R Commands for Clustering

```
> gg ggplot(iris, aes(petallength, petalwidth, idth, color =
class)) + geom_point()
> irisCluster <- kmeans(iris[, 3:4], 3, nstart = 20)
> irisCluster$cluster <- as.factor(irisCluster$cluster)
> ggplot(iris, aes(petallength, petalwidth, color =
iris$cluster)) + geom_point()
> intraclust = c("complete","average","centroid")
> interclust = c("single", "complete", "average","centroid",
"aveToCent", "hausdorff")
> # compute Dunn indicies (also Davies-Bouldin indicies)
> # 1. optimal solution:
> # compute intercluster distances and intracluster diameters
> install.packages("clv")
> cls.scatt <- cls.scatt.data(iris.data, irisCluster$cluster,
dist="manhattan")
> dunn1 <- clv.Dunn(cls.scatt, intraclust, interclust)
> davies1 <- clv.Davies.Bouldin(cls.scatt, intraclust,
interclust)

# Standardize glass, transform to a dataframe: glass_sc
> glass = glass[,1:9]
> glassCluster <- kmeans(glass, 7, nstart = 20)
> glass_sc <- as.data.frame(scale(glass))
OU
> glassCluster <- kmeans(glass, 7, nstart = 20)
>
> dun=clv.Dunn(cls.scatt.data(glass, glassCluster$cluster,
dist="euclidean"), c("centroid"), c("centroid"))
> dun
          cent
cent 0.8315106
> glass_pp = preProcess(glass, method=c("center", "scale"))
> glass_pp = predict(glass_pp, glass)
> glass_ppCluster <- kmeans(glass_pp, 7, nstart = 20)
>
> dun_pp=clv.Dunn(cls.scatt.data(glass_pp,
glass_ppCluster$cluster, dist="euclidean"), c("centroid"),
c("centroid"))
> dun_pp
          cent
cent 0.6324083
> dun_pp=clv.Dunn(cls.scatt.data(glass_pp,
glass_ppCluster$cluster, dist="euclidean"), c("average"),
c("aveToCent"))
> dun_pp
            ave
aveto 0.4748731
> dun=clv.Dunn(cls.scatt.data(glass, glassCluster$cluster,
dist="euclidean"), c("average"), c("aveToCent"))
> dun
            ave
aveto 0.5901392
```