Lab 17

quan le

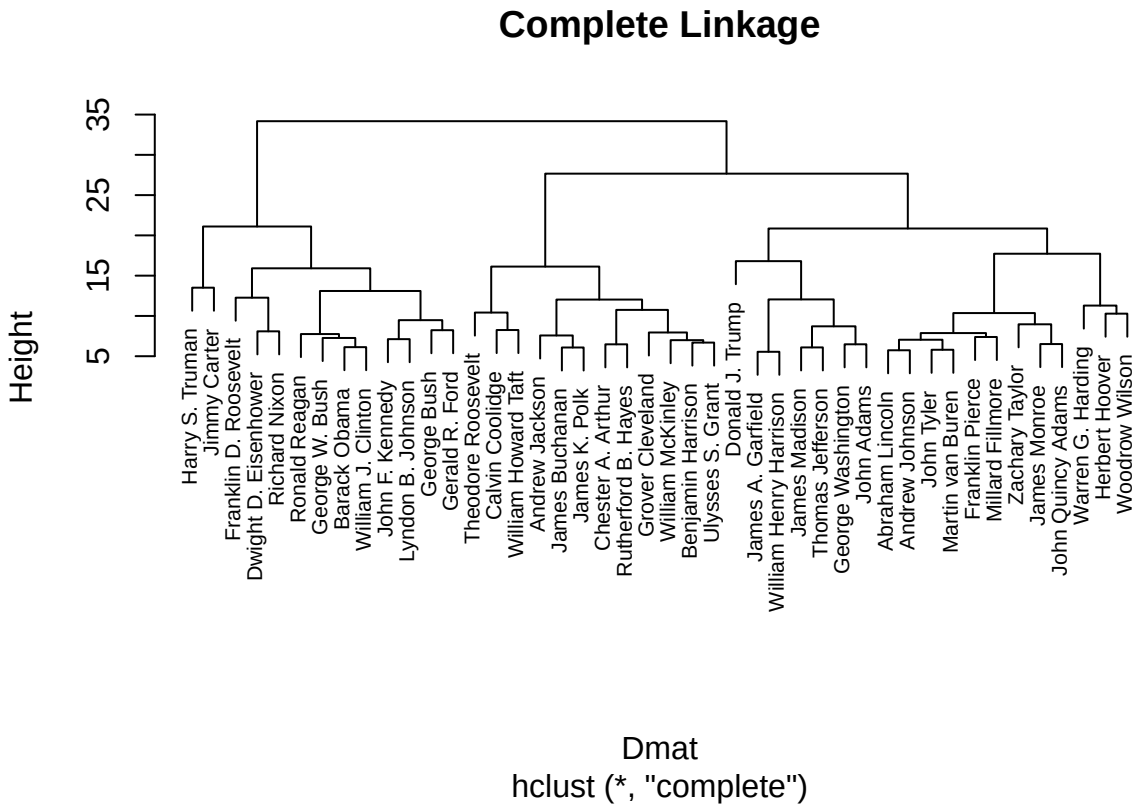
2022-10-24

Hierarchical Clustering

```
# data from https://github.com/DataSlingers/clustRviz/tree/master/data
load("data/presidential_speech.rda")
```

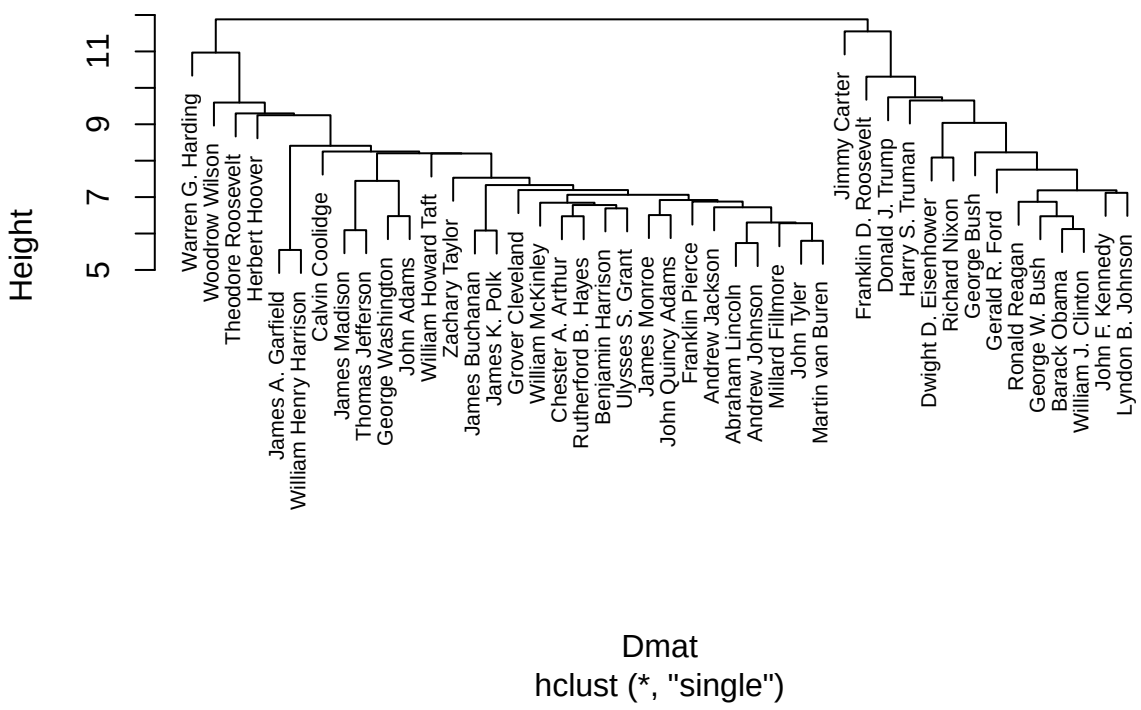
Complete linakge - Euclidean distance

```
cols = as.numeric(as.factor(rownames(presidential_speech)))
Dmat = dist(presidential_speech)
com.hclust = hclust(Dmat,method="complete")
plot(com.hclust,cex=.7,main="Complete Linkage")
```



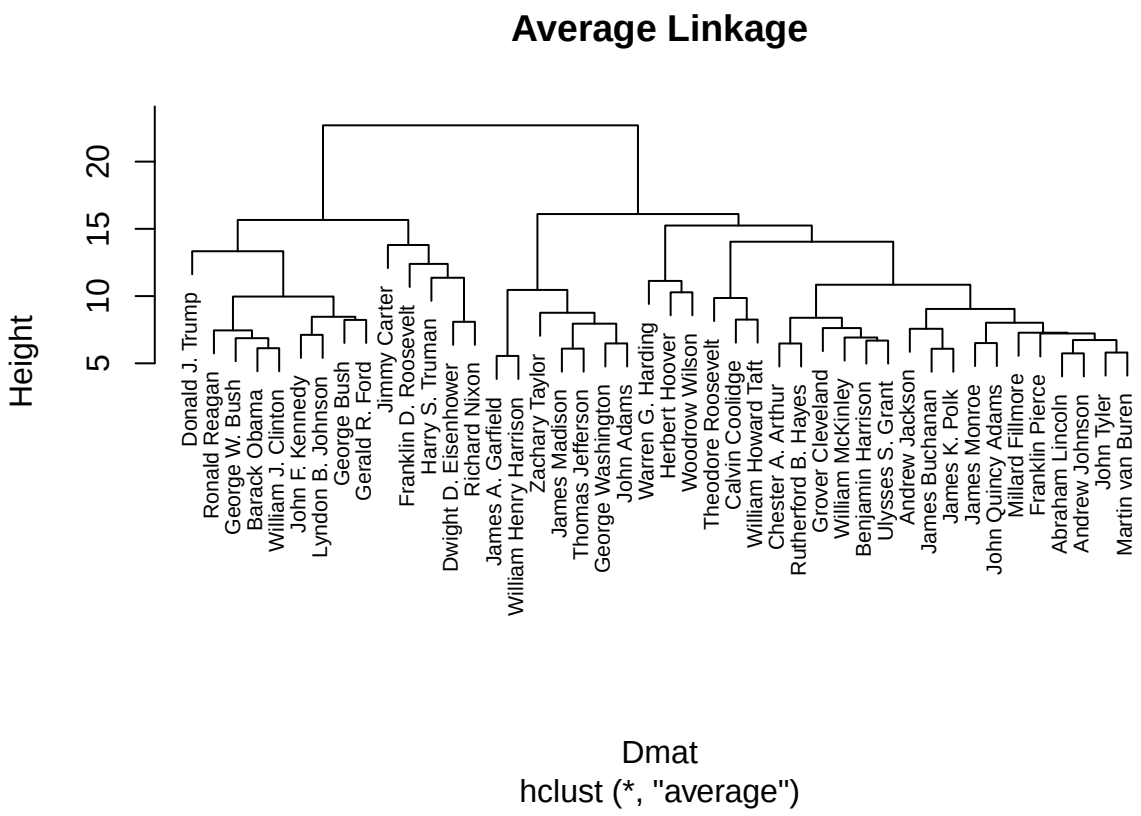
Single linakge

```
sing.hclust = hclust(Dmat,method="single")
plot(sing.hclust,cex=.7,main="Single Linkage")
```



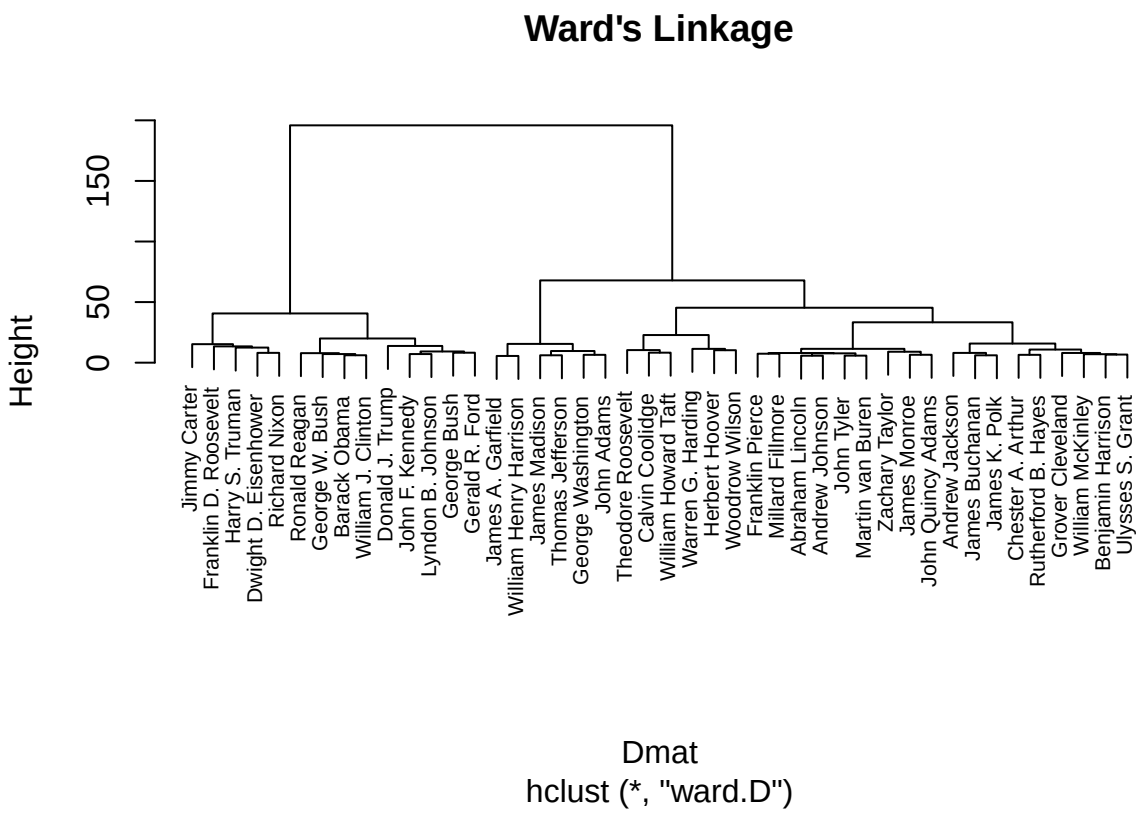
Average linakge

```
ave.hclust = hclust(Dmat,method="average")
plot(ave.hclust,cex=.7,main="Average Linkage")
```



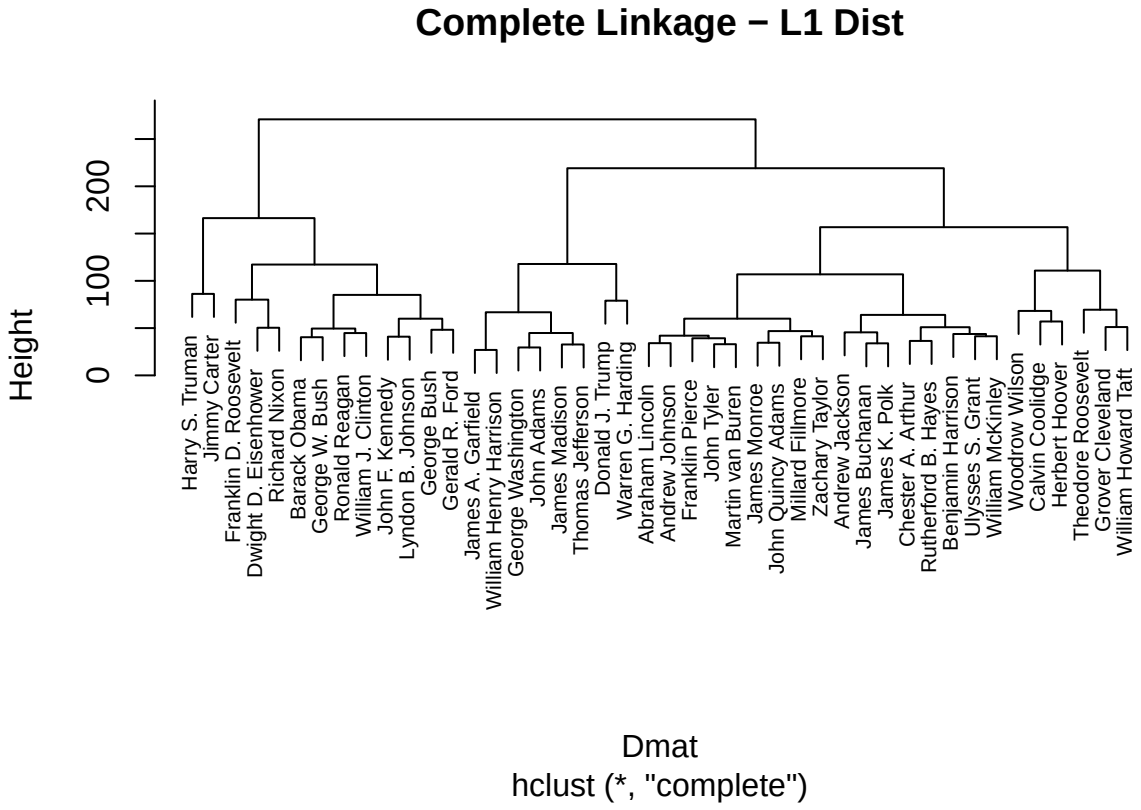
Ward's linakge

```
ward.hclust = hclust(Dmat,method="ward.D")
plot(ward.hclust,cex=.7,main="Ward's Linkage")
```



Complete linkage with different distances - L1 distance

```
Dmat = dist(presidential_speech,method="manhattan") #L1 distance
com.hclust = hclust(Dmat,method="complete")
plot(com.hclust,cex=.7,main="Complete Linkage - L1 Dist")
```



Biclustering - Cluster Heatmap

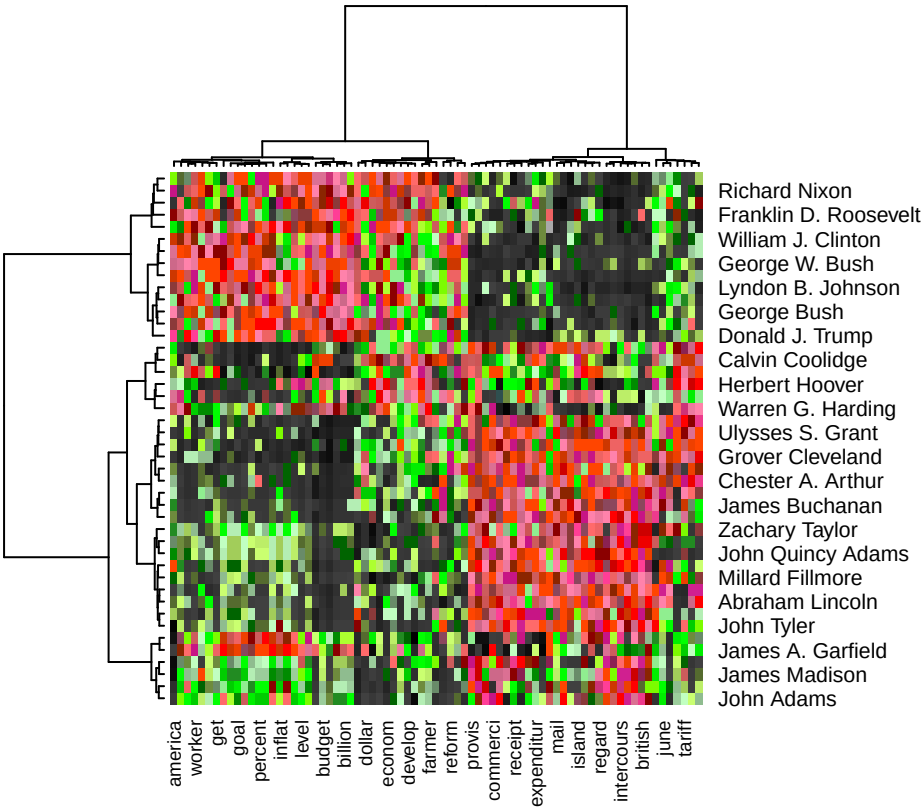
```
X = scale(presidential_speech,center=TRUE,scale=FALSE)
sv = svd(X)
V = sv$v
D = sv$d
U = sv$u
V = sv$v
Z = X%*%V


aa = grep("grey",colors())
bb = grep("green",colors())
cc = grep("red",colors())
gcol2 = colors()[c(aa[1:30],bb[1:20],rep(cc,2))]

j = 2
ord = order(abs(V[,j]),decreasing=TRUE)
x = as.matrix(X[,ord])
```

cluster heatmap - uses Ward's linkage (complete is default)

```
heatmap(x,col=gcol2,hclustfun=function(x) hclust(x,method="ward.D"))
```

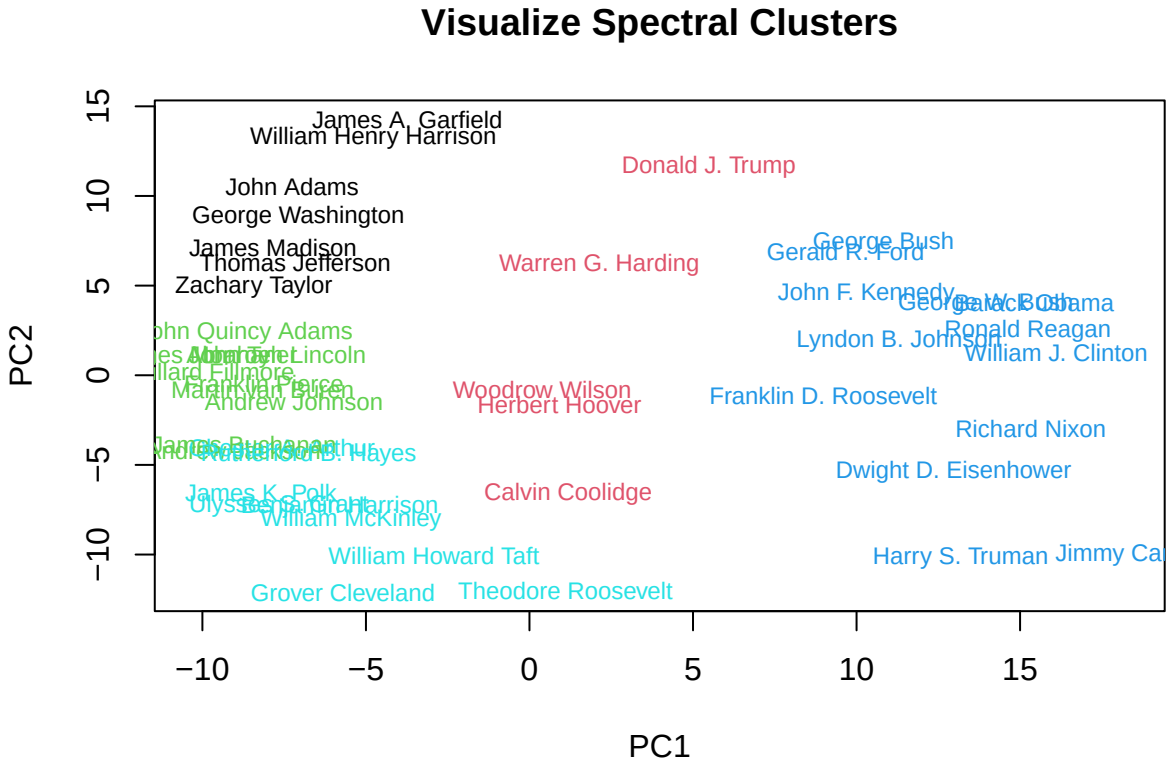


Spectral Clustering

```
library(kknn)
K = 5
pres.spec = specClust(presidential_speech, centers=K, nn = 7, method = "symmetric", gmax=NULL)
```

Visualize

```
plot(Z[,1],Z[,2],col = pres.spec$cluster, type = 'n',main = "Visualize Spectral Clusters", xlab = "PC1"
text(Z[,1],Z[,2],rownames(presidential_speech),cex=.75,col=pres.spec$cluster)
```


Visualize Spectral Clusters
James A. Garfield
William Henry Harrison
Donald J. Trump
John Adams
George Washington
James Madison
Thomas Jefferson
Warren G. Harding
George Bush
Gerald R. Ford
Zachary Taylor
John F. Kennedy
Barack Obama
Ronald Reagan
Lyndon B. Johnson
William J. Clinton
Woodrow Wilson
Herbert Hoover
Franklin D. Roosevelt
Andrew Johnson
Richard Nixon
Dwight D. Eisenhower
James K. Polk
Calvin Coolidge
William McKinley
William Howard Taft
Harry S. Truman
Grover Cleveland
Theodore Roosevelt
PC1
PC2
−10
−5
0
5
10
15