```
raw <- read.csv("unempstates.csv")
View(raw)

## time sequence plots of three series
plot(raw[,5],type="l",ylim=c(0,12),xlab="unemployment rate",ylab="Months")
##CA
points(raw[,32],type="l", cex = .5, col = "dark red") ## New York
points(raw[,15],type="l", cex = .5, col = "dark green") ## Iowa

## transpose the data
## then we have 50 rows (states) and 416 columns (time periods)
rawt=matrix(nrow=50,ncol=416)
rawt=t(raw)
View(rawt[1:3,])


## k-means clustering in 416 dimensions
set.seed(1)
grpunemp2 <- kmeans(rawt, centers=2, nstart=10)
sort(grpunemp2$cluster)

grpunemp3 <- kmeans(rawt, centers=3, nstart=10)
sort(grpunemp3$cluster)

grpunemp4 <- kmeans(rawt, centers=4, nstart=10)
sort(grpunemp4$cluster)

grpunemp5 <- kmeans(rawt, centers=5, nstart=10)
sort(grpunemp5$cluster)


## another analysis
## data set unemp.csv with means and standard deviations for each state
## k-means clustering on 2 dimensions (mean, stddev)
unemp <- read.csv("unemp.csv")
unemp[1:3,]

set.seed(1)
grpunemp <- kmeans(unemp[,c("mean","stddev")], centers=3, nstart=10)
sort(grpunemp3$cluster)
## list of cluster assignments
o=order(grpunemp$cluster)
data.frame(unemp$state[o],grpunemp$cluster[o])

text(x=unemp$mean,y=unemp$stddev,labels=unemp$state, col=grpunemp$cluster+1)
```