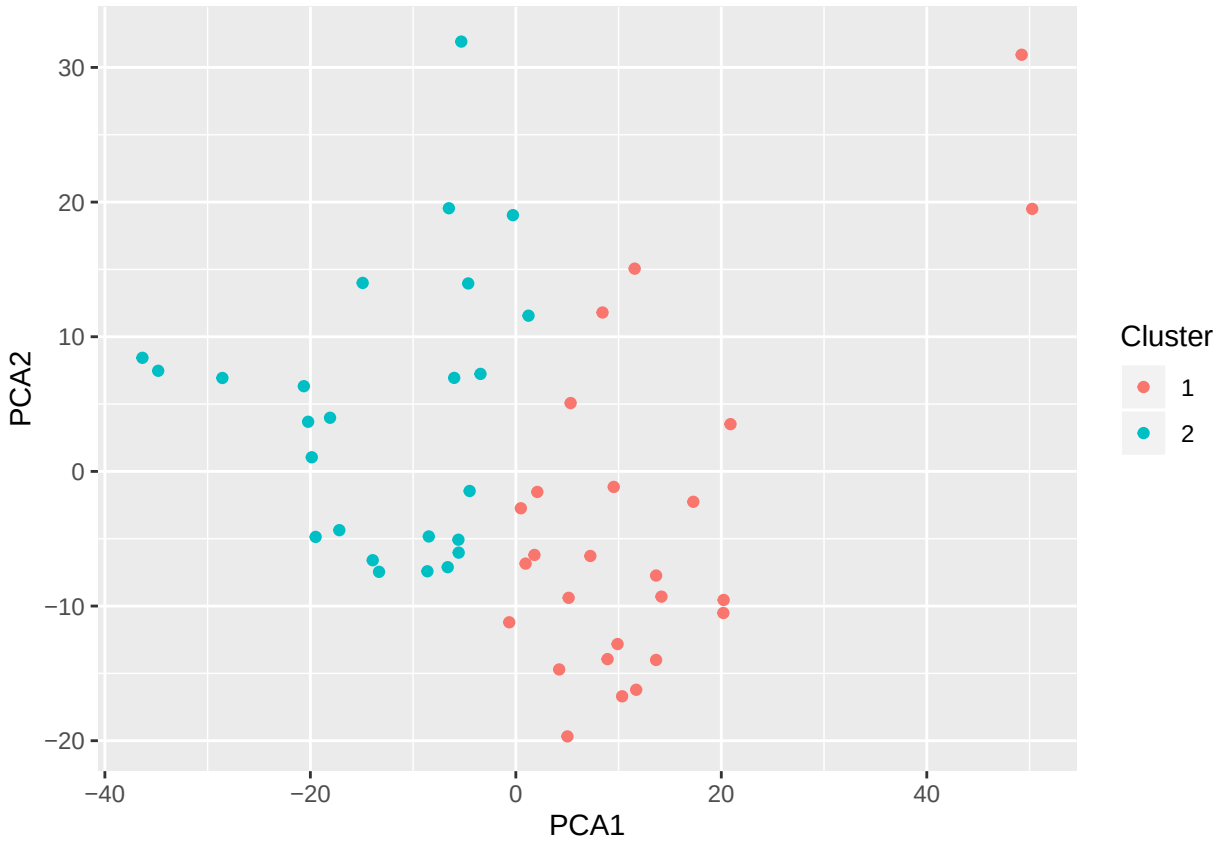
Activity 5

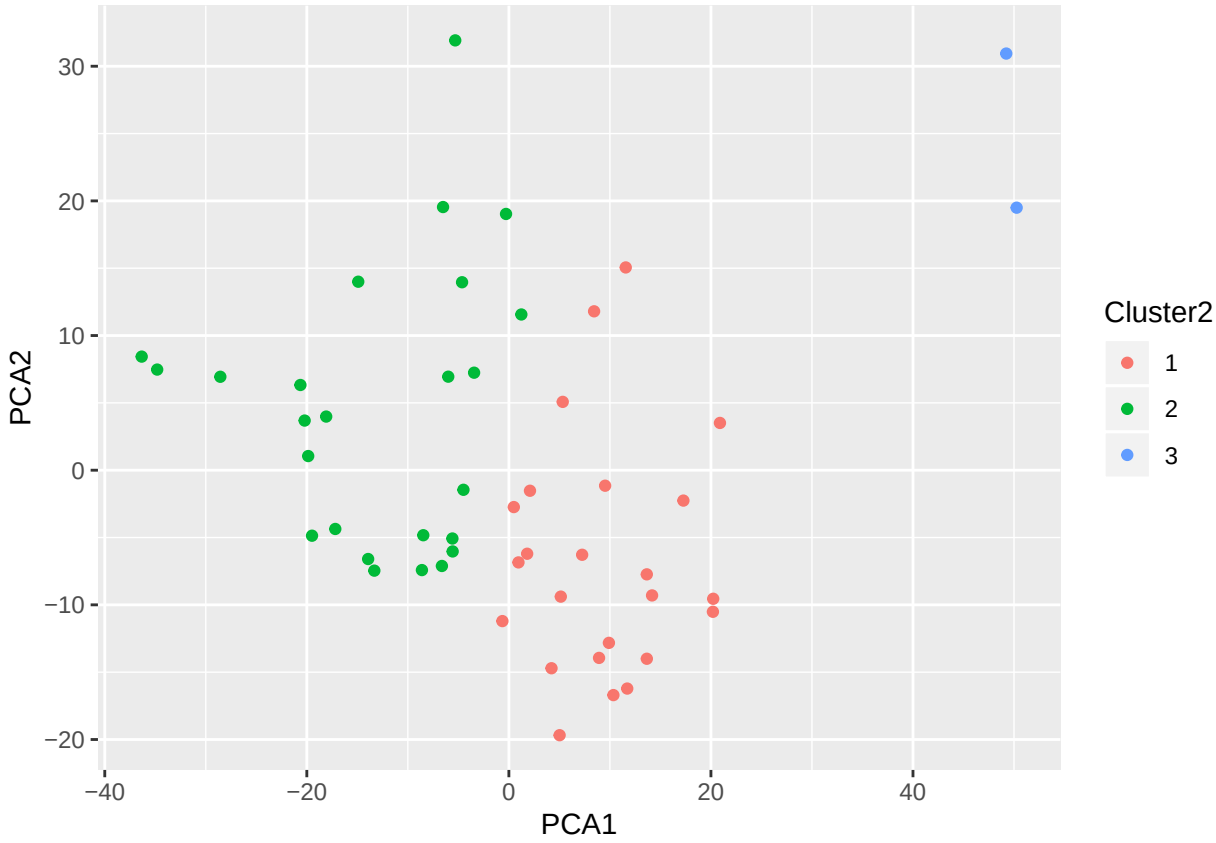
Alice Chang

```
library(ggplot2)
#Load data
poverty <- read.delim("https://bitly.com/381pd5e")

poverty <- poverty[-1]
set.seed(1)

pca <- prcomp(poverty)
dat.loadings <- pca$x[,1:2]
cl <- kmeans(dat.loadings, centers=2)
pca1 <-pca$x[,1]
pca2 <-pca$x[,2]

mydf<-data.frame(PCA1=pca1, PCA2=pca2, Cluster=factor(cl$cluster))

ggplot(mydf, aes(x=PCA1, y=PCA2, color=Cluster)) +
       geom_point()
```

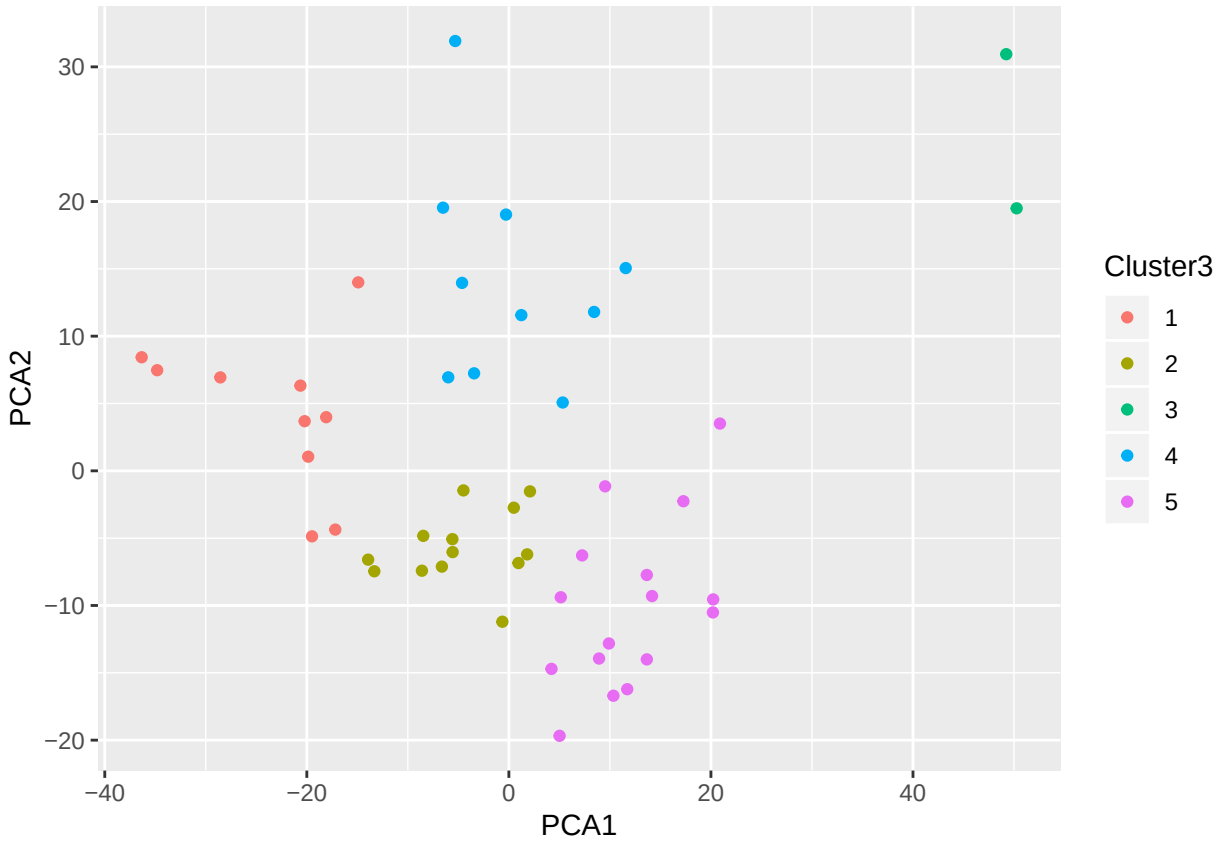



```
#k = 3
c2 <- kmeans(dat.loadings, centers=3)
mydf2<-data.frame(PCA1=pca1, PCA2=pca2, Cluster2=factor(c2$cluster))
ggplot(mydf2, aes(x=PCA1, y=PCA2, color=Cluster2)) +
```

```
        geom_point()
```




```
#k = 5
c3 <- kmeans(dat.loadings, centers=5)
mydf3<-data.frame(PCA1=pca1, PCA2=pca2, Cluster3=factor(c3$cluster))
ggplot(mydf3, aes(x=PCA1, y=PCA2, color=Cluster3)) +
        geom_point()
```


PCA2
PCA1
30
20
10
0
−10
−20
−40
−20
0
20
40
Cluster3
1
2
3
4
5