Assignment-5

rama krishna

11/29/2021

```
setwd("C:/Users/krish/OneDrive/Desktop/R_MLCODES/rmullapu_64060")
library(readr)
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse
1.3.1 --

## v ggplot2 3.3.5     v dplyr   1.0.7
## v tibble  3.1.4     v stringr 1.4.0
## v tidyr   1.1.3     v forcats 0.5.1
## v purrr   0.3.4

## -- Conflicts ------------------------------------------
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(cluster)
library(dplyr)
library(knitr)
library(caret)

## Loading required package: lattice

##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
##     lift

library(factoextra)

## Welcome! Want to learn more? See two factoextra-related books at
https://goo.gl/ve3WBa

library(dendextend)

##
## ---------------------
## Welcome to dendextend version 1.15.2
## Type citation('dendextend') for how to cite the package.
##
## Type browseVignettes(package = 'dendextend') for the package vignette.
```

```
## The github page is: https://github.com/talgalili/dendextend/
##
## Suggestions and bug-reports can be submitted at:
https://github.com/talgalili/dendextend/issues
## You may ask questions at stackoverflow, use the r and dendextend tags:
##   https://stackoverflow.com/questions/tagged/dendextend
##
##  To suppress this message use:
suppressPackageStartupMessages(library(dendextend))
## ----------------------

##
## Attaching package: 'dendextend'

## The following object is masked from 'package:stats':
##
##     cutree

cereals<- read.csv("Cereals.csv")
cereals

##                                     name mfr type calories protein fat
sodium
## 1                              100%_Bran   N    C       70       4   1
130
## 2                      100%_Natural_Bran   Q    C      120       3   5
15
## 3                               All-Bran   K    C       70       4   1
260
## 4               All-Bran_with_Extra_Fiber   K    C       50       4   0
140
## 5                          Almond_Delight   R    C      110       2   2
200
## 6                 Apple_Cinnamon_Cheerios   G    C      110       2   2
180
## 7                              Apple_Jacks   K    C      110       2   0
125
## 8                                  Basic_4   G    C      130       3   2
210
## 9                                Bran_Chex   R    C       90       2   1
200
## 10                             Bran_Flakes   P    C       90       3   0
210
## 11                           Cap'n'Crunch   Q    C      120       1   2
220
## 12                                Cheerios   G    C      110       6   2
290
## 13                  Cinnamon_Toast_Crunch   G    C      120       1   3
210
## 14                               Clusters   G    C      110       3   2
140
```

```
## 15                            Cocoa_Puffs   G    C      110        1   1
180
## 16                              Corn_Chex   R    C      110        2   0
280
## 17                            Corn_Flakes   K    C      100        2   0
290
## 18                              Corn_Pops   K    C      110        1   0
90
## 19                           Count_Chocula   G    C      110        1   1
180
## 20                       Cracklin'_Oat_Bran   K    C      110        3   3
140
## 21                  Cream_of_Wheat_(Quick)   N    H      100        3   0
80
## 22                                Crispix   K    C      110        2   0
220
## 23                  Crispy_Wheat_&_Raisins   G    C      100        2   1
140
## 24                            Double_Chex   R    C      100        2   0
190
## 25                            Froot_Loops   K    C      110        2   1
125
## 26                          Frosted_Flakes   K    C      110        1   0
200
## 27                     Frosted_Mini-Wheats   K    C      100        3   0
0
## 28 Fruit_&_Fibre_Dates,_Walnuts,_and_Oats   P    C      120        3   2
160
## 29                             Fruitful_Bran   K    C      120        3   0
240
## 30                          Fruity_Pebbles   P    C      110        1   1
135
## 31                            Golden_Crisp   P    C      100        2   0
45
## 32                          Golden_Grahams   G    C      110        1   1
280
## 33                       Grape_Nuts_Flakes   P    C      100        3   1
140
## 34                              Grape-Nuts   P    C      110        3   0
170
## 35                       Great_Grains_Pecan   P    C      120        3   3
75
## 36                         Honey_Graham_Ohs   Q    C      120        1   2
220
## 37                       Honey_Nut_Cheerios   G    C      110        3   1
250
## 38                              Honey-comb   P    C      110        1   0
180
## 39              Just_Right_Crunchy__Nuggets   K    C      110        2   1
170
```

```
## 40                 Just_Right_Fruit_&_Nut   K    C      140        3   1
170
## 41                                    Kix   G    C      110        2   1
260
## 42                                   Life   Q    C      100        4   2
150
## 43                           Lucky_Charms   G    C      110        2   1
180
## 44                                  Maypo   A    H      100        4   1
0
## 45       Muesli_Raisins,_Dates,_&_Almonds   R    C      150        4   3
95
## 46      Muesli_Raisins,_Peaches,_&_Pecans   R    C      150        4   3
150
## 47                  Mueslix_Crispy_Blend   K    C      160        3   2
150
## 48                  Multi-Grain_Cheerios   G    C      100        2   1
220
## 49                      Nut&Honey_Crunch   K    C      120        2   1
190
## 50              Nutri-Grain_Almond-Raisin   K    C      140        3   2
220
## 51                     Nutri-grain_Wheat   K    C       90        3   0
170
## 52                  Oatmeal_Raisin_Crisp   G    C      130        3   2
170
## 53                 Post_Nat._Raisin_Bran   P    C      120        3   1
200
## 54                             Product_19   K    C      100        3   0
320
## 55                            Puffed_Rice   Q    C       50        1   0
0
## 56                           Puffed_Wheat   Q    C       50        2   0
0
## 57                     Quaker_Oat_Squares   Q    C      100        4   1
135
## 58                         Quaker_Oatmeal   Q    H      100        5   2
0
## 59                            Raisin_Bran   K    C      120        3   1
210
## 60                        Raisin_Nut_Bran   G    C      100        3   2
140
## 61                         Raisin_Squares   K    C       90        2   0
0
## 62                              Rice_Chex   R    C      110        1   0
240
## 63                           Rice_Krispies   K    C      110        2   0
290
## 64                         Shredded_Wheat   N    C       80        2   0
0
```

```
## 65                  Shredded_Wheat_'n'Bran   N    C       90       3   0
0
## 66               Shredded_Wheat_spoon_size   N    C       90       3   0
0
## 67                                  Smacks   K    C      110       2   1
70
## 68                               Special_K   K    C      110       6   0
230
## 69                 Strawberry_Fruit_Wheats   N    C       90       2   0
15
## 70                       Total_Corn_Flakes   G    C      110       2   1
200
## 71                       Total_Raisin_Bran   G    C      140       3   1
190
## 72                       Total_Whole_Grain   G    C      100       3   1
200
## 73                                 Triples   G    C      110       2   1
250
## 74                                    Trix   G    C      110       1   1
140
## 75                               Wheat_Chex   R    C      100       3   1
230
## 76                                 Wheaties   G    C      100       3   1
200
## 77                       Wheaties_Honey_Gold   G    C      110       2   1
200
##    fiber carbo sugars potass vitamins shelf weight cups   rating
## 1   10.0   5.0      6    280       25     3   1.00 0.33 68.40297
## 2    2.0   8.0      8    135        0     3   1.00 1.00 33.98368
## 3    9.0   7.0      5    320       25     3   1.00 0.33 59.42551
## 4   14.0   8.0      0    330       25     3   1.00 0.50 93.70491
## 5    1.0  14.0      8     NA       25     3   1.00 0.75 34.38484
## 6    1.5  10.5     10     70       25     1   1.00 0.75 29.50954
## 7    1.0  11.0     14     30       25     2   1.00 1.00 33.17409
## 8    2.0  18.0      8    100       25     3   1.33 0.75 37.03856
## 9    4.0  15.0      6    125       25     1   1.00 0.67 49.12025
## 10   5.0  13.0      5    190       25     3   1.00 0.67 53.31381
## 11   0.0  12.0     12     35       25     2   1.00 0.75 18.04285
## 12   2.0  17.0      1    105       25     1   1.00 1.25 50.76500
## 13   0.0  13.0      9     45       25     2   1.00 0.75 19.82357
## 14   2.0  13.0      7    105       25     3   1.00 0.50 40.40021
## 15   0.0  12.0     13     55       25     2   1.00 1.00 22.73645
## 16   0.0  22.0      3     25       25     1   1.00 1.00 41.44502
## 17   1.0  21.0      2     35       25     1   1.00 1.00 45.86332
## 18   1.0  13.0     12     20       25     2   1.00 1.00 35.78279
## 19   0.0  12.0     13     65       25     2   1.00 1.00 22.39651
## 20   4.0  10.0      7    160       25     3   1.00 0.50 40.44877
## 21   1.0  21.0      0     NA        0     2   1.00 1.00 64.53382
## 22   1.0  21.0      3     30       25     3   1.00 1.00 46.89564
## 23   2.0  11.0     10    120       25     3   1.00 0.75 36.17620
```

```
## 24   1.0  18.0      5     80       25     3   1.00 0.75 44.33086
## 25   1.0  11.0     13     30       25     2   1.00 1.00 32.20758
## 26   1.0  14.0     11     25       25     1   1.00 0.75 31.43597
## 27   3.0  14.0      7    100       25     2   1.00 0.80 58.34514
## 28   5.0  12.0     10    200       25     3   1.25 0.67 40.91705
## 29   5.0  14.0     12    190       25     3   1.33 0.67 41.01549
## 30   0.0  13.0     12     25       25     2   1.00 0.75 28.02576
## 31   0.0  11.0     15     40       25     1   1.00 0.88 35.25244
## 32   0.0  15.0      9     45       25     2   1.00 0.75 23.80404
## 33   3.0  15.0      5     85       25     3   1.00 0.88 52.07690
## 34   3.0  17.0      3     90       25     3   1.00 0.25 53.37101
## 35   3.0  13.0      4    100       25     3   1.00 0.33 45.81172
## 36   1.0  12.0     11     45       25     2   1.00 1.00 21.87129
## 37   1.5  11.5     10     90       25     1   1.00 0.75 31.07222
## 38   0.0  14.0     11     35       25     1   1.00 1.33 28.74241
## 39   1.0  17.0      6     60      100     3   1.00 1.00 36.52368
## 40   2.0  20.0      9     95      100     3   1.30 0.75 36.47151
## 41   0.0  21.0      3     40       25     2   1.00 1.50 39.24111
## 42   2.0  12.0      6     95       25     2   1.00 0.67 45.32807
## 43   0.0  12.0     12     55       25     2   1.00 1.00 26.73451
## 44   0.0  16.0      3     95       25     2   1.00 1.00 54.85092
## 45   3.0  16.0     11    170       25     3   1.00 1.00 37.13686
## 46   3.0  16.0     11    170       25     3   1.00 1.00 34.13976
## 47   3.0  17.0     13    160       25     3   1.50 0.67 30.31335
## 48   2.0  15.0      6     90       25     1   1.00 1.00 40.10596
## 49   0.0  15.0      9     40       25     2   1.00 0.67 29.92429
## 50   3.0  21.0      7    130       25     3   1.33 0.67 40.69232
## 51   3.0  18.0      2     90       25     3   1.00 1.00 59.64284
## 52   1.5  13.5     10    120       25     3   1.25 0.50 30.45084
## 53   6.0  11.0     14    260       25     3   1.33 0.67 37.84059
## 54   1.0  20.0      3     45      100     3   1.00 1.00 41.50354
## 55   0.0  13.0      0     15        0     3   0.50 1.00 60.75611
## 56   1.0  10.0      0     50        0     3   0.50 1.00 63.00565
## 57   2.0  14.0      6    110       25     3   1.00 0.50 49.51187
## 58   2.7    NA     NA    110        0     1   1.00 0.67 50.82839
## 59   5.0  14.0     12    240       25     2   1.33 0.75 39.25920
## 60   2.5  10.5      8    140       25     3   1.00 0.50 39.70340
## 61   2.0  15.0      6    110       25     3   1.00 0.50 55.33314
## 62   0.0  23.0      2     30       25     1   1.00 1.13 41.99893
## 63   0.0  22.0      3     35       25     1   1.00 1.00 40.56016
## 64   3.0  16.0      0     95        0     1   0.83 1.00 68.23588
## 65   4.0  19.0      0    140        0     1   1.00 0.67 74.47295
## 66   3.0  20.0      0    120        0     1   1.00 0.67 72.80179
## 67   1.0   9.0     15     40       25     2   1.00 0.75 31.23005
## 68   1.0  16.0      3     55       25     1   1.00 1.00 53.13132
## 69   3.0  15.0      5     90       25     2   1.00 1.00 59.36399
## 70   0.0  21.0      3     35      100     3   1.00 1.00 38.83975
## 71   4.0  15.0     14    230      100     3   1.50 1.00 28.59278
## 72   3.0  16.0      3    110      100     3   1.00 1.00 46.65884
## 73   0.0  21.0      3     60       25     3   1.00 0.75 39.10617
```

```
## 74   0.0  13.0     12     25       25     2   1.00 1.00 27.75330
## 75   3.0  17.0      3    115       25     1   1.00 0.67 49.78744
## 76   3.0  17.0      3    110       25     1   1.00 1.00 51.59219
## 77   1.0  16.0      8     60       25     1   1.00 0.75 36.18756
```

To view if the data has any missing values

```
#View(cereals)
```

To remove the missing values present in the data

```
cdata<- na.omit(cereals)
#view(cdata)
```

#QUESTION 1 Apply hierarchical clustering to the data using Euclidean distance to the normalized measurements. Use Agnes to compare the clustering from single linkage, complete linkage, average linkage, and Ward. Choose the best method. # Normalize the data

```
cdata_norm <- cbind(cdata[, 1:3], scale(cdata[, -c(1:3)]))
#view(cdata_norm)
```

#Apply hierarchical clustering to the data using Euclidean distance

```
dist <- dist(cdata, method = "euclidean", diag = FALSE, upper = FALSE, p =2)

## Warning in dist(cdata, method = "euclidean", diag = FALSE, upper = FALSE,
: NAs
## introduced by coercion

#dist
```

#clustering the data and comparing them using agnes

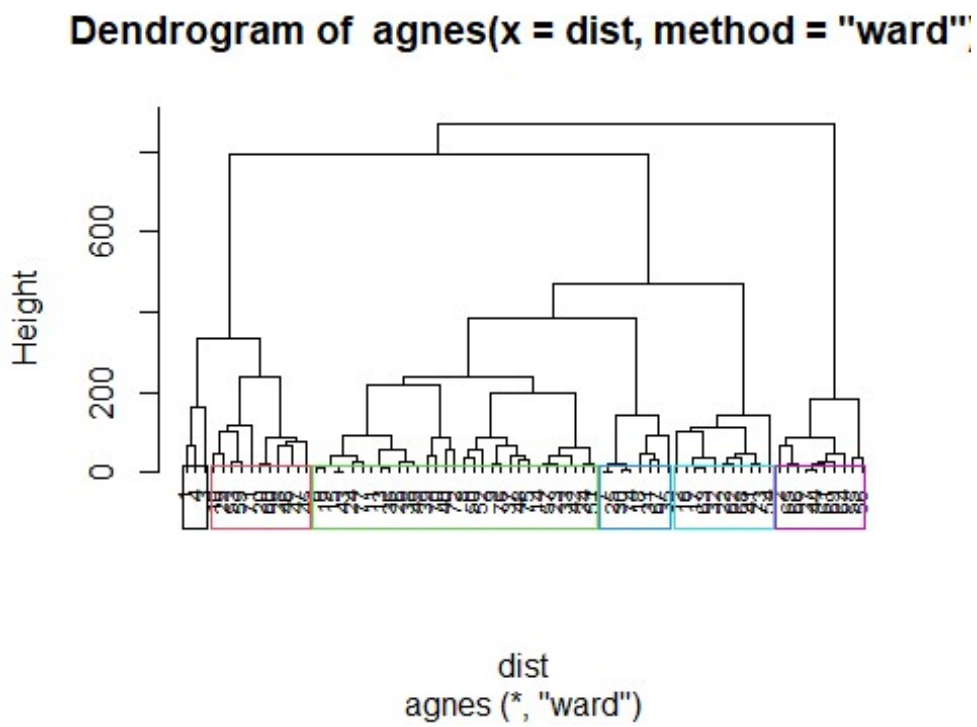
```
single_clust <- agnes(dist, method = "single")
print(single_clust)

## Call:     agnes(x = dist, method = "single")
## Agglomerative coefficient:  0.7311616
## Order of objects:
##  [1] 1  4  2  27 44 61 69 64 65 66 55 56 6  19 15 43 38 49 26 11 13 36 77
24 14
## [26] 57 33 42 23 60 20 34 51 16 17 63 32 22 62 41 73 68 9  76 48 75 37 8
50 52
## [51] 7  25 30 74 18 31 67 46 47 28 12 45 10 29 53 59 35 39 70 40 72 71 54
3
## Height (summary):
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.90   20.64   28.69   36.13   48.28  112.44
##
## Available components:
```

```
## [1] "order"     "height"    "ac"        "merge"     "diss"      "call"
## [7] "method"    "order.lab"

# the AGGLOMERATIVE COEFFICIENT using single linkage method is 0.7311616

complete_clust <- agnes(dist, method = "complete")
print(complete_clust)

## Call:     agnes(x = dist, method = "complete")
## Agglomerative coefficient:  0.922957
## Order of objects:
##  [1] 1  4  3  10 29 53 59 71 20 60 28 45 46 47 2  65 66 27 44 61 69 64 18
31 67
## [26] 35 55 56 6  19 15 43 77 11 13 36 26 38 49 7  25 30 74 8  50 52 9  76
24 34
## [51] 51 14 57 23 33 42 39 70 40 72 12 37 48 75 16 17 63 32 22 62 68 41 73
54
## Height (summary):
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.90   22.32   43.77   70.53   84.64  446.26
##
## Available components:
## [1] "order"     "height"    "ac"        "merge"     "diss"      "call"
## [7] "method"    "order.lab"

#using complete linkage method the AGGLOMERATIVE COEFFICIENT is 0.922957

avg_clust <- agnes(dist, method = "average")
print(avg_clust)

## Call:     agnes(x = dist, method = "average")
## Agglomerative coefficient:  0.8792621
## Order of objects:
##  [1] 1  4  3  2  27 44 61 69 64 65 66 35 55 56 6  19 15 43 77 26 38 49 11
13 36
## [26] 14 57 23 33 42 24 34 51 52 8  50 9  76 48 75 12 37 16 17 63 32 22 62
68 41
## [51] 73 39 70 40 72 7  25 30 74 18 31 67 10 29 53 59 71 20 60 28 46 47 45
54
## Height (summary):
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.90   22.32   37.90   54.82   71.16  280.59
##
## Available components:
## [1] "order"     "height"    "ac"        "merge"     "diss"      "call"
## [7] "method"    "order.lab"

#using average linkage method the AGGLOMERATIVE COEFFICIENT is 0.8792621

ward_clust <- agnes(dist, method = "ward")
print(ward_clust)
```

```
## Call:     agnes(x = dist, method = "ward")
## Agglomerative coefficient:  0.9597071
## Order of objects:
##  [1] 1  4  3  10 29 53 59 71 20 60 28 46 47 45 6  19 15 43 24 77 11 13 36
26 38
## [26] 49 39 70 40 72 8  50 52 9  76 37 48 75 14 57 23 33 42 34 51 7  25 30
74 18
## [51] 31 67 35 12 16 17 63 32 22 62 68 41 73 54 2  65 66 27 44 61 69 64 55
56
## Height (summary):
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.90   22.89   44.72   96.14   91.98  869.24
##
## Available components:
## [1] "order"     "height"    "ac"        "merge"     "diss"      "call"
## [7] "method"    "order.lab"

#using ward linkage method the AGGLOMERATIVE COEFFICIENT is 0.9597071
```

by comparing the agglomerative coefficient values ward linkage method is the best method.

```
pltree(ward_clust, cex = 0.6, hang = -1)
rect.hclust(ward_clust, k = 6, border = 1:6)
```



```
Model_1 <- cutree(ward_clust, 6)
data <- cbind(data,Model_1)
```

QUESTION 3 cluster structure and stability

Cluster partition A

```
set.seed(123)
dataA_index <- sample(seq_len(nrow(cdata_norm)), size = 67)
dataA <- cdata_norm[dataA_index,]
dataB <- cdata_norm[-dataA_index,]
dataA
```

```
##                                       name mfr type   calories    protein
## 33                       Grape_Nuts_Flakes   P    C -0.3541153  0.4522084
## 53                    Post_Nat._Raisin_Bran   P    C  0.6537514  0.4522084
## 15                               Cocoa_Puffs   G    C  0.1498180 -1.4068705
## 70                          Total_Corn_Flakes   G    C  0.1498180 -0.4773310
## 44                                      Maypo   A    H -0.3541153  1.3817478
## 52                     Oatmeal_Raisin_Crisp   G    C  1.1576848  0.4522084
## 45       Muesli_Raisins,_Dates,_&_Almonds   R    C  2.1655516  1.3817478
## 75                                Wheat_Chex   R    C -0.3541153  0.4522084
## 27                      Frosted_Mini-Wheats   K    C -0.3541153  0.4522084
## 60                          Raisin_Nut_Bran   G    C -0.3541153  0.4522084
## 30                           Fruity_Pebbles   P    C  0.1498180 -1.4068705
## 68                                Special_K   K    C  0.1498180  3.2408266
## 10                               Bran_Flakes   P    C -0.8580487  0.4522084
## 31                             Golden_Crisp   P    C -0.3541153 -0.4773310
## 37                       Honey_Nut_Cheerios   G    C  0.1498180  0.4522084
## 9                                   Bran_Chex   R    C -0.8580487 -0.4773310
## 28 Fruit_&_Fibre_Dates,_Walnuts,_and_Oats   P    C  0.6537514  0.4522084
## 8                                     Basic_4   G    C  1.1576848  0.4522084
## 73                                    Triples   G    C  0.1498180 -0.4773310
## 65                  Shredded_Wheat_'n'Bran   N    C -0.8580487  0.4522084
## 20                      Cracklin'_Oat_Bran   K    C  0.1498180  0.4522084
## 38                                Honey-comb   P    C  0.1498180 -1.4068705
## 74                                      Trix   G    C  0.1498180 -1.4068705
## 18                                 Corn_Pops   K    C  0.1498180 -1.4068705
## 71                       Total_Raisin_Bran   G    C  1.6616182  0.4522084
## 41                                       Kix   G    C  0.1498180 -0.4773310
## 13                    Cinnamon_Toast_Crunch   G    C  0.6537514 -1.4068705
## 16                                 Corn_Chex   R    C  0.1498180 -0.4773310
## 34                               Grape-Nuts   P    C  0.1498180  0.4522084
## 59                               Raisin_Bran   K    C  0.6537514  0.4522084
## 66               Shredded_Wheat_spoon_size   N    C -0.8580487  0.4522084
## 57                      Quaker_Oat_Squares   Q    C -0.3541153  1.3817478
## 43                             Lucky_Charms   G    C  0.1498180 -0.4773310
## 11                             Cap'n'Crunch   Q    C  0.6537514 -1.4068705
```

```
## 25                            Froot_Loops   K    C  0.1498180 -0.4773310
## 29                          Fruitful_Bran   K    C  0.6537514  0.4522084
## 46      Muesli_Raisins,_Peaches,_&_Pecans   R    C  2.1655516  1.3817478
## 51                      Nutri-grain_Wheat   K    C -0.8580487  0.4522084
## 48                  Multi-Grain_Cheerios   G    C -0.3541153 -0.4773310
## 69                Strawberry_Fruit_Wheats   N    C -0.8580487 -0.4773310
## 36                      Honey_Graham_Ohs   Q    C  0.6537514 -1.4068705
## 64                        Shredded_Wheat   N    C -1.3619821 -0.4773310
## 6                Apple_Cinnamon_Cheerios   G    C  0.1498180 -0.4773310
## 62                              Rice_Chex   R    C  0.1498180 -1.4068705
## 50            Nutri-Grain_Almond-Raisin   K    C  1.6616182  0.4522084
## 14                              Clusters   G    C  0.1498180  0.4522084
## 19                         Count_Chocula   G    C  0.1498180 -1.4068705
## 1                              100%_Bran   N    C -1.8659155  1.3817478
## 63                         Rice_Krispies   K    C  0.1498180 -0.4773310
## 61                        Raisin_Squares   K    C -0.8580487 -0.4773310
## 7                            Apple_Jacks   K    C  0.1498180 -0.4773310
## 23                Crispy_Wheat_&_Raisins   G    C -0.3541153 -0.4773310
## 49                      Nut&Honey_Crunch   K    C  0.6537514 -0.4773310
## 72                     Total_Whole_Grain   G    C -0.3541153  0.4522084
## 24                           Double_Chex   R    C -0.3541153 -0.4773310
## 17                           Corn_Flakes   K    C -0.3541153 -0.4773310
## 26                        Frosted_Flakes   K    C  0.1498180 -1.4068705
## 12                               Cheerios   G    C  0.1498180  3.2408266
## 77                   Wheaties_Honey_Gold   G    C  0.1498180 -0.4773310
## 67                                Smacks   K    C  0.1498180 -0.4773310
## 22                               Crispix   K    C  0.1498180 -0.4773310
## 40                Just_Right_Fruit_&_Nut   K    C  1.6616182  0.4522084
## 39            Just_Right_Crunchy__Nuggets   K    C  0.1498180 -0.4773310
## 54                             Product_19   K    C -0.3541153  0.4522084
## 2                     100%_Natural_Bran   Q    C  0.6537514  0.4522084
## 32                       Golden_Grahams   G    C  0.1498180 -1.4068705
## 47                  Mueslix_Crispy_Blend   K    C  2.6694849  0.4522084
##           fat      sodium       fiber       carbo      sugars      potass
## 33  0.0000000 -0.27020566  0.34015322  0.06944832 -0.48360961 -0.19065695
## 53  0.0000000  0.45469653  1.57808790 -0.95838683  1.58103142  2.27835060
## 15  0.0000000  0.21306247 -0.89778146 -0.70142805  1.35162686 -0.61391539
## 70  0.0000000  0.45469653 -0.89778146  1.61120105 -0.94241873 -0.89608768
## 44  0.0000000 -1.96164410 -0.89778146  0.32640711 -0.94241873 -0.04957081
## 52  0.9932203  0.09224544 -0.27881412 -0.31598986  0.66341318  0.30314456
## 45  1.9864405 -0.81388230  0.34015322  0.32640711  0.89281774  1.00857529
## 75  0.0000000  0.81714763  0.34015322  0.58336590 -0.94241873  0.23260148
## 27 -0.9932203 -1.96164410  0.34015322 -0.18751047 -0.02480049  0.02097226
## 60  0.9932203 -0.27020566  0.13383078 -1.08686623  0.20460407  0.58531685
## 30  0.0000000 -0.33061417 -0.89778146 -0.44446926  1.12222230 -1.03717383
## 68 -0.9932203  0.81714763 -0.48513656  0.32640711 -0.94241873 -0.61391539
## 10 -0.9932203  0.57551356  1.16544301 -0.44446926 -0.48360961  1.29074758
## 31 -0.9932203 -1.41796746 -0.89778146 -0.95838683  1.81043598 -0.82554461
## 37  0.0000000  1.05878169 -0.27881412 -0.82990744  0.66341318 -0.12011388
## 9   0.0000000  0.45469653  0.75279812  0.06944832 -0.25420505  0.37368763
```

```
## 28  0.9932203 -0.02857160  1.16544301 -0.70142805  0.66341318  1.43183372
## 8   0.9932203  0.57551356 -0.07249167  0.84032469  0.20460407  0.02097226
## 73  0.0000000  1.05878169 -0.89778146  1.61120105 -0.94241873 -0.54337232
## 65 -0.9932203 -1.96164410  0.75279812  1.09728348 -1.63063240  0.58531685
## 20  1.9864405 -0.27020566  0.75279812 -1.21534562 -0.02480049  0.86748914
## 38 -0.9932203  0.21306247 -0.89778146 -0.18751047  0.89281774 -0.89608768
## 74  0.0000000 -0.27020566 -0.89778146 -0.44446926  1.12222230 -1.03717383
## 18 -0.9932203 -0.87429082 -0.48513656 -0.44446926  1.12222230 -1.10771690
## 71  0.0000000  0.33387950  0.75279812  0.06944832  1.58103142  1.85509216
## 41  0.0000000  1.17959872 -0.89778146  1.61120105 -0.94241873 -0.82554461
## 13  1.9864405  0.57551356 -0.89778146 -0.44446926  0.43400862 -0.75500154
## 16 -0.9932203  1.42123279 -0.89778146  1.86815984 -0.94241873 -1.03717383
## 34 -0.9932203  0.09224544  0.34015322  0.58336590 -0.94241873 -0.12011388
## 59  0.0000000  0.57551356  1.16544301 -0.18751047  1.12222230  1.99617831
## 66 -0.9932203 -1.96164410  0.34015322  1.35424227 -1.63063240  0.30314456
## 57  0.0000000 -0.33061417 -0.07249167 -0.18751047 -0.25420505  0.16205841
## 43  0.0000000  0.21306247 -0.89778146 -0.70142805  1.12222230 -0.61391539
## 11  0.9932203  0.69633060 -0.89778146 -0.70142805  1.12222230 -0.89608768
## 25  0.0000000 -0.45143121 -0.48513656 -0.95838683  1.35162686 -0.96663076
## 29 -0.9932203  0.93796466  1.16544301 -0.18751047  1.12222230  1.29074758
## 46  1.9864405 -0.14938863  0.34015322  0.32640711  0.89281774  1.00857529
## 51 -0.9932203  0.09224544  0.34015322  0.84032469 -1.17182329 -0.12011388
## 48  0.0000000  0.69633060 -0.07249167  0.06944832 -0.25420505 -0.12011388
## 69 -0.9932203 -1.78041856  0.34015322  0.06944832 -0.48360961 -0.12011388
## 36  0.9932203  0.69633060 -0.48513656 -0.70142805  0.89281774 -0.75500154
## 64 -0.9932203 -1.96164410  0.34015322  0.32640711 -1.63063240 -0.04957081
## 6   0.9932203  0.21306247 -0.27881412 -1.08686623  0.66341318 -0.40228617
## 62 -0.9932203  0.93796466 -0.89778146  2.12511863 -1.17182329 -0.96663076
## 50  0.9932203  0.69633060  0.34015322  1.61120105 -0.02480049  0.44423070
## 14  0.9932203 -0.27020566 -0.07249167 -0.44446926 -0.02480049  0.09151534
## 19  0.0000000  0.21306247 -0.89778146 -0.70142805  1.35162686 -0.47282925
## 1   0.0000000 -0.39102269  3.22866747 -2.50013957 -0.25420505  2.56052289
## 63 -0.9932203  1.54204982 -0.89778146  1.86815984 -0.94241873 -0.89608768
## 61 -0.9932203 -1.96164410 -0.07249167  0.06944832 -0.25420505  0.16205841
## 7  -0.9932203 -0.45143121 -0.48513656 -0.95838683  1.58103142 -0.96663076
## 23  0.0000000 -0.27020566 -0.07249167 -0.95838683  0.66341318  0.30314456
## 49  0.0000000  0.33387950 -0.89778146  0.06944832  0.43400862 -0.82554461
## 72  0.0000000  0.45469653  0.34015322  0.32640711 -0.94241873  0.16205841
## 24 -0.9932203  0.33387950 -0.48513656  0.84032469 -0.48360961 -0.26120003
## 17 -0.9932203  1.54204982 -0.48513656  1.61120105 -1.17182329 -0.89608768
## 26 -0.9932203  0.45469653 -0.48513656 -0.18751047  0.89281774 -1.03717383
## 12  0.9932203  1.54204982 -0.07249167  0.58336590 -1.40122785  0.09151534
## 77  0.0000000  0.45469653 -0.48513656  0.32640711  0.20460407 -0.54337232
## 67  0.0000000 -1.11592488 -0.48513656 -1.47230441  1.81043598 -0.82554461
## 22 -0.9932203  0.69633060 -0.48513656  1.61120105 -0.94241873 -0.96663076
## 40  0.0000000  0.09224544 -0.07249167  1.35424227  0.43400862 -0.04957081
## 39  0.0000000  0.09224544 -0.48513656  0.58336590 -0.25420505 -0.54337232
## 54 -0.9932203  1.90450091 -0.48513656  1.35424227 -0.94241873 -0.75500154
## 2   3.9728810 -1.78041856 -0.07249167 -1.72926320  0.20460407  0.51477378
## 32  0.0000000  1.42123279 -0.89778146  0.06944832  0.43400862 -0.75500154
```

```
## 47  0.9932203 -0.14938863  0.34015322  0.58336590  1.35162686  0.86748914
##      vitamins      shelf     weight        cups      rating
## 33 -0.1818422  0.9419715 -0.2008324  0.24766475  0.69155685
## 53 -0.1818422  0.9419715  1.9501886 -0.64324039 -0.32287913
## 15 -0.1818422 -0.2598542 -0.2008324  0.75675340 -1.39915514
## 70  3.1822385  0.9419715 -0.2008324  0.75675340 -0.25168258
## 44 -0.1818422 -0.2598542 -0.2008324  0.75675340  0.88922515
## 52 -0.1818422  0.9419715  1.4287290 -1.36444931 -0.84945049
## 45 -0.1818422  0.9419715 -0.2008324  0.75675340 -0.37302488
## 75 -0.1818422 -1.4616799 -0.2008324 -0.64324039  0.52841741
## 27 -0.1818422 -0.2598542 -0.2008324 -0.09172768  1.13821301
## 60 -0.1818422  0.9419715 -0.2008324 -1.36444931 -0.19014120
## 30 -0.1818422 -0.2598542 -0.2008324 -0.30384795 -1.02225423
## 68 -0.1818422 -1.4616799 -0.2008324  0.75675340  0.76669214
## 10 -0.1818422  0.9419715 -0.2008324 -0.64324039  0.77969576
## 31 -0.1818422 -1.4616799 -0.2008324  0.24766475 -0.50730289
## 37 -0.1818422 -1.4616799 -0.2008324 -0.30384795 -0.80517325
## 9  -0.1818422 -1.4616799 -0.2008324 -0.64324039  0.48087533
## 28 -0.1818422  0.9419715  1.4287290 -0.64324039 -0.10366038
## 8  -0.1818422  0.9419715  1.9501886 -0.30384795 -0.38002951
## 73 -0.1818422  0.9419715 -0.2008324 -0.30384795 -0.23269772
## 65 -1.3032024 -1.4616799 -0.2008324 -0.64324039  2.28743193
## 20 -0.1818422  0.9419715 -0.2008324 -1.36444931 -0.13702824
## 38 -0.1818422 -1.4616799 -0.2008324  2.15674718 -0.97118798
## 74 -0.1818422 -0.2598542 -0.2008324  0.75675340 -1.04166919
## 18 -0.1818422 -0.2598542 -0.2008324  0.75675340 -0.46951197
## 71  3.1822385  0.9419715  3.0582904  0.75675340 -0.98185009
## 41 -0.1818422 -0.2598542 -0.2008324  2.87795610 -0.22308231
## 13 -0.1818422 -0.2598542 -0.2008324 -0.30384795 -1.60671768
## 16 -0.1818422 -1.4616799 -0.2008324  0.75675340 -0.06603869
## 34 -0.1818422  0.9419715 -0.2008324 -2.42505066  0.78377123
## 59 -0.1818422 -0.2598542  1.9501886 -0.30384795 -0.22179377
## 66 -1.3032024 -1.4616799 -0.2008324 -0.64324039  2.16834997
## 57 -0.1818422  0.9419715 -0.2008324 -1.36444931  0.50878106
## 43 -0.1818422 -0.2598542 -0.2008324  0.75675340 -1.11426481
## 11 -0.1818422 -0.2598542 -0.2008324 -0.30384795 -1.73360655
## 25 -0.1818422 -0.2598542 -0.2008324  0.75675340 -0.72427057
## 29 -0.1818422  0.9419715  1.9501886 -0.64324039 -0.09664548
## 46 -0.1818422  0.9419715 -0.2008324  0.75675340 -0.58658904
## 51 -0.1818422  0.9419715 -0.2008324  0.75675340  1.23068291
## 48 -0.1818422 -1.4616799 -0.2008324  0.75675340 -0.16145563
## 69 -0.1818422 -0.2598542 -0.2008324  0.75675340  1.21081332
## 36 -0.1818422 -0.2598542 -0.2008324  0.75675340 -1.46080340
## 64 -1.3032024 -1.4616799 -1.3089342  0.75675340  1.84299757
## 6  -0.1818422 -1.4616799 -0.2008324 -0.30384795 -0.91652483
## 62 -0.1818422 -1.4616799 -0.2008324  1.30826610 -0.02656845
## 50 -0.1818422  0.9419715  1.9501886 -0.64324039 -0.11967375
## 14 -0.1818422  0.9419715 -0.2008324 -1.36444931 -0.14048876
## 19 -0.1818422 -0.2598542 -0.2008324  0.75675340 -1.42337774
## 1  -0.1818422  0.9419715 -0.2008324 -2.08565823  1.85490376
```

```
## 63 -0.1818422 -1.4616799 -0.2008324  0.75675340 -0.12909114
## 61 -0.1818422  0.9419715 -0.2008324 -1.36444931  0.92358705
## 7  -0.1818422 -0.2598542 -0.2008324  0.75675340 -0.65539984
## 23 -0.1818422  0.9419715 -0.2008324 -0.30384795 -0.44147911
## 49 -0.1818422 -0.2598542 -0.2008324 -0.64324039 -0.88697142
## 72  3.1822385  0.9419715 -0.2008324  0.75675340  0.30548275
## 24 -0.1818422  0.9419715 -0.2008324 -0.30384795  0.13959735
## 17 -0.1818422 -1.4616799 -0.2008324  0.75675340  0.24879639
## 26 -0.1818422 -1.4616799 -0.2008324 -0.30384795 -0.77925310
## 12 -0.1818422 -1.4616799 -0.2008324  1.81735475  0.59807496
## 77 -0.1818422 -1.4616799 -0.2008324 -0.30384795 -0.44066942
## 67 -0.1818422 -0.2598542 -0.2008324 -0.30384795 -0.79392626
## 22 -0.1818422  0.9419715 -0.2008324  0.75675340  0.32235640
## 40  3.1822385  0.9419715  1.7546413 -0.30384795 -0.42043579
## 39  3.1822385  0.9419715 -0.2008324  0.75675340 -0.41671824
## 54  3.1822385  0.9419715 -0.2008324  0.75675340 -0.06186866
## 2  -1.3032024  0.9419715 -0.2008324  0.75675340 -0.59771126
## 32 -0.1818422 -0.2598542 -0.2008324 -0.30384795 -1.32308140
## 47 -0.1818422  0.9419715  3.0582904 -0.64324039 -0.85924775
```

#Use the cluster centroids from A to assign each record in partition B (each record is assigned to the cluster with the closest centroid).

```
dataA_dist <- dist(dataA[, -c(1:3)], method = "euclidean")
hc_A <- hclust(dataA_dist, method = "ward.D")
Model_A <- cutree(hc_A, 6)
dataA <- cbind(Model_A, dataA)
head(dataA)
```

```
##    Model_A                 name mfr type   calories    protein       fat
## 33       1     Grape_Nuts_Flakes   P    C -0.3541153  0.4522084 0.0000000
## 53       2 Post_Nat._Raisin_Bran   P    C  0.6537514  0.4522084 0.0000000
## 15       3           Cocoa_Puffs   G    C  0.1498180 -1.4068705 0.0000000
## 70       4     Total_Corn_Flakes   G    C  0.1498180 -0.4773310 0.0000000
## 44       5                 Maypo   A    H -0.3541153  1.3817478 0.0000000
## 52       2  Oatmeal_Raisin_Crisp   G    C  1.1576848  0.4522084 0.9932203
##         sodium      fiber       carbo     sugars      potass   vitamins
## 33 -0.27020566  0.3401532  0.06944832 -0.4836096 -0.19065695 -0.1818422
## 53  0.45469653  1.5780879 -0.95838683  1.5810314  2.27835060 -0.1818422
## 15  0.21306247 -0.8977815 -0.70142805  1.3516269 -0.61391539 -0.1818422
## 70  0.45469653 -0.8977815  1.61120105 -0.9424187 -0.89608768  3.1822385
## 44 -1.96164410 -0.8977815  0.32640711 -0.9424187 -0.04957081 -0.1818422
## 52  0.09224544 -0.2788141 -0.31598986  0.6634132  0.30314456 -0.1818422
##         shelf     weight       cups     rating
## 33  0.9419715 -0.2008324  0.2476647  0.6915569
## 53  0.9419715  1.9501886 -0.6432404 -0.3228791
## 15 -0.2598542 -0.2008324  0.7567534 -1.3991551
## 70  0.9419715 -0.2008324  0.7567534 -0.2516826
## 44 -0.2598542 -0.2008324  0.7567534  0.8892251
## 52  0.9419715  1.4287290 -1.3644493 -0.8494505
```

```
C1 <- colMeans(dataA[dataA$Model_A == 1,5:ncol(dataA)])
C2 <- colMeans(dataA[dataA$Model_A == 2,5:ncol(dataA)])
C3 <- colMeans(dataA[dataA$Model_A == 3,5:ncol(dataA)])
C4 <- colMeans(dataA[dataA$Model_A == 4,5:ncol(dataA)])
C5 <- colMeans(dataA[dataA$Model_A == 5,5:ncol(dataA)])
C6 <- colMeans(dataA[dataA$Model_A == 6,5:ncol(dataA)])
Centroids_A <- rbind(C1, C2, C3, C4, C5, C6)
dist <- dist(rbind(Centroids_A,dataB[,4:ncol(dataB)]))
dist
```

```
##          C1       C2       C3       C4       C5       C6        3        4
## C2 3.087519
## C3 3.171116 3.766283
## C4 4.080860 4.394851 4.465807
## C5 3.120031 5.017983 4.225386 5.574752
## C6 3.992549 5.064938 3.600257 4.537638 4.243322
## 3  5.001063 6.043032 7.197863 7.486394 6.477277 7.838851
## 4  7.300259 8.670076 9.647012 9.528843 7.407964 9.662959 4.031051
## 35 2.937467 3.483268 4.492378 5.395800 4.485411 5.823893 5.940473 8.282237
## 42 2.116481 3.253686 3.177192 4.677708 3.523530 4.044877 5.449421 7.913299
## 55 5.871507 8.431790 6.391470 8.139591 4.910302 6.737769 8.695724 9.343676
## 56 5.617653 8.188194 6.445799 8.129250 4.710236 6.816139 7.916319 8.492807
## 76 3.013139 4.262622 3.278687 4.537729 3.020609 2.070710 6.290066 8.004134
##          35       42       55       56
## C2
## C3
## C4
## C5
## C6
## 3
## 4
## 35
## 42 2.768632
## 55 7.106645 6.462448
## 56 6.861611 6.052896 1.377717
## 76 4.732649 2.856300 6.290658 6.120699
```

#c. Assess how consistent the cluster assignments are compared to the assignments based on all the data.

```
Assignments<- data.frame(data[-dataA_index, "Model_1"])
Assignments$Model_A <- 0
Assignments[1,2] <-
which.max(c(2.462441,3.146983,3.782227,5.983747,4.881300,3.826735))
Assignments[2,2] <- which.max(c(2.312626,3.855177, 4.581339, 8.455727
,4.827885 ,5.441721))
Assignments[3,2] <- which.max(c(3.437484, 1.869957, 4.404072, 8.395433,
4.851045, 4.429787))
Assignments[4,2] <- which.max(c(3.339714, 1.285089 ,4.752289, 8.059830,
4.687570, 3.657016))
```

```
Assignments[5,2] <- which.max(c(3.007575 ,3.930439 ,2.431885, 6.230808,
5.027546, 4.008456))
Assignments[6,2] <- which.max(c(3.159701 ,4.694503 ,4.418982, 6.075311,
4.930370, 3.901197))
Assignments[7,2] <- which.max(c(2.691188 ,3.357138, 4.095313, 6.119492,
4.677708, 2.705759))
Assignments
```

```
##   X3 X4 X35 X42 X55 X56 X76 Model_A
## 1  1  4   4   3   2   2   3       0
## 2 NA  4  NA  NA  NA  NA  NA      NA
## 3 NA  4  NA  NA  NA  NA  NA      NA
## 4 NA  4  NA  NA  NA  NA  NA      NA
## 5 NA  4  NA  NA  NA  NA  NA      NA
## 6 NA  4  NA  NA  NA  NA  NA      NA
## 7 NA  4  NA  NA  NA  NA  NA      NA
```

#The elementary public schools would like to choose a set of cereals to include in their daily cafeterias. Every day a different cereal is offered, but all cereals should support a healthy diet. For this goal, you are requested to find a cluster of "healthy cereals." Should the data be normalized? If not, how should they be used in the cluster analysis?

```
healthycereals <- cereals
healthycereals_na <- na.omit(healthycereals)
Clusthealthy <- cbind(healthycereals_na, Model_1)
Clusthealthy[Clusthealthy$Model_1==1,]
```

```
##                        name mfr type calories protein fat sodium fiber
carbo
## 1                 100%_Bran   N    C       70       4   1    130    10
5
## 3                  All-Bran   K    C       70       4   1    260     9
7
## 4 All-Bran_with_Extra_Fiber   K    C       50       4   0    140    14
8
##   sugars potass vitamins shelf weight cups   rating Model_1
## 1      6    280       25     3      1 0.33 68.40297       1
## 3      5    320       25     3      1 0.33 59.42551       1
## 4      0    330       25     3      1 0.50 93.70491       1
```

```
Clusthealthy[Clusthealthy$Model_1==2,]
```

```
##                         name mfr type calories protein fat sodium fiber
carbo
## 2          100%_Natural_Bran   Q    C      120       3   5     15     2
8
## 27       Frosted_Mini-Wheats   K    C      100       3   0      0     3
14
## 44                     Maypo   A    H      100       4   1      0     0
16
## 55               Puffed_Rice   Q    C       50       1   0      0     0
```

```
13
## 56              Puffed_Wheat   Q    C       50       2   0      0     1
10
## 61            Raisin_Squares   K    C       90       2   0      0     2
15
## 64            Shredded_Wheat   N    C       80       2   0      0     3
16
## 65    Shredded_Wheat_'n'Bran   N    C       90       3   0      0     4
19
## 66 Shredded_Wheat_spoon_size   N    C       90       3   0      0     3
20
## 69   Strawberry_Fruit_Wheats   N    C       90       2   0     15     3
15
##    sugars potass vitamins shelf weight cups   rating Model_1
## 2       8    135        0     3   1.00 1.00 33.98368       2
## 27      7    100       25     2   1.00 0.80 58.34514       2
## 44      3     95       25     2   1.00 1.00 54.85092       2
## 55      0     15        0     3   0.50 1.00 60.75611       2
## 56      0     50        0     3   0.50 1.00 63.00565       2
## 61      6    110       25     3   1.00 0.50 55.33314       2
## 64      0     95        0     1   0.83 1.00 68.23588       2
## 65      0    140        0     1   1.00 0.67 74.47295       2
## 66      0    120        0     1   1.00 0.67 72.80179       2
## 69      5     90       25     2   1.00 1.00 59.36399       2

Clusthealthy[Clusthealthy$Model_1==3,]

##                          name mfr type calories protein fat sodium fiber
carbo
## 6      Apple_Cinnamon_Cheerios   G    C      110       2   2    180   1.5
10.5
## 8                      Basic_4   G    C      130       3   2    210   2.0
18.0
## 9                    Bran_Chex   R    C       90       2   1    200   4.0
15.0
## 11               Cap'n'Crunch   Q    C      120       1   2    220   0.0
12.0
## 13       Cinnamon_Toast_Crunch   G    C      120       1   3    210   0.0
13.0
## 14                    Clusters   G    C      110       3   2    140   2.0
13.0
## 15                 Cocoa_Puffs   G    C      110       1   1    180   0.0
12.0
## 19               Count_Chocula   G    C      110       1   1    180   0.0
12.0
## 23      Crispy_Wheat_&_Raisins   G    C      100       2   1    140   2.0
11.0
## 24                 Double_Chex   R    C      100       2   0    190   1.0
18.0
## 26              Frosted_Flakes   K    C      110       1   0    200   1.0
```

```
14.0
## 33           Grape_Nuts_Flakes   P    C      100       3   1    140   3.0
15.0
## 34                 Grape-Nuts   P    C      110       3   0    170   3.0
17.0
## 36            Honey_Graham_Ohs   Q    C      120       1   2    220   1.0
12.0
## 37          Honey_Nut_Cheerios   G    C      110       3   1    250   1.5
11.5
## 38                 Honey-comb   P    C      110       1   0    180   0.0
14.0
## 39 Just_Right_Crunchy__Nuggets   K    C      110       2   1    170   1.0
17.0
## 40      Just_Right_Fruit_&_Nut   K    C      140       3   1    170   2.0
20.0
## 42                        Life   Q    C      100       4   2    150   2.0
12.0
## 43               Lucky_Charms   G    C      110       2   1    180   0.0
12.0
## 48        Multi-Grain_Cheerios   G    C      100       2   1    220   2.0
15.0
## 49            Nut&Honey_Crunch   K    C      120       2   1    190   0.0
15.0
## 50   Nutri-Grain_Almond-Raisin   K    C      140       3   2    220   3.0
21.0
## 51           Nutri-grain_Wheat   K    C       90       3   0    170   3.0
18.0
## 52        Oatmeal_Raisin_Crisp   G    C      130       3   2    170   1.5
13.5
## 57          Quaker_Oat_Squares   Q    C      100       4   1    135   2.0
14.0
## 70           Total_Corn_Flakes   G    C      110       2   1    200   0.0
21.0
## 72           Total_Whole_Grain   G    C      100       3   1    200   3.0
16.0
## 75                 Wheat_Chex   R    C      100       3   1    230   3.0
17.0
## 76                   Wheaties   G    C      100       3   1    200   3.0
17.0
## 77         Wheaties_Honey_Gold   G    C      110       2   1    200   1.0
16.0
##    sugars potass vitamins shelf weight cups   rating Model_1
## 6      10     70       25     1   1.00 0.75 29.50954       3
## 8       8    100       25     3   1.33 0.75 37.03856       3
## 9       6    125       25     1   1.00 0.67 49.12025       3
## 11     12     35       25     2   1.00 0.75 18.04285       3
## 13      9     45       25     2   1.00 0.75 19.82357       3
## 14      7    105       25     3   1.00 0.50 40.40021       3
## 15     13     55       25     2   1.00 1.00 22.73645       3
## 19     13     65       25     2   1.00 1.00 22.39651       3
```

```
## 23     10    120       25     3   1.00 0.75 36.17620       3
## 24      5     80       25     3   1.00 0.75 44.33086       3
## 26     11     25       25     1   1.00 0.75 31.43597       3
## 33      5     85       25     3   1.00 0.88 52.07690       3
## 34      3     90       25     3   1.00 0.25 53.37101       3
## 36     11     45       25     2   1.00 1.00 21.87129       3
## 37     10     90       25     1   1.00 0.75 31.07222       3
## 38     11     35       25     1   1.00 1.33 28.74241       3
## 39      6     60      100     3   1.00 1.00 36.52368       3
## 40      9     95      100     3   1.30 0.75 36.47151       3
## 42      6     95       25     2   1.00 0.67 45.32807       3
## 43     12     55       25     2   1.00 1.00 26.73451       3
## 48      6     90       25     1   1.00 1.00 40.10596       3
## 49      9     40       25     2   1.00 0.67 29.92429       3
## 50      7    130       25     3   1.33 0.67 40.69232       3
## 51      2     90       25     3   1.00 1.00 59.64284       3
## 52     10    120       25     3   1.25 0.50 30.45084       3
## 57      6    110       25     3   1.00 0.50 49.51187       3
## 70      3     35      100     3   1.00 1.00 38.83975       3
## 72      3    110      100     3   1.00 1.00 46.65884       3
## 75      3    115       25     1   1.00 0.67 49.78744       3
## 76      3    110       25     1   1.00 1.00 51.59219       3
## 77      8     60       25     1   1.00 0.75 36.18756       3

Clusthealthy[Clusthealthy$Model_1==4,]

##                 name mfr type calories protein fat sodium fiber carbo
sugars
## 7         Apple_Jacks   K    C      110       2   0    125     1    11
14
## 18          Corn_Pops   K    C      110       1   0     90     1    13
12
## 25        Froot_Loops   K    C      110       2   1    125     1    11
13
## 30     Fruity_Pebbles   P    C      110       1   1    135     0    13
12
## 31       Golden_Crisp   P    C      100       2   0     45     0    11
15
## 35 Great_Grains_Pecan   P    C      120       3   3     75     3    13
4
## 67             Smacks   K    C      110       2   1     70     1     9
15
## 74               Trix   G    C      110       1   1    140     0    13
12
##    potass vitamins shelf weight cups   rating Model_1
## 7      30       25     2      1 1.00 33.17409       4
## 18     20       25     2      1 1.00 35.78279       4
## 25     30       25     2      1 1.00 32.20758       4
## 30     25       25     2      1 0.75 28.02576       4
## 31     40       25     1      1 0.88 35.25244       4
```

```
## 35    100       25     3      1 0.33 45.81172       4
## 67     40       25     2      1 0.75 31.23005       4
## 74     25       25     2      1 1.00 27.75330       4
```

```
mean(Clusthealthy[Clusthealthy$Model_1==1,"rating"])
```

```
## [1] 73.84446
```

```
mean(Clusthealthy[Clusthealthy$Model_1==2,"rating"])
```

```
## [1] 60.11492
```

```
mean(Clusthealthy[Clusthealthy$Model_1==3,"rating"])
```

```
## [1] 37.30956
```

```
mean(Clusthealthy[Clusthealthy$Model_1==4,"rating"])
```

```
## [1] 33.65472
```

hence we see that mean value of cluster 1 is high (73.84446), we can choose cluster1