A17

Bryana Benson
Conner Bryan

A17

Libraries

```r
library(readxl)
library(readr)
#install.packages("NbClust")
library(NbClust)
library(tidyverse)
#install.packages("rattle")
library(rattle)
#install.packages("cluster")
library(cluster)
```

Reading the Data

```r
#url<-"http://im.ft-static.com/content/images/b38c350e-169d-11e5-b07f-00144feabdc0.xls"
#data<-read_xls(url)

data<-read_excel("FTGlobal500.xls")
```

Cleaning the Data

```r
colnames(data)<-data[4,]
data<-data[-(1:4),]

companies<-data$Company

data <- sapply(data, as.numeric)
data<-as.data.frame(data)
data<-data[, -c(3,4,6)]
data$Company<-companies

data<-na.omit(data)

top_20<- data %>%
  arrange(desc(`Market value $m`))

top_20<-top_20[1:20,]

companies<-top_20$Company

top_20<-top_20[,-12]

rownames(top_20)<-companies

head(top_20)
```

```
##                 Global rank 2015 Global rank 2014 Market value $m
## Apple                          1               1         724773.1
## Exxon Mobil                    2               2         356548.7
## Microsoft                      5               3         333524.8
## PetroChina                     6              16         329715.1
## Johnson & Johnson              8               6         279723.9
## Novartis                      10              14         267897.0
##                 Turnover $m Net income $m Total assets $m Employees
## Apple              182795.00      39510.00         231839.0     92600
## Exxon Mobil        364763.00      32520.00         349493.0     75300
## Microsoft           86833.00      22074.00         172384.0    128000
## PetroChina         367853.67      17268.62         385177.7    534652
## Johnson & Johnson   74331.00      16323.00         127723.0    126500
## Novartis            49550.66       9432.68         117393.4    133413
##                 Price $ P/e ratio Dividend yield (%) Year End
## Apple           124.430  19.29147          1.454633    41909
## Exxon Mobil      85.000  11.19895          3.176471    42004
## Microsoft        40.655  15.45817          2.754889    41820
## PetroChina        1.107  12.30000          3.613369    42004
## Johnson & Johnson 100.600 17.64912         2.743539    42001
## Novartis         98.994  25.44833          2.646625    42004
```

Creating Clusters

```
clusters<-kmeans(top_20, 3)

clusters$centers
```

```
##   Global rank 2015 Global rank 2014 Market value $m Turnover $m
## 1         15.71429         15.64286        270283.7    97053.67
## 2         12.00000         10.00000        265107.3   485651.00
## 3         12.40000         12.20000        273419.6   303158.95
##   Net income $m Total assets $m Employees  Price $ P/e ratio
## 1      15651.45        164640.8  140273.4 175.9915  18.67859
## 2      16363.00        203706.0 2200000.0  82.2500  16.28713
## 3      19257.55        425139.8  269565.4  42.1254  13.26887
##   Dividend yield (%) Year End
## 1          2.737447 41955.43
## 2          2.285714 42035.00
## 3          3.774412 41949.00
```
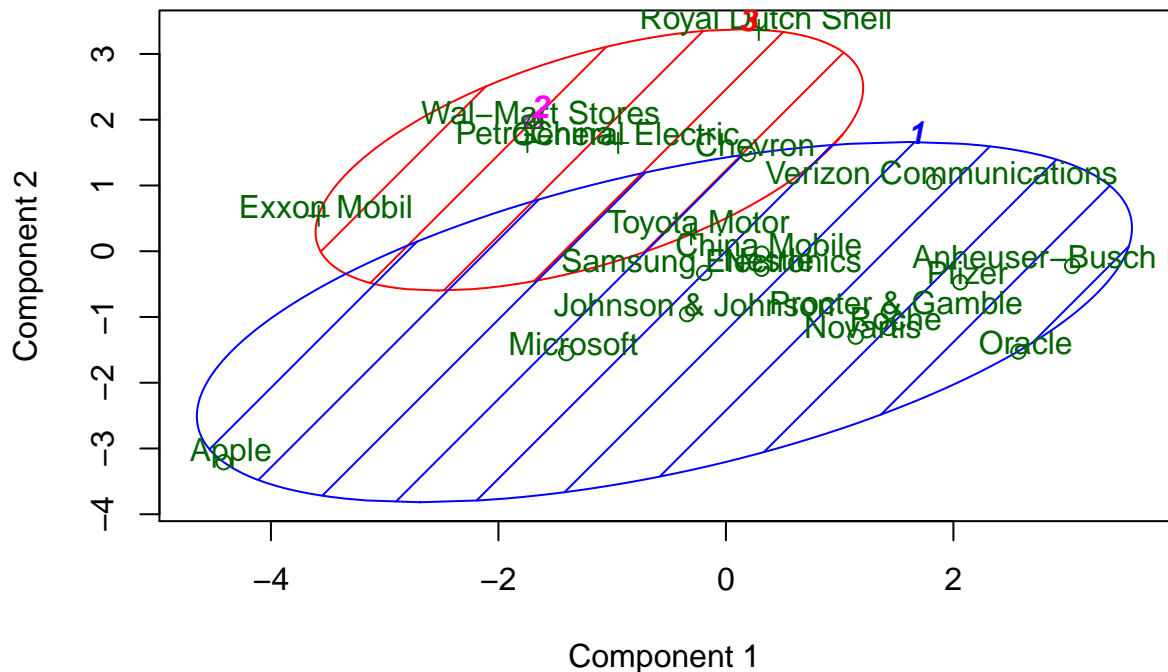
```
clusters$size
```

```
## [1] 14  1  5
```

Graphing two principal components

```
clusplot(top_20, clusters$cluster, main='2D representation of the Cluster solution',
         color=TRUE, shade=TRUE,
         labels=2, lines=0)
```
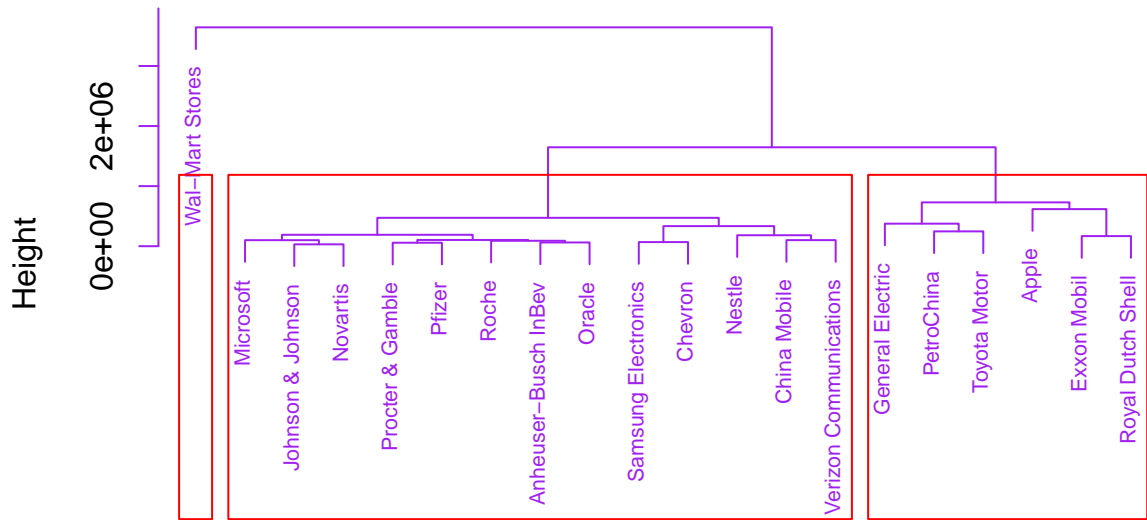
2D representation of the Cluster solution

Component 1
These two components explain 52.32 % of the point variability.

Plotting a dendrogram

```
d <- dist(top_20, method = "euclidean")
H.fit <- hclust(d, method="ward.D")
plot(H.fit, xlab="Company", sub="Clusters", cex=.7, col="purple")        # display dendogram
groups <- cutree(H.fit, k=3)              # cut tree into 5 clusters
rect.hclust(H.fit, k=3, border="red")
```

Cluster Dendrogram

Height

Company
Clusters