**CSE – 6005 – Machine Learning**

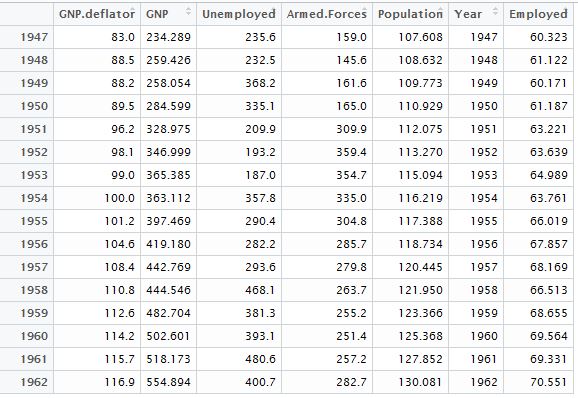
**Lab Experiment – 09 - Implement K-means Clustering to find natural pattern in data.**

#Kmeans

#first step you insert a dataset follow below commands

s <- longley

View(s)



#after partition cluster 4 parts

c <- kmeans(s,4)

summary(c)

Length Class Mode

cluster 16 -none- numeric

centers 28 -none- numeric

totss 1 -none- numeric

withinss 4 -none- numeric

tot.withinss 1 -none- numeric

betweenss 1 -none- numeric

size 4 -none- numeric

iter 1 -none- numeric

ifault 1 -none- numeric

c #to view full summary of k-means

K-means clustering with 4 clusters of sizes 3, 4, 4, 5

Cluster means:

GNP.deflator GNP Unemployed Armed.Forces Population

1 97.76667 347.1197 196.70 341.3333 113.4797

2 103.55000 405.6325 306.00 301.3250 118.1965

3 87.30000 259.0920 292.85 157.8000 109.2355

4 114.04000 500.5836 424.76 262.0400 125.7234

Year Employed

1 1952.0 63.94967

2 1955.5 66.45150

3 1948.5 60.70075

4 1960.0 68.92280

Clustering vector:

1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958

3 3 3 3 1 1 1 2 2 2 2 4

1959 1960 1961 1962

4 4 4 4

Within cluster sum of squares by cluster:

[1] 2448.990 9007.647 15903.502 15888.535

(between\_SS / total\_SS = 87.8 %)

Available components:

[1] "cluster" "centers" "totss"

[4] "withinss" "tot.withinss" "betweenss"

[7] "size" "iter" "ifault"

# to plot a graph

plot(s$GNP,s$Year,col=c$cluster, pch=16)

