Fuel Cost Analysis in Power Generation

Tejesh Varma Maddana

#Loading all the required libraries to perform the analysis

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(caret)

## Loading required package: ggplot2

## Loading required package: lattice

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ lubridate 1.9.3 ✔ tibble 3.2.1  
## ✔ purrr 1.0.2 ✔ tidyr 1.3.0  
## ✔ readr 2.1.4

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ purrr::lift() masks caret::lift()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(class)  
library(e1071)  
library(gridExtra)

##   
## Attaching package: 'gridExtra'  
##   
## The following object is masked from 'package:dplyr':  
##   
## combine

library(pracma)

##   
## Attaching package: 'pracma'  
##   
## The following object is masked from 'package:e1071':  
##   
## sigmoid  
##   
## The following object is masked from 'package:purrr':  
##   
## cross

rm(list = ls()) #cleaning the environment  
library(readr)  
library(cluster)  
library(tidyr)  
library(ggplot2)  
library(pander)  
library(knitr)  
library(factoextra)

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

library(flexclust)

## Loading required package: grid  
## Loading required package: modeltools  
## Loading required package: stats4  
##   
## Attaching package: 'flexclust'  
##   
## The following object is masked from 'package:e1071':  
##   
## bclust

library(cowplot)

##   
## Attaching package: 'cowplot'  
##   
## The following object is masked from 'package:lubridate':  
##   
## stamp

**Loading the dataset to perform the K-Means Cluster analysis**

Energy\_data <- read.csv("~/Documents/KSU/Fundamentals of Machine Learning - 64060/Assignments/Final Exam/fuel\_receipts1.csv")  
head(Energy\_data)

## X rowid plant\_id\_eia energy\_source\_code fuel\_type\_code\_pudl fuel\_group\_code  
## 1 1 1 3 BIT coal coal  
## 2 2 2 3 BIT coal coal  
## 3 3 3 3 NG gas natural\_gas  
## 4 4 4 7 BIT coal coal  
## 5 5 5 7 BIT coal coal  
## 6 6 6 7 BIT coal coal  
## supplier\_name fuel\_received\_units fuel\_mmbtu\_per\_unit sulfur\_content\_pct  
## 1 interocean coal 259412 23.100 0.49  
## 2 interocean coal 52241 22.800 0.48  
## 3 bay gas pipeline 2783619 1.039 0.00  
## 4 alabama coal 25397 24.610 1.69  
## 5 d & e mining 764 24.446 0.84  
## 6 alabama coal 603 24.577 1.54  
## ash\_content\_pct fuel\_cost\_per\_mmbtu  
## 1 5.4 2.135  
## 2 5.7 2.115  
## 3 0.0 8.631  
## 4 14.7 2.776  
## 5 15.5 3.381  
## 6 14.6 2.199

**Using the summary fuction finding all the summary statistics**

summary(Energy\_data)

## X rowid plant\_id\_eia energy\_source\_code  
## Min. : 1.0 Min. : 1.0 Min. : 3.0 Length:756   
## 1st Qu.:189.8 1st Qu.: 221.8 1st Qu.: 527.0 Class :character   
## Median :378.5 Median : 450.5 Median : 728.0 Mode :character   
## Mean :378.5 Mean : 465.7 Mean : 808.7   
## 3rd Qu.:567.2 3rd Qu.: 704.2 3rd Qu.:1252.0   
## Max. :756.0 Max. :1000.0 Max. :1710.0   
## fuel\_type\_code\_pudl fuel\_group\_code supplier\_name fuel\_received\_units  
## Length:756 Length:756 Length:756 Min. : 1   
## Class :character Class :character Class :character 1st Qu.: 1300   
## Mode :character Mode :character Mode :character Median : 12570   
## Mean : 89258   
## 3rd Qu.: 42025   
## Max. :4823176   
## fuel\_mmbtu\_per\_unit sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## Min. : 0.857 Min. :0.0000 Min. : 0.000 Min. : 0.343   
## 1st Qu.: 1.028 1st Qu.:0.0000 1st Qu.: 0.000 1st Qu.: 1.947   
## Median :17.141 Median :0.3200 Median : 5.000 Median : 3.082   
## Mean :13.349 Mean :0.7314 Mean : 4.995 Mean : 5.786   
## 3rd Qu.:23.870 3rd Qu.:0.9400 3rd Qu.: 9.600 3rd Qu.: 8.400   
## Max. :29.400 Max. :6.6100 Max. :20.900 Max. :29.514

#Thus from the above results of summary function, each of the variable Min, 1st Qu, median, mean, 3rd Qu, Max values are resulted for analysis and to have a general idea about the values.

**Finding the structure of dataset**

str(Energy\_data)

## 'data.frame': 756 obs. of 12 variables:  
## $ X : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ rowid : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ plant\_id\_eia : int 3 3 3 7 7 7 7 8 8 8 ...  
## $ energy\_source\_code : chr "BIT" "BIT" "NG" "BIT" ...  
## $ fuel\_type\_code\_pudl: chr "coal" "coal" "gas" "coal" ...  
## $ fuel\_group\_code : chr "coal" "coal" "natural\_gas" "coal" ...  
## $ supplier\_name : chr "interocean coal" "interocean coal" "bay gas pipeline" "alabama coal" ...  
## $ fuel\_received\_units: int 259412 52241 2783619 25397 764 603 2341 8869 75442 206741 ...  
## $ fuel\_mmbtu\_per\_unit: num 23.1 22.8 1.04 24.61 24.45 ...  
## $ sulfur\_content\_pct : num 0.49 0.48 0 1.69 0.84 1.54 0 2.16 1.24 1.9 ...  
## $ ash\_content\_pct : num 5.4 5.7 0 14.7 15.5 14.6 0 15.4 11.9 15.4 ...  
## $ fuel\_cost\_per\_mmbtu: num 2.13 2.12 8.63 2.78 3.38 ...

#From the above results, i came to know that there are 756 observations and 12 variables .

**Checking the missing or not available values in the provided dataset by using the is.na() function**

colMeans(is.na(Energy\_data))

## X rowid plant\_id\_eia energy\_source\_code   
## 0 0 0 0   
## fuel\_type\_code\_pudl fuel\_group\_code supplier\_name fuel\_received\_units   
## 0 0 0 0   
## fuel\_mmbtu\_per\_unit sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu   
## 0 0 0 0

#Since, the results show “0” against the each variable, this indicate that there are no missing or not available values in the provided dataset and the data is clean which can be used for proceeding the analysis.

**Now, i would like to form the clusters using the variables of “Sulfur Content”, “Ash Content”, “Fuel Cost”**

**Creating a Sub-set which includes only the values of required variable such as “Sulfur Content”, “Ash Content”, “Fuel Cost” to perform the cluster analysis**

#Using the numerical variables from 1 to 9 to cluster the 21 firms.  
Energy\_data\_1 <- Energy\_data[,c(10:12)]  
head(Energy\_data\_1)

## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## 1 0.49 5.4 2.135  
## 2 0.48 5.7 2.115  
## 3 0.00 0.0 8.631  
## 4 1.69 14.7 2.776  
## 5 0.84 15.5 3.381  
## 6 1.54 14.6 2.199

#Finding the summary statistics of "Energy\_data\_1"  
summary(Energy\_data\_1)

## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## Min. :0.0000 Min. : 0.000 Min. : 0.343   
## 1st Qu.:0.0000 1st Qu.: 0.000 1st Qu.: 1.947   
## Median :0.3200 Median : 5.000 Median : 3.082   
## Mean :0.7314 Mean : 4.995 Mean : 5.786   
## 3rd Qu.:0.9400 3rd Qu.: 9.600 3rd Qu.: 8.400   
## Max. :6.6100 Max. :20.900 Max. :29.514

##Finding the Structure of "Energy\_data\_1"  
str(Energy\_data\_1)

## 'data.frame': 756 obs. of 3 variables:  
## $ sulfur\_content\_pct : num 0.49 0.48 0 1.69 0.84 1.54 0 2.16 1.24 1.9 ...  
## $ ash\_content\_pct : num 5.4 5.7 0 14.7 15.5 14.6 0 15.4 11.9 15.4 ...  
## $ fuel\_cost\_per\_mmbtu: num 2.13 2.12 8.63 2.78 3.38 ...

**Normalizing the data - Since the data in the variable has different units, normalization is being done so that the cluster analysis is not affected by the different units of the variable. Normalization is being done using the “scale” function**

Norm\_Energy\_data\_1 <- scale(Energy\_data\_1)  
pandoc.table(head(Norm\_Energy\_data\_1),style="grid", split.tables = Inf)

##   
##   
## +--------------------+-----------------+---------------------+  
## | sulfur\_content\_pct | ash\_content\_pct | fuel\_cost\_per\_mmbtu |  
## +====================+=================+=====================+  
## | -0.2306 | 0.07903 | -0.6846 |  
## +--------------------+-----------------+---------------------+  
## | -0.2402 | 0.1376 | -0.6884 |  
## +--------------------+-----------------+---------------------+  
## | -0.6988 | -0.9756 | 0.5335 |  
## +--------------------+-----------------+---------------------+  
## | 0.9159 | 1.895 | -0.5644 |  
## +--------------------+-----------------+---------------------+  
## | 0.1038 | 2.052 | -0.451 |  
## +--------------------+-----------------+---------------------+  
## | 0.7726 | 1.876 | -0.6726 |  
## +--------------------+-----------------+---------------------+

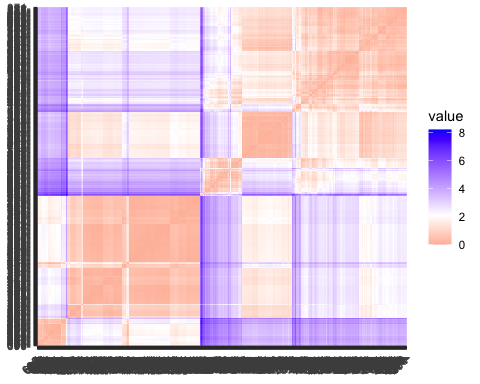
**The data is being clustered using euclidean distance method and plotting the graph to interpret the results by using the Euclidean distance formula**

#Finding the distances between observations in the data by using the above Euclidean distance formula  
Energy\_data\_1\_dist <- get\_dist(Norm\_Energy\_data\_1)   
corr <- cor(Norm\_Energy\_data\_1)  
corr

## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## sulfur\_content\_pct 1.0000000 0.5671048 -0.4364371  
## ash\_content\_pct 0.5671048 1.0000000 -0.6515051  
## fuel\_cost\_per\_mmbtu -0.4364371 -0.6515051 1.0000000

**Considering the distance matrix “Energy\_data\_1\_dist” as its main argument, visualizing the distances by using the ‘fviz\_dist()’ function which displays the heat-map.**

fviz\_dist(Energy\_data\_1\_dist , order = TRUE, show\_labels = TRUE)



**Interpretation of Heat-Map**

#The above heat-map indicates that the correlation between the three variables fuel cost, ash content, and sulfur content. The correlation is stronger the darker the colour.

#According to the heatmap,Ash content and sulphur content have a positive correlation. Accordingly, companies that use fuel with a higher sulphur level also typically use a higher ash content.

#Ash and sulphur concentration have a positive correlation with fuel cost. This indicates that businesses typically incur greater fuel expenses when their fuel contains higher levels of ash or sulphur.

#I can suggest that firms can reduce their fuel costs by using fuels with lower sulfur and ash content.

#There is a larger association between ash and sulphur content than there is between each of these two parameters and fuel price. This implies that factors like gasoline type or transportation expenses may not have as much of an impact on fuel cost as sulphur and ash concentration.

#The imperfect correlation between the three variables is also evident in the heatmap. This indicates that even among businesses with comparable levels of ash and sulphur, there is some variance in fuel costs. This implies that there might be more variables, such as the cost of fuel transportation or the effectiveness of fuel-burning machinery, that affect fuel prices as well.

#All things considered, the heatmap offers insightful information about the connection between fuel cost, ash content, and sulphur content. Businesses can utilise this data to create plans that will save fuel expenses while enhancing environmental performance.

**Using the ELbow Method Finding the best value of “K”**

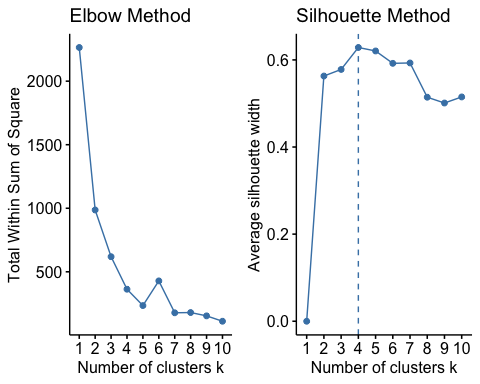
Elbow\_method <- fviz\_nbclust(Norm\_Energy\_data\_1, kmeans, method = "wss")+ggtitle("Elbow Method")

**Using the Silhouette Method Finding the best value of “K”**

Silhouette\_method <- fviz\_nbclust(Norm\_Energy\_data\_1, kmeans, method = "silhouette")+ggtitle("Silhouette Method")

**Plotting the graphs for the elbow menthod and silhouette method to determine the best value of K**

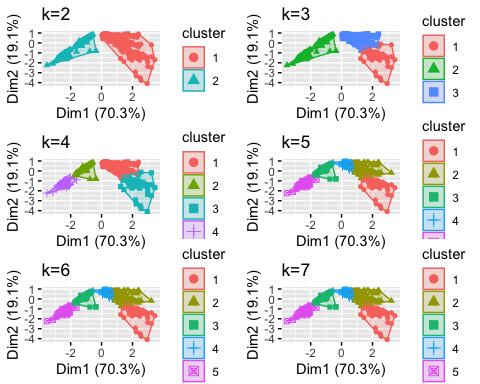
plot\_grid(Elbow\_method, Silhouette\_method, nrow = 1)



#From the above results of elbow method indicate Graph is not clear to choose k=2, 4, 5 or 7. from the silhouette method, got to know that K = 4. So, considering the silhouette method the value of k is 4

#Attempting to determine the ideal value of k. will examine every number between 2 and 7 and considering number of restarts = 30

k\_2<-kmeans(Norm\_Energy\_data\_1,centers =2,nstart=30)  
k\_3<-kmeans(Norm\_Energy\_data\_1,centers =3,nstart=30)  
k\_4<-kmeans(Norm\_Energy\_data\_1,centers =4,nstart=30)  
k\_5<-kmeans(Norm\_Energy\_data\_1,centers =5,nstart=30)  
k\_6<-kmeans(Norm\_Energy\_data\_1,centers =6,nstart=30)  
k\_7<-kmeans(Norm\_Energy\_data\_1,centers =7,nstart=30)  
p\_2<-fviz\_cluster(k\_2,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=2")  
p\_3<-fviz\_cluster(k\_3,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=3")  
p\_4<-fviz\_cluster(k\_4,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=4")  
p\_5<-fviz\_cluster(k\_5,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=5")  
p\_6<-fviz\_cluster(k\_5,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=6")  
p\_7<-fviz\_cluster(k\_5,geom = "point", data=Norm\_Energy\_data\_1)+ggtitle("k=7")  
grid.arrange(p\_2,p\_3,p\_4,p\_5,p\_6,p\_7)



*Thus from the above cluster plots of different k values, i can see that K=4 cluster is creating 4 best clusters, Hence choosing the value of k=4*

**Thus, i can conclude that the best value of K is 4**

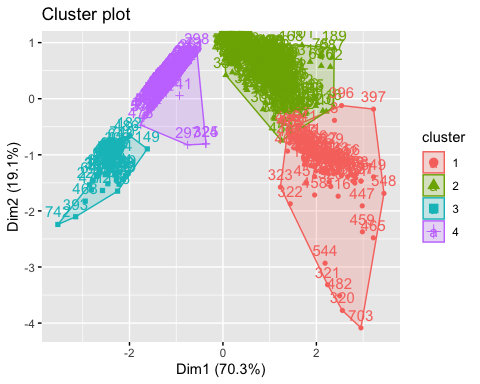
**Using the K-Means, Clustering the data**

set.seed(3695)  
k\_m <- kmeans(Norm\_Energy\_data\_1, centers = 4, nstart = 50) # k = 4, number of restarts = 50  
#Output Visualization  
  
#Centroids  
k\_m$centers

## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## 1 2.31460610 0.8087649 -0.7318205  
## 2 0.03743433 0.7622339 -0.6717529  
## 3 -0.57343169 -0.9756003 2.6109757  
## 4 -0.68285701 -0.9756003 0.4267807

**Visualizing Clustering Results**

fviz\_cluster(k\_m, data = Norm\_Energy\_data\_1)



k\_m

## K-means clustering with 4 clusters of sizes 90, 332, 67, 267  
##   
## Cluster means:  
## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## 1 2.31460610 0.8087649 -0.7318205  
## 2 0.03743433 0.7622339 -0.6717529  
## 3 -0.57343169 -0.9756003 2.6109757  
## 4 -0.68285701 -0.9756003 0.4267807  
##   
## Clustering vector:  
## [1] 2 2 4 2 2 2 4 1 2 2 2 2 2 3 4 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [38] 2 2 4 4 2 2 2 2 2 2 1 2 2 1 2 2 2 1 3 3 3 3 3 2 2 4 4 2 2 2 2 2 2 2 3 4 4  
## [75] 2 4 4 4 4 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 2 2 2 2 4 2 2 2 4 4 4 4 4 4 4 2  
## [112] 2 2 2 2 2 1 2 2 2 2 1 1 1 1 1 3 3 4 3 4 4 4 4 4 4 4 4 2 4 2 2 2 4 4 2 2 2  
## [149] 3 4 4 2 2 3 4 3 4 4 4 4 4 4 4 4 4 2 2 2 4 2 2 2 2 4 4 2 2 2 2 4 2 3 2 2 4  
## [186] 4 2 4 2 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 3 4  
## [223] 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4  
## [260] 4 4 4 4 2 2 2 2 2 2 3 4 4 4 4 4 4 4 4 4 4 4 4 2 2 2 2 2 2 2 2 2 2 2 2 4 4  
## [297] 4 4 2 2 4 2 2 2 2 3 2 2 3 4 4 2 4 4 4 4 4 2 4 1 1 1 1 4 4 4 4 4 3 4 4 2 2  
## [334] 2 2 2 2 4 4 4 4 4 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 3 2 2  
## [371] 2 2 2 3 2 2 2 2 2 2 2 2 2 2 3 3 4 4 1 1 3 1 3 4 4 1 1 1 1 1 1 1 1 1 3 3 2  
## [408] 2 2 2 2 3 3 2 2 1 1 3 4 2 3 1 1 1 1 1 1 3 1 4 1 4 2 2 2 1 1 4 2 1 1 1 2 2  
## [445] 3 4 1 4 4 4 4 4 4 2 4 2 1 1 1 1 2 2 2 2 1 2 1 3 1 2 3 2 4 4 2 4 4 4 2 2 4  
## [482] 1 2 2 2 3 4 2 2 2 2 2 2 2 2 2 3 3 4 2 2 2 2 2 2 2 2 2 4 4 2 2 2 2 2 2 3 4  
## [519] 2 2 2 2 3 2 4 1 2 4 4 4 2 2 4 2 2 1 4 4 4 4 4 4 4 1 2 4 4 1 1 3 3 2 2 2 2  
## [556] 2 2 2 2 2 2 2 4 4 4 2 4 2 4 4 4 4 2 2 4 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [593] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 3 3 4 2 2 2 2 1 1 2 2 2 1 1 1  
## [630] 2 2 1 1 1 3 3 1 1 1 1 1 3 3 3 4 2 2 1 1 4 1 1 1 1 1 1 1 1 4 4 1 1 1 1 1 1  
## [667] 1 1 1 3 1 2 2 2 2 2 1 3 3 3 2 2 2 2 2 2 2 2 2 3 3 2 2 2 3 3 3 2 4 4 4 4 1  
## [704] 2 2 2 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 3 4 3 4 3 4 4 4 3  
## [741] 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
##   
## Within cluster sum of squares by cluster:  
## [1] 99.90285 212.64214 14.54854 35.68235  
## (between\_SS / total\_SS = 84.0 %)  
##   
## Available components:  
##   
## [1] "cluster" "centers" "totss" "withinss" "tot.withinss"  
## [6] "betweenss" "size" "iter" "ifault"

#K-means clustering with 4 clusters of sizes 90, 332, 67, 267

#K-Means Cluster Analysis- Fit the data with 4 clusters  
fit<-kmeans(Norm\_Energy\_data\_1,4)  
  
#Finding the mean value of all quantitative variables for each cluster  
aggregate(Norm\_Energy\_data\_1,by=list(fit$cluster),FUN=mean)

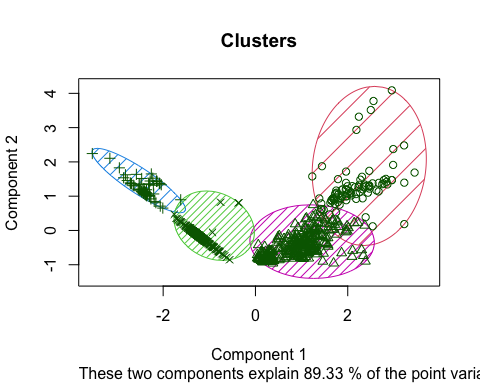
## Group.1 sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu  
## 1 1 2.31460610 0.8087649 -0.7318205  
## 2 2 -0.57343169 -0.9756003 2.6109757  
## 3 3 0.03743433 0.7622339 -0.6717529  
## 4 4 -0.68285701 -0.9756003 0.4267807

Norm\_Energy\_data\_2 <-data.frame(Norm\_Energy\_data\_1,fit$cluster)  
head(Norm\_Energy\_data\_2)

## sulfur\_content\_pct ash\_content\_pct fuel\_cost\_per\_mmbtu fit.cluster  
## 1 -0.2306101 0.07903011 -0.6846187 3  
## 2 -0.2401646 0.13762069 -0.6883690 3  
## 3 -0.6987812 -0.97560030 0.5334835 4  
## 4 0.9159315 1.89533804 -0.5644211 3  
## 5 0.1037979 2.05157959 -0.4509741 3  
## 6 0.7726138 1.87580785 -0.6726176 3

**visulizing the clusterplot with the help of cluspot function in R**

clusplot(Norm\_Energy\_data\_1,k\_m$cluster, main="Clusters" ,shade=TRUE ,color = TRUE, labels = 8,lines = 0)



**Summary Section**

1. What is the best value of K ?

**The Best value of K is 4**

1. Describe your clusters, provide relevant tables and graphs to support your conclusion?

**The 4 clusters as defined with the relevant tables and graphs to support my conclusion in the above explanation**

1. What can you say about the relative composition of the different fuel types in relation to your clusters?

**Interpretation of the four clusters about the relative composition of the different fuel types in relation to the clusters**

**Cluster 1:** All three of the variables in this cluster have high values. This implies that these businesses are paying high fuel expenses in addition to utilizing fuels with high sulfur and ash contents.

**Cluster 2:** This cluster has a lower ash level but a higher sulfur content and higher fuel expenditures. This indicates that these businesses are using fuels with a high sulfur content. But, by switching to fuels with a lower ash level, they would be able to minimize their fuel expenses.

**Cluster 3:** This cluster has a lower sulfur content but a higher ash content and higher fuel expenditures. This indicates that these businesses are using fuels with a high ash content, but by switching to fuels with a reduced sulphur level, they would be able to minimize their fuel expenses.

**Cluster 4:** This cluster is distinguished by low fuel prices, low ash and sulfur levels. This implies that these companies are using fuels with low levels of ash content and sulfur content and that their fuel expenses are likewise low.

**With the use of my clustering data, power generation companies can target particular groups of businesses with interventions designed to meet their unique requirements. For instance, companies that generate electricity could:**

Assist companies in Cluster 1 in cutting expenses and fuel usage by providing them with energy-efficient and renewable energy initiatives.

Provide incentives to companies in Cluster 2 to convert to fuels with less ash content.

Provide incentives to companies in Cluster 3 so they will convert to fuels with less sulfur.

Give companies in Cluster 4 advice on how to keep their fuel economy and low fuel expenses.

Through the application of cluster analysis findings, energy producing firms can assist their clients in lowering fuel expenses and enhancing their environmental performance.

**My segmentation will help to understand power generation in the US to focus on each cluster is as follows**

**Cluster 1:** Companies in this cluster may be eligible for a discount from power generation companies on installations of renewable energy sources or energy efficiency improvements. This would assist businesses in Cluster 1 in lowering their fuel expenses and consumption.

**Cluster 2:** Because natural gas has less ash than other fossil fuels, power companies could provide incentives to businesses in this cluster to switch to it. By doing this, businesses in Cluster 2 would be able to lower their fuel expenses without having to significantly alter their current infrastructure.

**Cluster 3:** Since ultra-low sulfur diesel has a lower sulfur content than other fossil fuels, power companies could provide incentives to businesses in this cluster to switch to it. By doing this, businesses in Cluster 3 would be able to lower their fuel expenses without having to significantly alter their current infrastructure.

**Cluster 4:** Businesses in this cluster could receive advice from power generation businesses about how to keep their fuel efficiency and low fuel prices. This could contain details on the best ways to maintain equipment, manage fuel, and train drivers.

**Finally, i can say that Power generation businesses can assist their clients in lowering fuel costs, enhancing their environmental performance, and gaining competitiveness by putting these kinds of interventions into practice.**

**Thanks and Regards**

**Tejesh Varma Maddana**