ThesisSurveyCluster

07 March, 2021

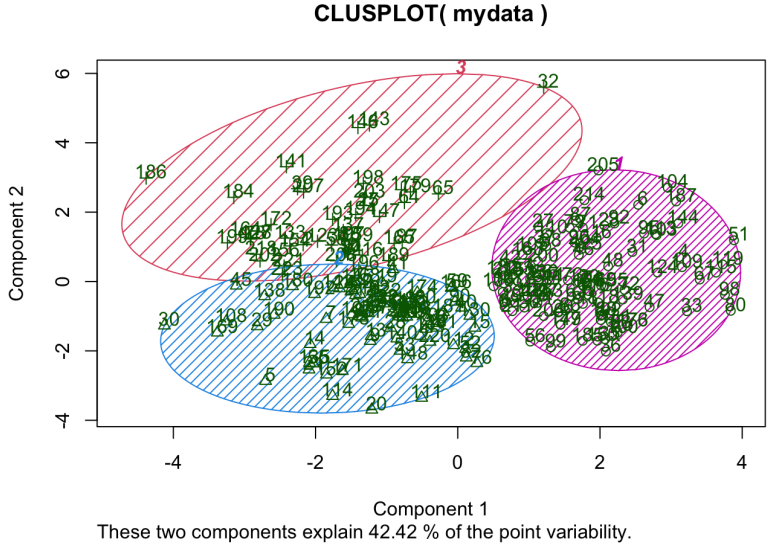
https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a

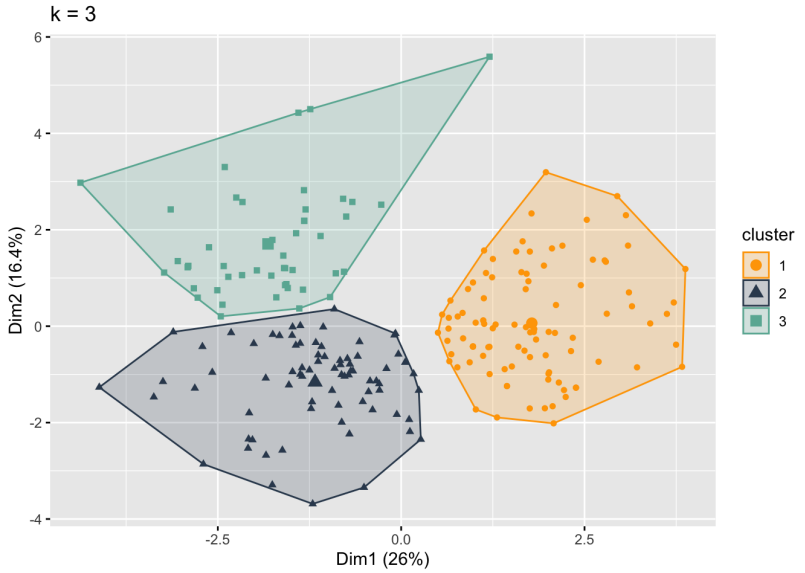
One hot encoding

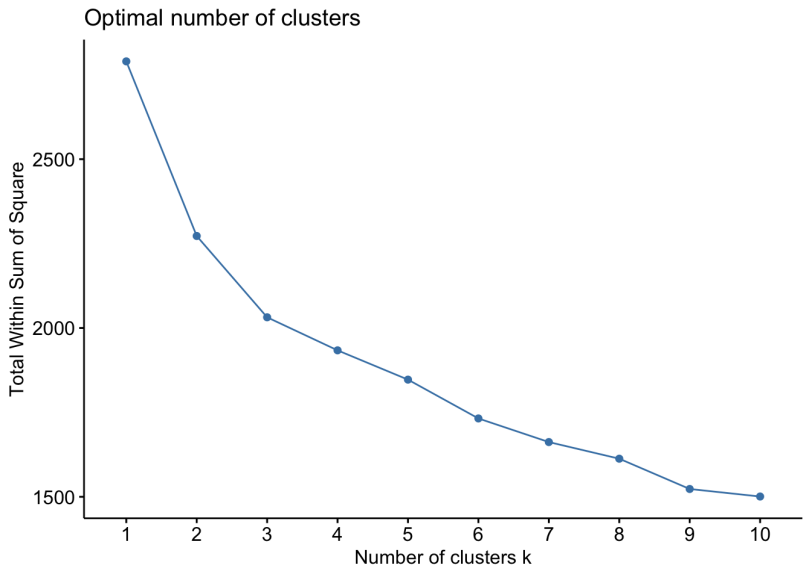
Caret package

Clustering Numerical

```
  Group.1 AmountWeek AmountOutMonth MoneyCoffee MoneyGroceries KnowledgeCoffee
1       1  0.2443646     -0.1856383  0.01523532     -0.1030381      -0.1460830
2       2 -0.5176410     -0.2381982 -0.41641927     -0.1409261      -0.2033228
3       3  0.3418313      0.7788835  0.64901217      0.4356492       0.6721079
  Purchase_Price Purchase_Sustainability Purchase_Certificate
1     -0.3232527              -0.7029086           -0.5833119
2      0.4808903               0.6990571            0.6656516
3     -0.0924339               0.3098094            0.1414696
  Purchase_Fairtrade Purchase_Packaging Subscription_Likely  App_Likely
1         -0.7068063         -0.5578558         -0.47089152 -0.42106889
2          0.6617415          0.6916814         -0.05904157 -0.02451035
3          0.3777212          0.0119494          1.08633570  0.92940313
```




k = 3
Dim2 (16.4%)
Dim1 (26%)
cluster
1
2
3


Optimal number of clusters
Total Within Sum of Square
Number of clusters k