

## Code

```
##Clustering another example
getwd()
setwd('/Volumes/GoogleDrive/My Drive/ML_AI/Edureka/DS with R/Class 6')
getwd()
install.packages("readxl")
library(readxl)
Insurance_data <- read_excel("InsuranceData.xlsx")
summary(Insurance_data)
#CustomerData <- read.csv(file.choose())
Insurance_data <- data.frame(Insurance_data)
summary(Insurance_data)
sum(is.na(Insurance_data))
data <- Insurance_data[, -c(1,3,4)]
<.max <- 10
vss <- sapply(1:k.max, function(k){kmeans(data,k, nstart = 1)$tot.withinss})
plot(1:k.max,vss, type="b", frame=FALSE, xlab = "Number of clusters k", ylab = "total within clusters sum of squares")
set.seed(123)
install.packages("cluster")
library(cluster)
install.packages("fpc")
library(fpc)
22 library(fpc)
23 clus <- kmeans(data, centers=3)
24 # Fig 01
25 plotcluster(data, clus$cluster)
26 # More complex
27 clusplot(data, clus$cluster, color=TRUE, shade=TRUE,
28          labels=2, lines=0)
29 data_1 <- cbind(data, clus)
30 ##C-means clustering
31 library(e1071)
32 c2 <- cmeans(mycustomer, 4)
33 c2
34
35 ###W
36
```

## Results

Assignment -6\_soln.R

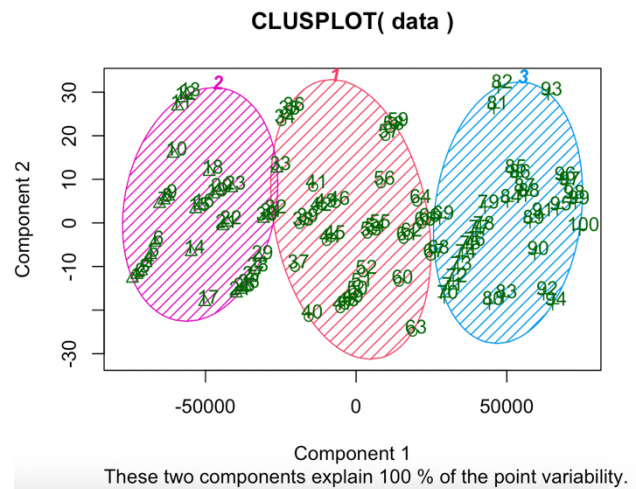
Insurance\_data

Clustering\_example\_7o3up.R

Filter

	Premiums.Paid	Age	Days.to.Renew	Claims.made	Income
1	2800	26	233	3890.076	28000
2	2950	27	130	2294.444	29500
3	3100	28	144	2564.545	31000
4	3250	30	65	1978.261	32500
5	3400	32	56	2009.091	34000
6	3550	35	89	2349.455	35500
7	3700	44	95	2503.346	37000
8	3850	45	48	2217.405	38500
9	4000	46	76	2527.778	40000
10	6225	56	200	6908.232	41500
11	6450	67	211	7672.549	43000
12	6675	69	245	10208.824	44500
13	6900	70	261	12192.233	46000
14	4750	34	278	10052.326	47500
15	4900	44	182	4900.000	49000
16	7575	45	60	4535.033	50500

Showing 1 to 16 of 100 entries, 5 total columns



Code