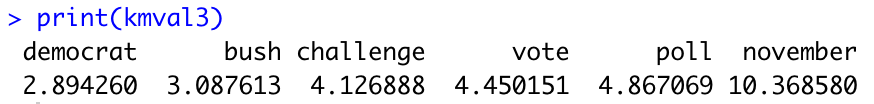


Kmcluster2:

Text

Description automatically generated

Kmcluster3:

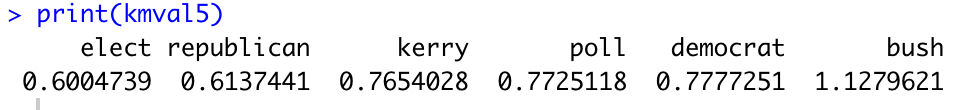


Kmcluster4:

Text

Description automatically generated

Kmcluster5:



Kmcluster6:

