install.packages("psych")

install.packages(xlxs)

> data <- read.xlsx("C:\\Users\\dell\\Desktop\\USA.xlsx")

summary(data)

library(psych)

describe(data[vars])

data.scale <- (data$score-min(data$score))/(max(data$score)-min(data$score))

data.scale

data.scale <- (max(data$score -data$score))/(max(data$score)-min(data$score))

data.scale

ls(data)

head(data)

data

#PCA

> data1 <- read.xlsx("C:\\Users\\dell\\Desktop\\PCA.xlsx")

head(data1)

head(data)

> library(openxlsx)

> data1 <- read.xlsx("D:\\Users\\dell\\Desktop\\PCA.xlsx")

> head(data1)

Parallel analysis suggests that the number of factors = NA and the number of components = 1

> head(data1)

head(data1)

fa.parallel(data1, fa = 'pc', n.iter = 100, show.legend = FALSE)

data1\_pca <- principal(data1, nfactors = 2, rotate = "none")

> data1\_pca

Principal Components Analysis

Call: principal(r = data1, nfactors = 2, rotate = "none")

Standardized loadings (pattern matrix) based upon correlation matrix

data1\_pca <- principal(data1, nfactors = 2, rotate = "none")

data1\_pca

> data1\_pca

data1\_pca <- principal(data1, nfactors = 3, rotate = "none")

data1\_pca

data1\_pca\_xz <- principal(data1, nfactors = 2, rotate = "varimax")

> data1\_pca\_xz

data1\_pca\_xz <- principal(data1, nfactors = 3, rotate = "varimax")

data1\_pca\_xz

data1=read.xlsx("C:\\Users\\dell\\Desktop\\PCA.xlsx",1)

fa.parallel(data1, fa = 'pc', n.iter = 100, show.legend = FALSE)

data1\_pca\_xz <- principal(data1, nfactors = 2, rotate = "varimax")

data1\_pca\_xz

data1\_pca\_xz1 <- principal(data1, nfactors = 2, rotate = "varimax", scores = TRUE)

options(digits = 2)

head(data1\_pca\_xz1$scores)

data1\_pca\_xz1$scores

data1\_pca\_xz1$scores

# spatial autocorrelation analysis

score=data1\_pca\_xz1$scores

View(score)

write(score,file="C:\\Users\\dell\\Desktop\\score.xlsx")

View(score)

library(rgdal)  
shpt <- readOGR("states.shp")  
cdatashpt <- merge(shpt, cdata, by = "STATE\_NAME")

moran(cdatashpt$gdp2017,listw=score, =length(cdatashpt$gdp2017), S0=Szero(score))

moran.plot(cdatashpt$gdp2017, zero.policy=NULL, spChk=NULL, labels=TRUE, xlab=NULL, ylab=NULL, quiet=NULL)

localmoran(cdatashpt$gdp2017, alternative = "greater")

install.packages("factoextra")

#cluster

library(factoextra)

View(data3)

fviz\_nbclust(data3, kmeans, method = "wss") + geom\_vline(xintercept = 4, linetype = 2)

distance <- dist(data3)

data3.hc <- hclust(distance)

plot(china.hc, hang = -1)

re <- rect.hclust(data.hc, k = 4)

refor (i in 1:4) { print(paste("",i,"")) print(score[re[[i]],]$state)}

distance <- dist(data3)

data3.hc <- hclust(distance)

plot(data3.hc, hang = -1)

result <- dist(data3, method = "euclidean")

result\_hc <- hclust(d = result, method = "ward.D2")

fviz\_dend(result\_hc, cex = 0.6)

fviz\_nbclust(data3, kmeans, method = "wss") + geom\_vline(xintercept = 4, linetype = 2)

km\_result <- kmeans(data3, 4, nstart = 24)

options(stringsAsFactors = TRUE)

class(data3)

data3.hc <- hclust(distance)

plot(data3.hc, hang = -1)

result <- dist(data3, method = "euclidean")

as.character(data3)

a=data3[,1]

as.character(a)

data3[,1]=a

hc<-hclust(dist(data3),method = "ave")

plot(hc,hang=-1)

plot(hc,hang=-1,labels=data3$Species[a])

fviz\_nbclust(data3, kmeans, method = "wss") + geom\_vline(xintercept = 4, linetype = 2)

km\_result <- kmeans(data3, 4, nstart = 24)

print(km\_result)

fviz\_cluster(km\_result, data = data3,

palette = c("#2E9FDF", "#00AFBB", "#E7B800", "#FC4E07"),

ellipse.type = "euclid",

star.plot = TRUE,

repel = TRUE,

ggtheme = theme\_minimal()

)

fviz\_dend(result\_hc, k = 4,

cex = 0.5,

k\_colors = c("#2E9FDF", "#00AFBB", "#E7B800", "#FC4E07"),

color\_labels\_by\_k = TRUE,

rect = TRUE

)

data4=read.xlsx("C:\\Users\\dell\\Desktop\\cluster1.xlsx","Sheet1")

> distance <- dist(data4)

> data4.hc <- hclust(distance)

> plot(data4.hc, hang = -1)

# regression analysis

install.packages("gcookbook")

library(gcookbook)

data5=read.xlsx("C:\\Users\\dell\\Desktop\\USA.xlsx","Sheet1")

View(data5)

model <- lm(urbanization.rate~ci,data5)

model

summary(model)

ggplot(data5, aes(x=ci, y=urbanization.rate,color="#5e616d"))+

geom\_point()+

stat\_smooth(method = lm,color="blue")+

annotate("text", label = "R^2==0.22")+

annotate("text", label = "y=91.72+-6.99x",x=4,y=85)