library(corrplot) #correlation plot

library(ggplot2) #support scatterplot

library(GPArotation) #support rotation

library(psych) #PCA package

url="<http://textuploader.com/ae6t4/raw>"

nhl = as.data.frame(read.csv(url, header=FALSE))

nhl

str(nhl)

names(nhl) = c("rank","team","played","wins","losses","OTL","pts","ROW","HROW","RROW","ppc","gg","gag","five","PPP","PKP","shots","sag","sc1","tr1","lead1","lead2","wop","wosp","face")

nhl=nhl[order(nhl$gg),]

nhl[1,2]

nhl[30,2]

pca.df = nhl[,c(-1:-11)]

pca.df = as.data.frame(lapply(pca.df, as.numeric))

str(pca.df)

nhl.cor = cor(pca.df)

corrplot(nhl.cor, method="ellipse")

pca = principal(pca.df, nfactors=5, rotate="none")

pca

plot(pca$values, type="b", ylab="Eigenvalues", xlab="Component")

pca.rotate = principal(pca.df, nfactors=3, rotate = "varimax")

pca.rotate

pca.scores = pca.rotate$scores

pca.scores = as.data.frame(pca.scores)

pca.scores

nhl$pts = as.numeric(nhl$pts)

nhl$Def = pca.scores$PC1

nhl$Off = pca.scores$PC3

nhl$PPlay = pca.scores$PC2

nhl.lm = lm(pts~Def+Off+PPlay, data=nhl)

summary(nhl.lm)

nhl.lm2 = lm(pts~Def+Off, data=nhl)

summary(nhl.lm2)

plot(nhl.lm2$fitted.values, nhl$pts, main="Predicted versus Actual",xlab="Predicted",ylab="Actual")

nhl$pred = nhl.lm2$fitted.values

nhl=nhl[order(-nhl$pts),]

nhl.best = nhl[1:15,]

p = ggplot(nhl.best, aes(x=pred, y=pts, label=team))

p + geom\_point() +

geom\_text(size=3.5, hjust=.2, vjust=-0.5, angle=15) +

xlim(90,120) + ylim(90, 120) +

stat\_smooth(method="lm", se=FALSE)

#biplot(pca.rotate)

p2 = ggplot(nhl, aes(x=Def, y=Off, label=team))

p2 + geom\_point() +

geom\_text(size=2.75, hjust=.2, vjust=-0.75, angle=0) +

xlim(-3,3) + ylim(-2.5,2.5)

###########