Module Five

Chol Stephen Jok

2022-09-14

#Selecting cases

#load data

```
#?mtcars
#data("mtcars")
#mtcars
```

#mean quarter mile time (for all cars)

```
#mean(mtcars$qsec)
```

#mean quarter mile for 8 cylinder cars #Use square bracket to indicate what to select #In this format: [row]

```
#mean(mtcars$qsec[mtcars$cyl==8])
```

#Median horsepower (for all cars)

```
#median(mtcars$hp)
```

#Mean MPG for cars above median horsepower

```
#mean(mtcars$mpg[mtcars$hp>median(mtcars$hp)])
```

#Create new dataframe for 8 cylinder cars #To create a new dataframe, must indicate which row and which column to copy in this format #Format : [rows: columns] to sellect all column, leave second part blank

```
#cyl.8<-mtcars[mtcars$cyl==8, ]
```

#select 4 cylinder cars with 4+ barrels carburators

```
#mtcars[mtcars$cyl==8 & mtcars$carb>=4, ]
```

#Analyzing by subgroup #load datasets

```
#?iris
#data("iris")
#iris
#mean(iris$Petal.Width)
```

#split the data file and repeat analysis with “aggregate” #Compare group with one variable

```
#aggregate(iris$Petal.Width~iris$Species, FUN=mean)
```

#Compare group of several variable #Use cbind to list outcome variables

```
#aggregate(cbind(iris$Petal.Width,iris$Petal.Length)~iris$Species, FUN=mean)
```

#MERGING FILES #load data

```
#?longley
#data("longley")
```

#split up longley

```
#a1<-longley[1:14, 1:6]  #starting data
#a2<-longley[1:14, 6:7]  # New column to add with "year" to match
#b<-longley[15:16, ]   # New row to add

#write.table(a1, "~desktop/R/longley.a1.txt", sep = "\t")
#write.table(a2, "~desktop/R/longley.a2.txt", sep = "\t")
#write.table(b, "~desktop/R/longley.b.txt", sep = "\t")
```

#import data

```
#a1t<-read.table("~desktop/R/longley.a1.txt", sep = "\t")
#a2t<-read.table("~desktop/R/longley.a2.txt", sep = "\t")
```

#TAKE early year (a1t) and add column (a2t) #must specify variable to match cases

```
#a.1.2<-merge(a1t,a2t, by="year")   # merge two data frame
#a.1.2   # check results
```

#add two more cases at the bottom

```
#b<-read.table("~desktop/R/longley.b.txt", sep = "\t")
#all.data<-rbind(a.1.2,b)   # row bind
#row.names(all.data)<-NULL     # RESET ROW DATA
```