

Assignment #5 ML

#Packages used

```
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(cluster)
library(factoextra)

## Loading required package: ggplot2

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

#install.packages(dendextend)
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
```

#Reading the dataset

```
Myfile <- read.csv("Cereals.csv")
head(Myfile)
```

```
##              name mfr type calories protein fat sodium fiber
carbo
## 1          100%_Bran  N   C        70         4  1   130  10.0
5.0
## 2      100%_Natural_Bran  Q   C       120         3  5    15   2.0
8.0
## 3           All-Bran  K   C        70         4  1   260   9.0
7.0
## 4 All-Bran_with_Extra_Fiber  K   C        50         4  0   140  14.0
8.0
## 5           Almond_Delight  R   C       110         2  2   200   1.0
14.0
## 6  Apple_Cinnamon_Cheerios  G   C       110         2  2   180   1.5
10.5
##  sugars potass vitamins shelf weight cups  rating
## 1      6    280        25    3      1 0.33 68.40297
## 2      8    135         0    3      1 1.00 33.98368
## 3      5    320        25    3      1 0.33 59.42551
## 4      0    330        25    3      1 0.50 93.70491
## 5      8     NA        25    3      1 0.75 34.38484
## 6     10     70        25    1      1 0.75 29.50954
```

```
summary(Myfile)
```

```
##      name              mfr              type              calories
## Length:77          Length:77          Length:77          Min.   : 50.0
## Class :character    Class :character    Class :character    1st Qu.:100.0
## Mode  :character    Mode  :character    Mode  :character    Median :110.0
##                                     Mean   :106.9
##                                     3rd Qu.:110.0
##                                     Max.   :160.0
##
##      protein          fat          sodium          fiber
## Min.   :1.000    Min.   :0.000    Min.   :  0.0    Min.   : 0.000
## 1st Qu.:2.000    1st Qu.:0.000    1st Qu.:130.0    1st Qu.: 1.000
## Median :3.000    Median :1.000    Median :180.0    Median : 2.000
## Mean   :2.545    Mean   :1.013    Mean   :159.7    Mean   : 2.152
## 3rd Qu.:3.000    3rd Qu.:2.000    3rd Qu.:210.0    3rd Qu.: 3.000
## Max.   :6.000    Max.   :5.000    Max.   :320.0    Max.   :14.000
##
##      carbo          sugars          potass          vitamins
## Min.   : 5.0    Min.   : 0.000    Min.   : 15.00    Min.   : 0.00
```

```
## 1st Qu.:12.0 1st Qu.: 3.000 1st Qu.: 42.50 1st Qu.: 25.00
## Median :14.5 Median : 7.000 Median : 90.00 Median : 25.00
## Mean :14.8 Mean : 7.026 Mean : 98.67 Mean : 28.25
## 3rd Qu.:17.0 3rd Qu.:11.000 3rd Qu.:120.00 3rd Qu.: 25.00
## Max. :23.0 Max. :15.000 Max. :330.00 Max. :100.00
## NA's :1 NA's :1 NA's :2
## shelf weight cups rating
## Min. :1.000 Min. :0.50 Min. :0.250 Min. :18.04
## 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:0.670 1st Qu.:33.17
## Median :2.000 Median :1.00 Median :0.750 Median :40.40
## Mean :2.208 Mean :1.03 Mean :0.821 Mean :42.67
## 3rd Qu.:3.000 3rd Qu.:1.00 3rd Qu.:1.000 3rd Qu.:50.83
## Max. :3.000 Max. :1.50 Max. :1.500 Max. :93.70
##
```

#Removing and scaling the data

```
#removing missing values
X <- na.omit(Myfile)
Y <- X[,c(-1,-2,-3)]
Myfile <- na.omit(Myfile)
```

```
#Scale the data
XY <- scale(Y)
```

##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.

```
library(dplyr)
library(cluster)
library(factoextra)
library(dendextend)

#dissimilarity matrix
diss <- dist(XY, method = "euclidean")
#complete linkage
complete <- hclust(diss, method= "complete")
#average linkage
average <- hclust(diss, method= "average")

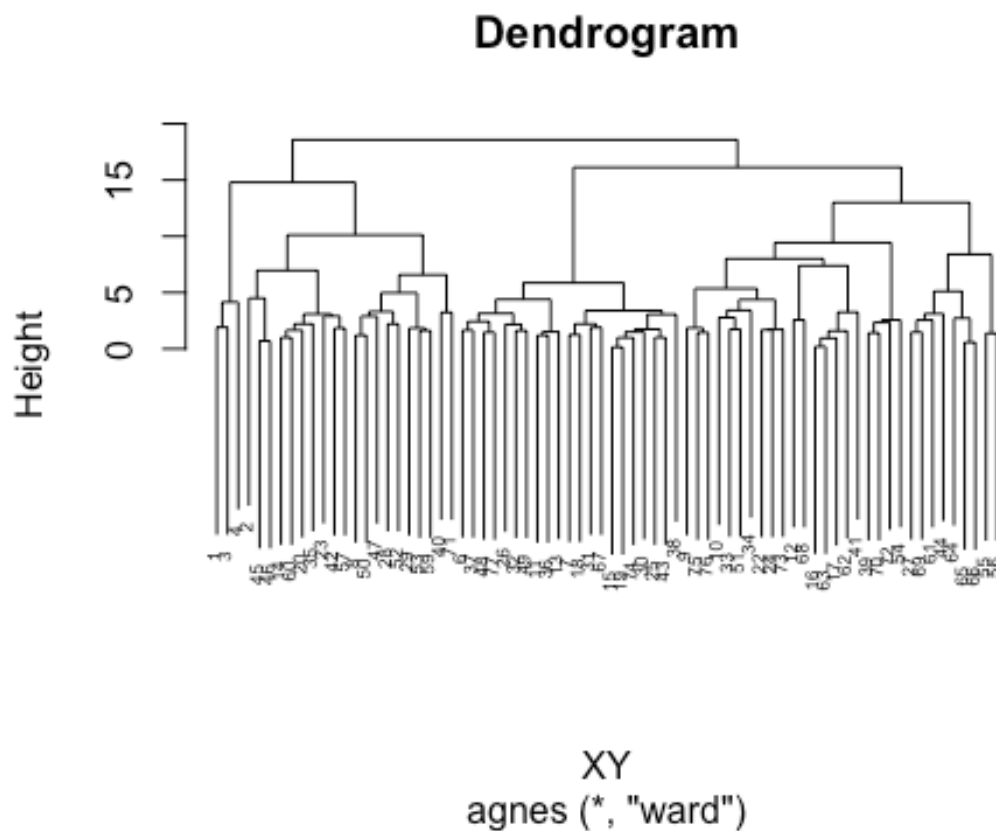
#using Agnes to compare the clustering
ag_single <- agnes(XY, method = "single")
ag_complete <- agnes(XY, method = "complete")
ag_average <- agnes(XY, method= "average")
ag_ward <- agnes(XY, method = "ward")

#comparing the coefficients
ag_single$ac
```

```
## [1] 0.6067859
ag_complete$ac
## [1] 0.8353712
ag_average$ac
## [1] 0.7766075
ag_ward$ac
## [1] 0.9046042

##the best method is Ward because the coefficient is the highest at .9046042

#plotting the dendrogram
plot1 <- pltree(ag_ward, cex= .5, hang=1, main="Dendrogram")
```

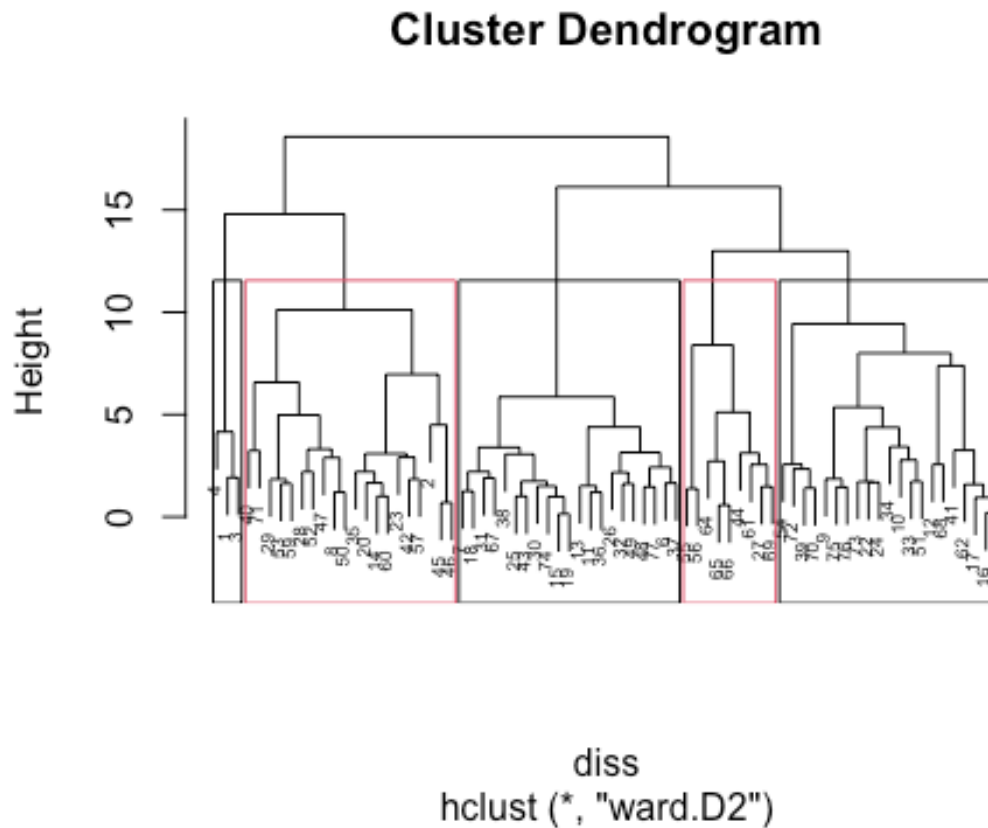


##How many clusters would you choose?

```
matrix <- dist(XY, method = "euclidean")
wardclust <- hclust(diss, method = "ward.D2")

#plotting dendrogram
```

```
plot(wardclust, cex = .5)
rect.hclust(wardclust, k=5, border= 1:2)
```



#Using cutree to identify the clusters

```
clust <- cutree(wardclust, k=5)
table(clust)
```

```
## clust
##  1  2  3  4  5
##  3 20 21 21  9
```

#I would choose 5 clusters. This is because after observing the height of the dendrogram, it appears that 5 clusters will work the best.

##Comment on the structure of the clusters and on their stability. Hint: To check stability, partition the data and see how well clusters formed based on one part apply to the other part. To do this: Cluster partition A

```
set.seed(100)
new <- Myfile
```

#removing missing values

```
cc <- na.omit(new)
cc1 <- cc[,c(-1,-2,-3)]
```

```

cc2 <- scale(cc1)
cc3 <- as.data.frame(cc2)

#partition
n1 <- cc[1:55,]
n2 <- cc[56:74,]

#cluster with agnes
clus1 <- agnes(scale(n1[, -c(1:3)]), method = "ward")
clus2 <- agnes(scale(n1[, -c(1:3)]), method = "average")
clus3 <- agnes(scale(n1[, -c(1:3)]), method = "complete")
clus4 <- agnes(scale(n1[, -c(1:3)]), method = "single")
cbind(ward = clus1$ac, average = clus2$ac, complete = clus3$ac,
single=clus4$ac)

##           ward   average  complete    single
## [1,] 0.8808195 0.7449303 0.8120228 0.6564842

tree <- cutree(clus1, k=5)

#The clusters are stable.

```

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

```

#the centers
x <- as.data.frame(cbind(scale(n1[, -c(1:3)]), tree))
center <- colMeans(x[x$tree==1,])
center2 <- colMeans(x[x$tree==2,])
center3 <- colMeans(x[x$tree==3,])
center4 <- colMeans(x[x$tree==4,])
center5 <- colMeans(x[x$tree==5,])

#binding the centers together
cent <- rbind(center, center2, center3, center4, center5)
cent

##           calories  protein    fat    sodium    fiber    carbo
## center -2.0643907  1.3722190 -0.4227336 0.1207793 3.2987660 -1.96959911
## center2 0.7279288  0.6244015  0.9156909 -0.2370392 0.3023157 -0.09643751
## center3 0.0924354 -0.9045372 -0.1168080 0.1575818 -0.5604394 -0.35725011
## center4 -0.1232472 0.3291094 -0.6062889 0.3218132 -0.1857336 1.04074035
## center5 -2.6806267 -0.9287580 -1.0345847 -2.1954810 -0.6494980 -0.68577787
##           sugars    potass    vitamins    shelf    weight    cups
## center -0.9815196 2.8376723 -0.1479841 0.8445408 -0.18608826 -1.7018175
## center2 0.3192895 0.6218875 -0.2327666 0.7638988 0.72999439 -0.5356524
## center3 0.7942560 -0.6498707 -0.1479841 -0.9210927 -0.18608826 0.2889243
## center4 -0.9002190 -0.2882521 0.6659284 0.0703784 -0.05973204 0.4454601
## center5 -1.8758258 -0.8959461 -1.5045049 0.8445408 -3.34499397 0.7522140

```

```
##           rating tree
## center    2.4007287    1
## center2  -0.1391348    2
## center3  -0.8182967    3
## center4   0.5003692    4
## center5   1.5330352    5
```

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

```
#calculating distance
y<- as.data.frame(rbind(cent[, -14], scale(n2[, -c(1:3)])))
y1 <- get_dist(y)
y2 <- as.matrix(y1)
d <- data.frame(data=seq(1,nrow(n2),1), clusters=rep(0,nrow(n2)))
for( i in 1:nrow(n2))
{d[i,2] <- which.min(y2[i+5,1:5])}
```

d

```
##      data clusters
## 1      1         2
## 2      2         2
## 3      3         2
## 4      4         4
## 5      5         4
## 6      6         5
## 7      7         4
## 8      8         4
## 9      9         3
## 10     10        4
## 11     11        4
## 12     12        4
## 13     13        2
## 14     14        4
## 15     15        4
## 16     16        3
## 17     17        2
## 18     18        4
## 19     19        3
```

```
y3 <- as.data.frame(cbind(XY,clust))
cbind(y3$clust[56:74], d$clusters)
```

```
##      [,1] [,2]
## [1,]    2    2
## [2,]    2    2
## [3,]    5    2
## [4,]    4    4
## [5,]    4    4
## [6,]    5    5
```

```
## [7,]      5      4
## [8,]      5      4
## [9,]      3      3
## [10,]     4      4
## [11,]     5      4
## [12,]     4      4
## [13,]     2      2
## [14,]     4      4
## [15,]     4      4
## [16,]     3      3
## [17,]     4      2
## [18,]     4      4
## [19,]     3      3
```

```
table(y3$clust[56:74]==d$clusters)
```

```
##
## FALSE  TRUE
##      5    14
```

#Assesing the cluster assignments, the data is stable.

##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?

```
cereal <- cbind(cc3, clust)
cereal[cereal$clust==1,]
```

```
##      calories  protein      fat      sodium  fiber      carbo      sugars
## 1 -1.865915  1.381748  0.0000000 -0.3910227  3.228667 -2.500140 -0.2542051
## 3 -1.865915  1.381748  0.0000000  1.1795987  2.816023 -1.986222 -0.4836096
## 4 -2.873782  1.381748 -0.9932203 -0.2702057  4.879247 -1.729263 -1.6306324
##      potass  vitamins  shelf      weight      cups      rating clust
## 1  2.560523 -0.1818422  0.9419715 -0.2008324 -2.085658  1.854904      1
## 3  3.124867 -0.1818422  0.9419715 -0.2008324 -2.085658  1.215196      1
## 4  3.265954 -0.1818422  0.9419715 -0.2008324 -1.364449  3.657844      1
```

```
cereal[cereal$clust==2,]
```

```
##      calories  protein      fat      sodium  fiber      carbo
## 2  0.6537514  0.4522084  3.9728810 -1.78041856 -0.07249167 -1.72926320
## 8  1.1576848  0.4522084  0.9932203  0.57551356 -0.07249167  0.84032469
## 14 0.1498180  0.4522084  0.9932203 -0.27020566 -0.07249167 -0.44446926
## 20 0.1498180  0.4522084  1.9864405 -0.27020566  0.75279812 -1.21534562
## 23 -0.3541153 -0.4773310  0.0000000 -0.27020566 -0.07249167 -0.95838683
## 28 0.6537514  0.4522084  0.9932203 -0.02857160  1.16544301 -0.70142805
## 29 0.6537514  0.4522084 -0.9932203  0.93796466  1.16544301 -0.18751047
## 35 0.6537514  0.4522084  1.9864405 -1.05551637  0.34015322 -0.44446926
## 40 1.6616182  0.4522084  0.0000000  0.09224544 -0.07249167  1.35424227
```



```

## 42 -0.3541153 1.3817478 0.9932203 -0.14938863 -0.07249167 -0.70142805
## 45 2.1655516 1.3817478 1.9864405 -0.81388230 0.34015322 0.32640711
## 46 2.1655516 1.3817478 1.9864405 -0.14938863 0.34015322 0.32640711
## 47 2.6694849 0.4522084 0.9932203 -0.14938863 0.34015322 0.58336590
## 50 1.6616182 0.4522084 0.9932203 0.69633060 0.34015322 1.61120105
## 52 1.1576848 0.4522084 0.9932203 0.09224544 -0.27881412 -0.31598986
## 53 0.6537514 0.4522084 0.0000000 0.45469653 1.57808790 -0.95838683
## 57 -0.3541153 1.3817478 0.0000000 -0.33061417 -0.07249167 -0.18751047
## 59 0.6537514 0.4522084 0.0000000 0.57551356 1.16544301 -0.18751047
## 60 -0.3541153 0.4522084 0.9932203 -0.27020566 0.13383078 -1.08686623
## 71 1.6616182 0.4522084 0.0000000 0.33387950 0.75279812 0.06944832
##          sugars      potass  vitamins      shelf      weight      cups
## 2  0.20460407 0.51477378 -1.3032024 0.9419715 -0.2008324 0.7567534
## 8  0.20460407 0.02097226 -0.1818422 0.9419715 1.9501886 -0.3038480
## 14 -0.02480049 0.09151534 -0.1818422 0.9419715 -0.2008324 -1.3644493
## 20 -0.02480049 0.86748914 -0.1818422 0.9419715 -0.2008324 -1.3644493
## 23 0.66341318 0.30314456 -0.1818422 0.9419715 -0.2008324 -0.3038480
## 28 0.66341318 1.43183372 -0.1818422 0.9419715 1.4287290 -0.6432404
## 29 1.12222230 1.29074758 -0.1818422 0.9419715 1.9501886 -0.6432404
## 35 -0.71301417 0.02097226 -0.1818422 0.9419715 -0.2008324 -2.0856582
## 40 0.43400862 -0.04957081 3.1822385 0.9419715 1.7546413 -0.3038480
## 42 -0.25420505 -0.04957081 -0.1818422 -0.2598542 -0.2008324 -0.6432404
## 45 0.89281774 1.00857529 -0.1818422 0.9419715 -0.2008324 0.7567534
## 46 0.89281774 1.00857529 -0.1818422 0.9419715 -0.2008324 0.7567534
## 47 1.35162686 0.86748914 -0.1818422 0.9419715 3.0582904 -0.6432404
## 50 -0.02480049 0.44423070 -0.1818422 0.9419715 1.9501886 -0.6432404
## 52 0.66341318 0.30314456 -0.1818422 0.9419715 1.4287290 -1.3644493
## 53 1.58103142 2.27835060 -0.1818422 0.9419715 1.9501886 -0.6432404
## 57 -0.25420505 0.16205841 -0.1818422 0.9419715 -0.2008324 -1.3644493
## 59 1.12222230 1.99617831 -0.1818422 -0.2598542 1.9501886 -0.3038480
## 60 0.20460407 0.58531685 -0.1818422 0.9419715 -0.2008324 -1.3644493
## 71 1.58103142 1.85509216 3.1822385 0.9419715 3.0582904 0.7567534
##          rating clust
## 2 -0.59771126      2
## 8 -0.38002951      2
## 14 -0.14048876      2
## 20 -0.13702824      2
## 23 -0.44147911      2
## 28 -0.10366038      2
## 29 -0.09664548      2
## 35 0.24511896      2
## 40 -0.42043579      2
## 42 0.21065609      2
## 45 -0.37302488      2
## 46 -0.58658904      2
## 47 -0.85924775      2
## 50 -0.11967375      2
## 52 -0.84945049      2
## 53 -0.32287913      2
## 57 0.50878106      2

```

```
## 59 -0.22179377      2
## 60 -0.19014120      2
## 71 -0.98185009      2
```

```
cereal[cereal$clust==3,]
```

##	calories	protein	fat	sodium	fiber	carbo
## 6	0.1498180	-0.4773310	0.9932203	0.2130625	-0.27881412	-1.08686623
## 7	0.1498180	-0.4773310	-0.9932203	-0.4514312	-0.48513656	-0.95838683
## 11	0.6537514	-1.4068705	0.9932203	0.6963306	-0.89778146	-0.70142805
## 13	0.6537514	-1.4068705	1.9864405	0.5755136	-0.89778146	-0.44446926
## 15	0.1498180	-1.4068705	0.0000000	0.2130625	-0.89778146	-0.70142805
## 18	0.1498180	-1.4068705	-0.9932203	-0.8742908	-0.48513656	-0.44446926
## 19	0.1498180	-1.4068705	0.0000000	0.2130625	-0.89778146	-0.70142805
## 25	0.1498180	-0.4773310	0.0000000	-0.4514312	-0.48513656	-0.95838683
## 26	0.1498180	-1.4068705	-0.9932203	0.4546965	-0.48513656	-0.18751047
## 30	0.1498180	-1.4068705	0.0000000	-0.3306142	-0.89778146	-0.44446926
## 31	-0.3541153	-0.4773310	-0.9932203	-1.4179675	-0.89778146	-0.95838683
## 32	0.1498180	-1.4068705	0.0000000	1.4212328	-0.89778146	0.06944832
## 36	0.6537514	-1.4068705	0.9932203	0.6963306	-0.48513656	-0.70142805
## 37	0.1498180	0.4522084	0.0000000	1.0587817	-0.27881412	-0.82990744
## 38	0.1498180	-1.4068705	-0.9932203	0.2130625	-0.89778146	-0.18751047
## 43	0.1498180	-0.4773310	0.0000000	0.2130625	-0.89778146	-0.70142805
## 48	-0.3541153	-0.4773310	0.0000000	0.6963306	-0.07249167	0.06944832
## 49	0.6537514	-0.4773310	0.0000000	0.3338795	-0.89778146	0.06944832
## 67	0.1498180	-0.4773310	0.0000000	-1.1159249	-0.48513656	-1.47230441
## 74	0.1498180	-1.4068705	0.0000000	-0.2702057	-0.89778146	-0.44446926
## 77	0.1498180	-0.4773310	0.0000000	0.4546965	-0.48513656	0.32640711
##	sugars	potass	vitamins	shelf	weight	cups
rating						
## 6	0.6634132	-0.4022862	-0.1818422	-1.4616799	-0.2008324	-0.3038480 - 0.9165248
## 7	1.5810314	-0.9666308	-0.1818422	-0.2598542	-0.2008324	0.7567534 - 0.6553998
## 11	1.1222223	-0.8960877	-0.1818422	-0.2598542	-0.2008324	-0.3038480 - 1.7336066
## 13	0.4340086	-0.7550015	-0.1818422	-0.2598542	-0.2008324	-0.3038480 - 1.6067177
## 15	1.3516269	-0.6139154	-0.1818422	-0.2598542	-0.2008324	0.7567534 - 1.3991551
## 18	1.1222223	-1.1077169	-0.1818422	-0.2598542	-0.2008324	0.7567534 - 0.4695120
## 19	1.3516269	-0.4728292	-0.1818422	-0.2598542	-0.2008324	0.7567534 - 1.4233777
## 25	1.3516269	-0.9666308	-0.1818422	-0.2598542	-0.2008324	0.7567534 - 0.7242706
## 26	0.8928177	-1.0371738	-0.1818422	-1.4616799	-0.2008324	-0.3038480 - 0.7792531
## 30	1.1222223	-1.0371738	-0.1818422	-0.2598542	-0.2008324	-0.3038480 - 1.0222542

```

## 31  1.8104360 -0.8255446 -0.1818422 -1.4616799 -0.2008324  0.2476647 -
0.5073029
## 32  0.4340086 -0.7550015 -0.1818422 -0.2598542 -0.2008324 -0.3038480 -
1.3230814
## 36  0.8928177 -0.7550015 -0.1818422 -0.2598542 -0.2008324  0.7567534 -
1.4608034
## 37  0.6634132 -0.1201139 -0.1818422 -1.4616799 -0.2008324 -0.3038480 -
0.8051733
## 38  0.8928177 -0.8960877 -0.1818422 -1.4616799 -0.2008324  2.1567472 -
0.9711880
## 43  1.1222223 -0.6139154 -0.1818422 -0.2598542 -0.2008324  0.7567534 -
1.1142648
## 48 -0.2542051 -0.1201139 -0.1818422 -1.4616799 -0.2008324  0.7567534 -
0.1614556
## 49  0.4340086 -0.8255446 -0.1818422 -0.2598542 -0.2008324 -0.6432404 -
0.8869714
## 67  1.8104360 -0.8255446 -0.1818422 -0.2598542 -0.2008324 -0.3038480 -
0.7939263
## 74  1.1222223 -1.0371738 -0.1818422 -0.2598542 -0.2008324  0.7567534 -
1.0416692
## 77  0.2046041 -0.5433723 -0.1818422 -1.4616799 -0.2008324 -0.3038480 -
0.4406694
##      clust
## 6        3
## 7        3
## 11       3
## 13       3
## 15       3
## 18       3
## 19       3
## 25       3
## 26       3
## 30       3
## 31       3
## 32       3
## 36       3
## 37       3
## 38       3
## 43       3
## 48       3
## 49       3
## 67       3
## 74       3
## 77       3

cereal[cereal$clust==4,]

##      calories  protein      fat      sodium      fiber      carbo
## 9  -0.8580487 -0.4773310  0.0000000  0.45469653  0.75279812  0.06944832
## 10 -0.8580487  0.4522084 -0.9932203  0.57551356  1.16544301 -0.44446926

```

```

## 12 0.1498180 3.2408266 0.9932203 1.54204982 -0.07249167 0.58336590
## 16 0.1498180 -0.4773310 -0.9932203 1.42123279 -0.89778146 1.86815984
## 17 -0.3541153 -0.4773310 -0.9932203 1.54204982 -0.48513656 1.61120105
## 22 0.1498180 -0.4773310 -0.9932203 0.69633060 -0.48513656 1.61120105
## 24 -0.3541153 -0.4773310 -0.9932203 0.33387950 -0.48513656 0.84032469
## 33 -0.3541153 0.4522084 0.0000000 -0.27020566 0.34015322 0.06944832
## 34 0.1498180 0.4522084 -0.9932203 0.09224544 0.34015322 0.58336590
## 39 0.1498180 -0.4773310 0.0000000 0.09224544 -0.48513656 0.58336590
## 41 0.1498180 -0.4773310 0.0000000 1.17959872 -0.89778146 1.61120105
## 51 -0.8580487 0.4522084 -0.9932203 0.09224544 0.34015322 0.84032469
## 54 -0.3541153 0.4522084 -0.9932203 1.90450091 -0.48513656 1.35424227
## 62 0.1498180 -1.4068705 -0.9932203 0.93796466 -0.89778146 2.12511863
## 63 0.1498180 -0.4773310 -0.9932203 1.54204982 -0.89778146 1.86815984
## 68 0.1498180 3.2408266 -0.9932203 0.81714763 -0.48513656 0.32640711
## 70 0.1498180 -0.4773310 0.0000000 0.45469653 -0.89778146 1.61120105
## 72 -0.3541153 0.4522084 0.0000000 0.45469653 0.34015322 0.32640711
## 73 0.1498180 -0.4773310 0.0000000 1.05878169 -0.89778146 1.61120105
## 75 -0.3541153 0.4522084 0.0000000 0.81714763 0.34015322 0.58336590
## 76 -0.3541153 0.4522084 0.0000000 0.45469653 0.34015322 0.58336590
##      sugars      potass    vitamins      shelf      weight      cups
## 9  -0.2542051  0.37368763 -0.1818422 -1.4616799 -0.2008324 -0.6432404
## 10 -0.4836096  1.29074758 -0.1818422  0.9419715 -0.2008324 -0.6432404
## 12 -1.4012278  0.09151534 -0.1818422 -1.4616799 -0.2008324  1.8173547
## 16 -0.9424187 -1.03717383 -0.1818422 -1.4616799 -0.2008324  0.7567534
## 17 -1.1718233 -0.89608768 -0.1818422 -1.4616799 -0.2008324  0.7567534
## 22 -0.9424187 -0.96663076 -0.1818422  0.9419715 -0.2008324  0.7567534
## 24 -0.4836096 -0.26120003 -0.1818422  0.9419715 -0.2008324 -0.3038480
## 33 -0.4836096 -0.19065695 -0.1818422  0.9419715 -0.2008324  0.2476647
## 34 -0.9424187 -0.12011388 -0.1818422  0.9419715 -0.2008324 -2.4250507
## 39 -0.2542051 -0.54337232  3.1822385  0.9419715 -0.2008324  0.7567534
## 41 -0.9424187 -0.82554461 -0.1818422 -0.2598542 -0.2008324  2.8779561
## 51 -1.1718233 -0.12011388 -0.1818422  0.9419715 -0.2008324  0.7567534
## 54 -0.9424187 -0.75500154  3.1822385  0.9419715 -0.2008324  0.7567534
## 62 -1.1718233 -0.96663076 -0.1818422 -1.4616799 -0.2008324  1.3082661
## 63 -0.9424187 -0.89608768 -0.1818422 -1.4616799 -0.2008324  0.7567534
## 68 -0.9424187 -0.61391539 -0.1818422 -1.4616799 -0.2008324  0.7567534
## 70 -0.9424187 -0.89608768  3.1822385  0.9419715 -0.2008324  0.7567534
## 72 -0.9424187  0.16205841  3.1822385  0.9419715 -0.2008324  0.7567534
## 73 -0.9424187 -0.54337232 -0.1818422  0.9419715 -0.2008324 -0.3038480
## 75 -0.9424187  0.23260148 -0.1818422 -1.4616799 -0.2008324 -0.6432404
## 76 -0.9424187  0.16205841 -0.1818422 -1.4616799 -0.2008324  0.7567534
##      rating clust
## 9  0.48087533    4
## 10 0.77969576    4
## 12 0.59807496    4
## 16 -0.06603869    4
## 17 0.24879639    4
## 22 0.32235640    4
## 24 0.13959735    4
## 33 0.69155685    4

```

```
## 34 0.78377123      4
## 39 -0.41671824     4
## 41 -0.22308231     4
## 51 1.23068291      4
## 54 -0.06186866     4
## 62 -0.02656845     4
## 63 -0.12909114     4
## 68 0.76669214      4
## 70 -0.25168258     4
## 72 0.30548275      4
## 73 -0.23269772     4
## 75 0.52841741      4
## 76 0.65701831      4
```

```
cereal[cereal$clust==5,]
```

```
##      calories  protein      fat  sodium      fiber      carbo
## 27 -0.3541153  0.4522084 -0.9932203 -1.961644  0.34015322 -0.18751047
## 44 -0.3541153  1.3817478  0.0000000 -1.961644 -0.89778146  0.32640711
## 55 -2.8737823 -1.4068705 -0.9932203 -1.961644 -0.89778146 -0.44446926
## 56 -2.8737823 -0.4773310 -0.9932203 -1.961644 -0.48513656 -1.21534562
## 61 -0.8580487 -0.4773310 -0.9932203 -1.961644 -0.07249167  0.06944832
## 64 -1.3619821 -0.4773310 -0.9932203 -1.961644  0.34015322  0.32640711
## 65 -0.8580487  0.4522084 -0.9932203 -1.961644  0.75279812  1.09728348
## 66 -0.8580487  0.4522084 -0.9932203 -1.961644  0.34015322  1.35424227
## 69 -0.8580487 -0.4773310 -0.9932203 -1.780419  0.34015322  0.06944832
##      sugars      potass  vitamins      shelf      weight      cups
## 27 -0.02480049  0.02097226 -0.1818422 -0.2598542 -0.2008324 -0.09172768
## 44 -0.94241873 -0.04957081 -0.1818422 -0.2598542 -0.2008324  0.75675340
## 55 -1.63063240 -1.17825998 -1.3032024  0.9419715 -3.4599552  0.75675340
## 56 -1.63063240 -0.68445846 -1.3032024  0.9419715 -3.4599552  0.75675340
## 61 -0.25420505  0.16205841 -0.1818422  0.9419715 -0.2008324 -1.36444931
## 64 -1.63063240 -0.04957081 -1.3032024 -1.4616799 -1.3089342  0.75675340
## 65 -1.63063240  0.58531685 -1.3032024 -1.4616799 -0.2008324 -0.64324039
## 66 -1.63063240  0.30314456 -1.3032024 -1.4616799 -0.2008324 -0.64324039
## 69 -0.48360961 -0.12011388 -0.1818422 -0.2598542 -0.2008324  0.75675340
##      rating clust
## 27 1.1382130     5
## 44 0.8892251     5
## 55 1.3100115     5
## 56 1.4703065     5
## 61 0.9235870     5
## 64 1.8429976     5
## 65 2.2874319     5
## 66 2.1683500     5
## 69 1.2108133     5
```

#Using mean rating calculation to determine the best cluster of healthy cereals

```
mean(cereal[cereal$clust==1,"rating"])
```

```
## [1] 2.242648
mean(cereal[cereal$clust==2, "rating"])
## [1] -0.2928786
mean(cereal[cereal$clust==3, "rating"])
## [1] -0.9636465
mean(cereal[cereal$clust==4, "rating"])
## [1] 0.2916795
mean(cereal[cereal$clust==5, "rating"])
## [1] 1.471215
```

#After calculating the mean rating, cluster #1 is the healthier option. This was concluded because cluster #1 has a higher mean rating of 2.242648. Since cluster #1 has a higher mean rating of healthy attributes, the elementary public schools should choose cluster #1.