Chapter 04

# 07.

(1)

str(iris)

apply(iris[,1:4], 1, sum)

> apply(iris[,1:4], 1, sum)

[1] 10.2 9.5 9.4 9.4 10.2 11.4 9.7 10.1 8.9 9.6 10.8 10.0 9.3

[14] 8.5 11.2 12.0 11.0 10.3 11.5 10.7 10.7 10.7 9.4 10.6 10.3 9.8

[27] 10.4 10.4 10.2 9.7 9.7 10.7 10.9 11.3 9.7 9.6 10.5 10.0 8.9

[40] 10.2 10.1 8.4 9.1 10.7 11.2 9.5 10.7 9.4 10.7 9.9 16.3 15.6

[53] 16.4 13.1 15.4 14.3 15.9 11.6 15.4 13.2 11.5 14.6 13.2 15.1 13.4

[66] 15.6 14.6 13.6 14.4 13.1 15.7 14.2 15.2 14.8 14.9 15.4 15.8 16.4

[79] 14.9 12.8 12.8 12.6 13.6 15.4 14.4 15.5 16.0 14.3 14.0 13.3 13.7

[92] 15.1 13.6 11.6 13.8 14.1 14.1 14.7 11.7 13.9 18.1 15.5 18.1 16.6

[105] 17.5 19.3 13.6 18.3 16.8 19.4 16.8 16.3 17.4 15.2 16.1 17.2 16.8

[118] 20.4 19.5 14.7 18.1 15.3 19.2 15.7 17.8 18.2 15.6 15.8 16.9 17.6

[131] 18.2 20.1 17.0 15.7 15.7 19.1 17.7 16.8 15.6 17.5 17.8 17.4 15.5

[144] 18.2 18.2 17.2 15.7 16.7 17.3 15.8

(2)

apply(iris[,1:4], 2, max)

> apply(iris[,1:4], 2, max)

Sepal.Length Sepal.Width Petal.Length Petal.Width

7.9 4.4 6.9 2.5

# 08.

(1)

apply(mtcars, 2, sum)

> apply(mtcars, 2, sum)

mpg cyl disp hp drat wt qsec vs

642.900 198.000 7383.100 4694.000 115.090 102.952 571.160 14.000

am gear carb

13.000 118.000 90.000

(2)

apply(mtcars, 2, max)

> apply(mtcars, 2, max)

mpg cyl disp hp drat wt qsec vs am

33.900 8.000 472.000 335.000 4.930 5.424 22.900 1.000 1.000

gear carb

5.000 8.000

(3)

apply(mtcars, 2, sd)

> apply(mtcars, 2, sd)

mpg cyl disp hp drat wt

6.0269481 1.7859216 123.9386938 68.5628685 0.5346787 0.9784574

qsec vs am gear carb

1.7869432 0.5040161 0.4989909 0.7378041 1.6152000

# 09.

lgm <- function(x,y) {

min <- min(x,y)

max <- max(x,y)

remainder <- max %% min

result <- ifelse(remainder == 0, min, lgm(min, remainder))

return(result)

}

result1 <- lgm(10,8)

result1

result2 <- lgm(10,20)

result2

> lgm <- function(x,y) {

+ min <- min(x,y)

+ max <- max(x,y)

+ remainder <- max %% min

+ result <- ifelse(remainder == 0, min, lgm(min, remainder))

+ return(result)

+ }

>

> result1 <- lgm(10,8)

> result1

[1] 2

> result2 <- lgm(10,20)

> result2

[1] 10

# 10.

maxmin <- function(vec) {

max\_val <- max(vec)

min\_val <- min(vec)

return(list(max=max\_val, min=min\_val))

}

> maxmin <- function(vec) {

+ max\_val <- max(vec)

+ min\_val <- min(vec)

+ return(list(max=max\_val, min=min\_val))

+ }

> v1 <- c(7,1,2,8,9)

> result <- maxmin(v1)

> result$max ; result$min

[1] 9

[1] 1

> result <- maxmin(iris[,1])

> result$max ; result$min

[1] 7.9

[1] 4.3

# 11.

(1)

weight <- c(69, 50, 55, 71, 89, 64, 59, 70, 71, 80)

which.max(weight)

> which.max(weight)

[1] 5

(2)

which.min(weight)

> which.min(weight)

[1] 2

(3)

which(weight > 61 & weight < 69)

> which(weight > 61 & weight < 69)

[1] 6

(4)

wtIndx <- which(weight <= 60)

weight.2 <- weight[wtIndx]

weight.2

> wtIndx <- which(weight <= 60)

> wtIndx

[1] 2 3 7

> weight.2 <- weight[wtIndx]

> weight.2

[1] 50 55 59

# 12.

(1)

str(iris)

maxIndex <- which.max(iris$Petal.Length)

iris[maxIndex,]

> maxIndex <- which.max(iris$Petal.Length)

> iris[maxIndex,]

Sepal.Length Sepal.Width Petal.Length Petal.Width Species

119 7.7 2.6 6.9 2.3 virginica

(2)

Index <- which(iris$Petal.Width >= 0.3 & iris$Petal.Width <= 0.4)

iris[Index,]

Index <- which(iris$Petal.Width >= 0.3 & iris$Petal.Width <= 0.4)

> iris[Index,]

Sepal.Length Sepal.Width Petal.Length Petal.Width Species

6 5.4 3.9 1.7 0.4 setosa

7 4.6 3.4 1.4 0.3 setosa

16 5.7 4.4 1.5 0.4 setosa

17 5.4 3.9 1.3 0.4 setosa

18 5.1 3.5 1.4 0.3 setosa

19 5.7 3.8 1.7 0.3 setosa

20 5.1 3.8 1.5 0.3 setosa

22 5.1 3.7 1.5 0.4 setosa

27 5.0 3.4 1.6 0.4 setosa

32 5.4 3.4 1.5 0.4 setosa

41 5.0 3.5 1.3 0.3 setosa

42 4.5 2.3 1.3 0.3 setosa

45 5.1 3.8 1.9 0.4 setosa

46 4.8 3.0 1.4 0.3 setosa

Chapter 05

# 05.

(1)

wt <- mtcars$wt

mean(wt)

median(wt)

mean(wt, trim = 0.15)

sd(wt)

> wt <- mtcars$wt

> mean(wt)

[1] 3.21725

> median(wt)

[1] 3.325

> mean(wt, trim = 0.15)

[1] 3.165208

> sd(wt)

[1] 0.9784574

(2)

summary(wt)

> summary(wt)

Min. 1st Qu. Median Mean 3rd Qu. Max.

1.513 2.581 3.325 3.217 3.610 5.424

(3)

str(mtcars)

unique(mtcars$cyl)

mtcars$cyl <- as.factor(mtcars$cyl)

table(mtcars$cyl)

> str(mtcars)

'data.frame': 32 obs. of 11 variables:

$ mpg : num 21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...

$ cyl : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...

$ disp: num 160 160 108 258 360 ...

$ hp : num 110 110 93 110 175 105 245 62 95 123 ...

$ drat: num 3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...

$ wt : num 2.62 2.88 2.32 3.21 3.44 ...

$ qsec: num 16.5 17 18.6 19.4 17 ...

$ vs : num 0 0 1 1 0 1 0 1 1 1 ...

$ am : num 1 1 1 0 0 0 0 0 0 0 ...

$ gear: num 4 4 4 3 3 3 3 4 4 4 ...

$ carb: num 4 4 1 1 2 1 4 2 2 4 ...

> unique(mtcars$cyl)

[1] 6 4 8

Levels: 4 6 8

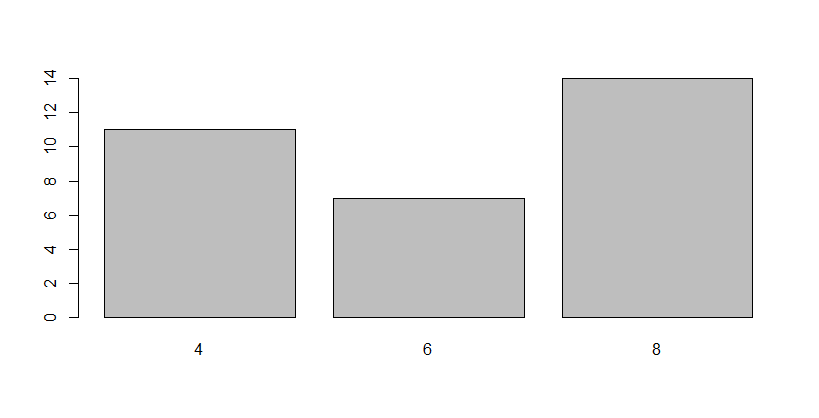
> mtcars$cyl <- as.factor(mtcars$cyl)

> table(mtcars$cyl)

4 6 8

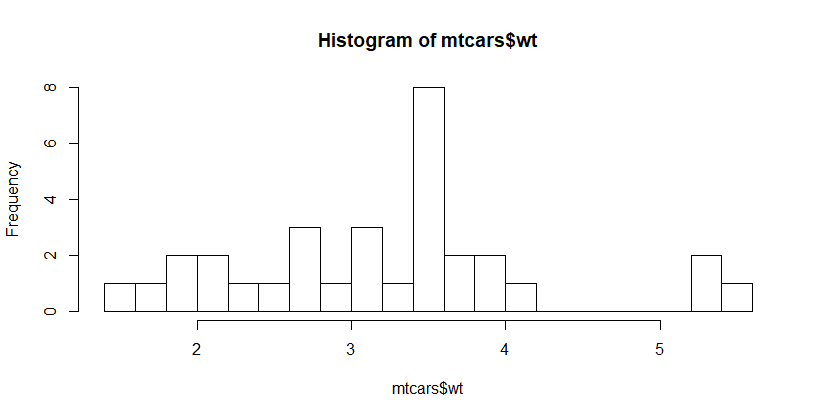
11 7 14

(4) barplot(table(mtcars$cyl))



(5)

hist(mtcars$wt, breaks = 15)



(6)

boxplot(mtcars$wt, horizontal = T)

q <- quantile(mtcars$wt)

Q1 <- q[2]

Q3 <- q[4]

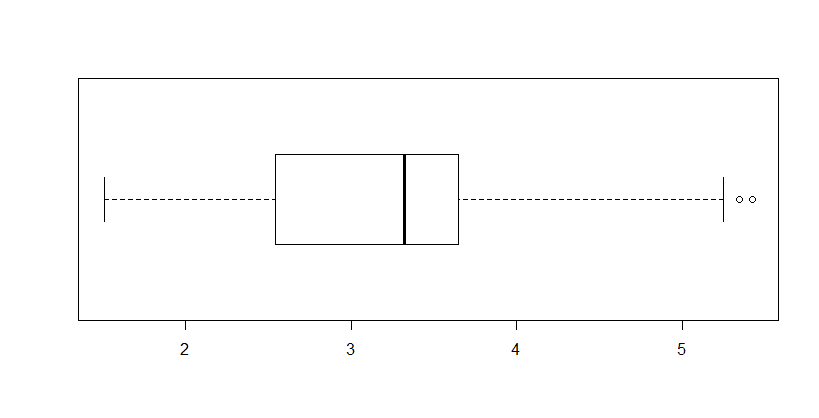
IQR <- Q3 - Q1

outIdx <- which(mtcars$wt > Q3 + 1.5\*IQR

| mtcars$wt < Q1 - 1.5\*IQR)

outlier <- mtcars$wt[outIdx]

summary(mtcars$wt)



> summary(mtcars$wt)

Min. 1st Qu. Median Mean 3rd Qu. Max.

1.513 2.581 3.325 3.217 3.610 5.424

> outlier

[1] 5.250 5.424 5.345

(7)

boxplot(mtcars$disp, horizontal = T)

q <- quantile(mtcars$disp)

Q1 <- q[2]

Q3 <- q[4]

IQR <- Q3 - Q1

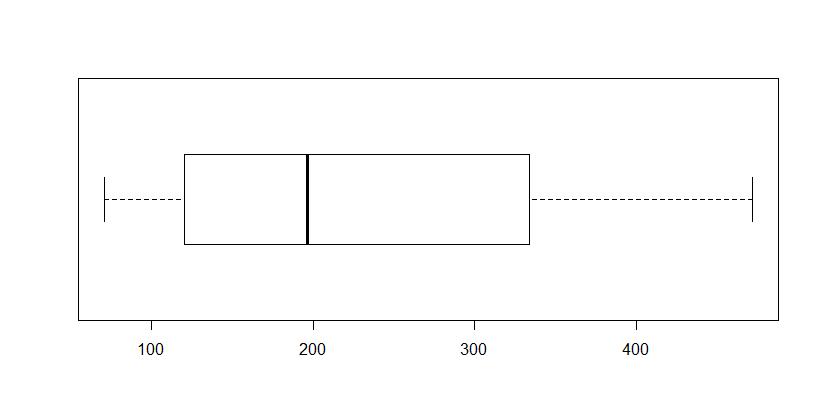
outIdx <- which(mtcars$disp > Q3 + 1.5\*IQR

| mtcars$disp < Q1 - 1.5\*IQR)

outlier <- mtcars$disp[outIdx]

summary(mtcars$disp)

outlier



> summary(mtcars$disp)

Min. 1st Qu. Median Mean 3rd Qu. Max.

71.1 120.8 196.3 230.7 326.0 472.0

> outlier

numeric(0)

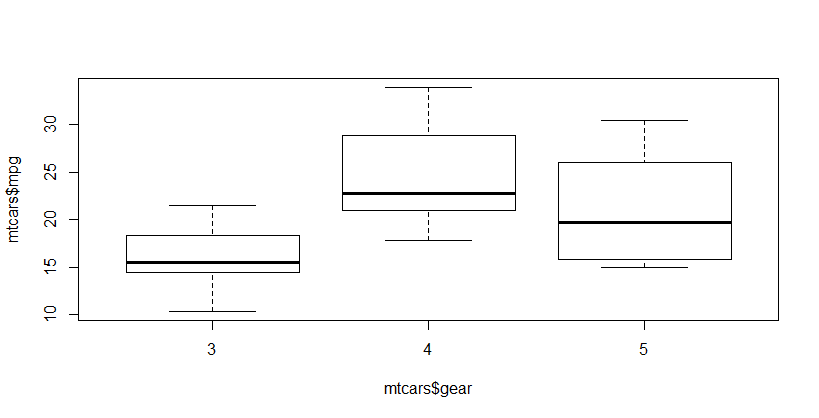
(8)

str(mtcars)

unique(mtcars$gear)

mtcars$gear <- factor(mtcars$gear, levels = c(3, 4, 5))

boxplot(mtcars$mpg ~ mtcars$gear)



> str(mtcars)

'data.frame': 32 obs. of 11 variables:

$ mpg : num 21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...

$ cyl : num 6 6 4 6 8 6 8 4 4 6 ...

$ disp: num 160 160 108 258 360 ...

$ hp : num 110 110 93 110 175 105 245 62 95 123 ...

$ drat: num 3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...

$ wt : num 2.62 2.88 2.32 3.21 3.44 ...

$ qsec: num 16.5 17 18.6 19.4 17 ...

$ vs : num 0 0 1 1 0 1 0 1 1 1 ...

$ am : num 1 1 1 0 0 0 0 0 0 0 ...

$ gear: Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...

$ carb: num 4 4 1 1 2 1 4 2 2 4 ...

> unique(mtcars$gear)

[1] 4 3 5

Levels: 3 4 5

> mtcars$gear <- factor(mtcars$gear, levels = c(3, 4, 5))

> boxplot(mtcars$mpg ~ mtcars$gear)

# 06.

(1)

data(trees)

head(trees)

> data(trees)

> head(trees)

Girth Height Volume

1 8.3 70 10.3

2 8.6 65 10.3

3 8.8 63 10.2

4 10.5 72 16.4

5 10.7 81 18.8

6 10.8 83 19.7

(2)

G <- trees$Girth

mean(G)

median(G)

mean(G, trim = 0.15)

sd(G)

> mean(G)

[1] 13.24839

> median(G)

[1] 12.9

> mean(G, trim = 0.15)

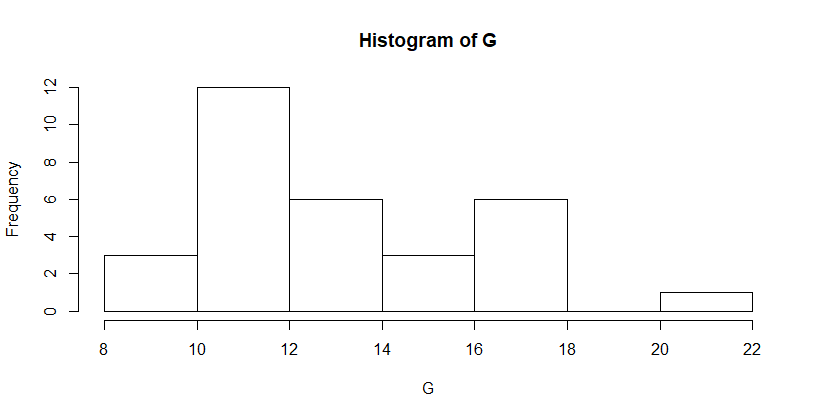
[1] 13.04348

> sd(G)

[1] 3.138139

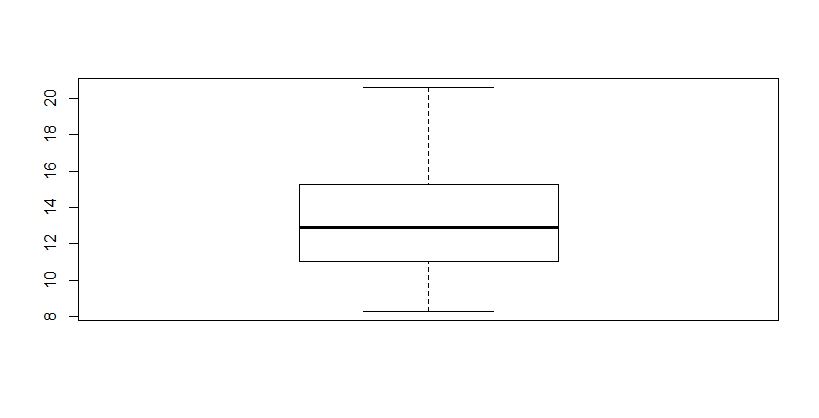
(3)

hist(G)



(4)

boxplot(G)



(5)

h <- trees$Height

mean(h)

median(h)

mean(h, trim = 0.15)

sd(h)

> h <- trees$Height

> mean(h)

[1] 76

> median(h)

[1] 76

> mean(h, trim = 0.15)

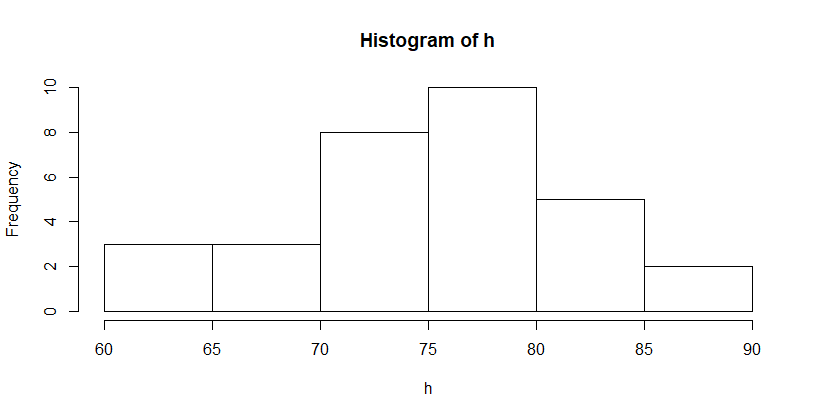
[1] 76.3913

> sd(h)

[1] 6.371813

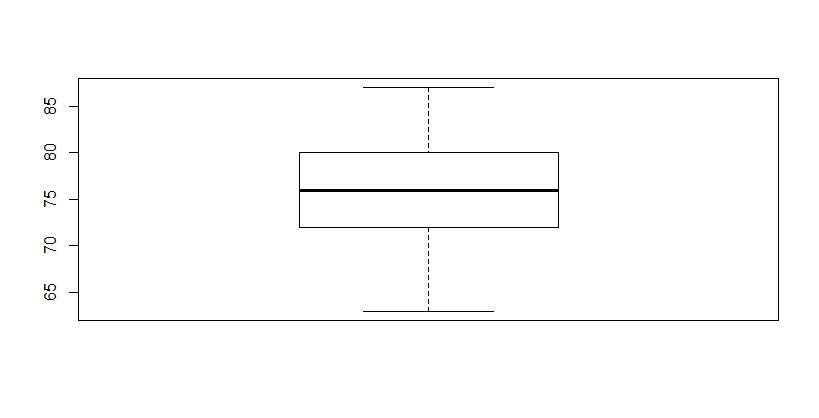
(6)

hist(h)



(7)

boxplot(h)



# 07.

(1)

data("Orange")

head(Orange)

> head(Orange)

Tree age circumference

1 1 118 30

2 1 484 58

3 1 664 87

4 1 1004 115

5 1 1231 120

6 1 1372 142

(2)

age <- Orange$age

mean(age)

median(age)

mean(age, trim = 0.15)

sd(age)

> age <- Orange$age

> mean(age)

[1] 922.1429

> median(age)

[1] 1004

> mean(age, trim = 0.15)

[1] 951

> sd(age)

[1] 491.8645

(3)

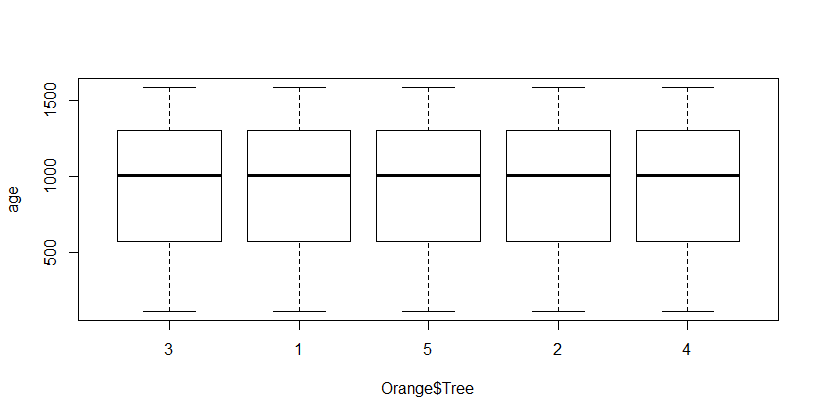
hist(age)



(4)

str(Orange)

boxplot(age ~ Orange$Tree)



(5)

str(Orange)

unique(Orange$Tree)

cir <- subset(Orange$circumference, Orange$Tree != 2)

mean(cir)

median(cir)

mean(cir, trim = 0.15)

sd(cir)

> str(Orange)

Classes ‘nfnGroupedData’, ‘nfGroupedData’, ‘groupedData’ and 'data.frame': 35 obs. of 3 variables:

$ Tree : Ord.factor w/ 5 levels "3"<"1"<"5"<"2"<..: 2 2 2 2 2 2 2 4 4 4 ...

$ age : num 118 484 664 1004 1231 ...

$ circumference: num 30 58 87 115 120 142 145 33 69 111 ...

- attr(\*, "formula")=Class 'formula' language circumference ~ age | Tree

.. ..- attr(\*, ".Environment")=<environment: R\_EmptyEnv>

- attr(\*, "labels")=List of 2

..$ x: chr "Time since December 31, 1968"

..$ y: chr "Trunk circumference"

- attr(\*, "units")=List of 2

..$ x: chr "(days)"

..$ y: chr "(mm)"

> unique(Orange$Tree)

[1] 1 2 3 4 5

Levels: 3 < 1 < 5 < 2 < 4

> cir <- subset(Orange$circumference, Orange$Tree != 2)

> mean(cir)

[1] 111

> median(cir)

[1] 115

> mean(cir, trim = 0.15)

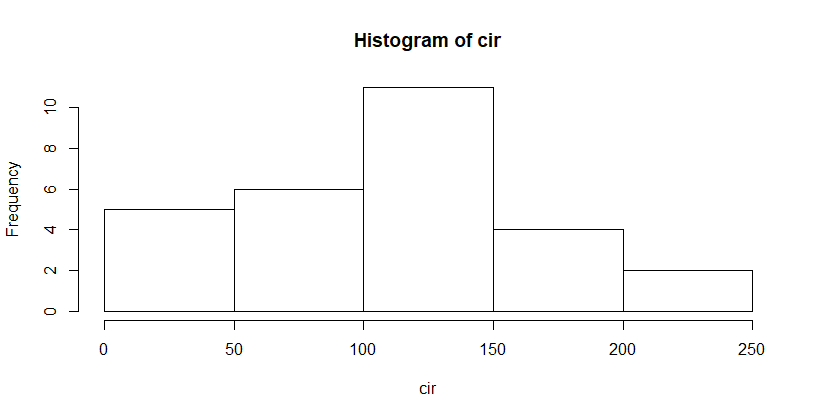
[1] 110.35

> sd(cir)

[1] 55.334

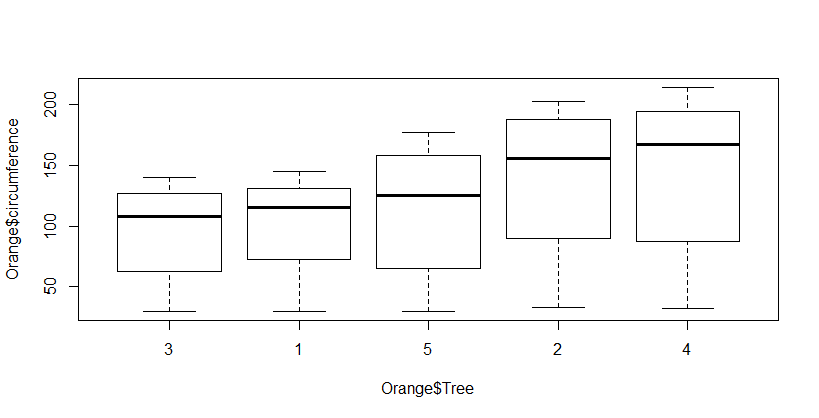
(6)

hist(cir)



(7)

boxplot(Orange$circumference ~ Orange$Tree)



Chapter 06

# 12.

(1)

> library(mlbench)

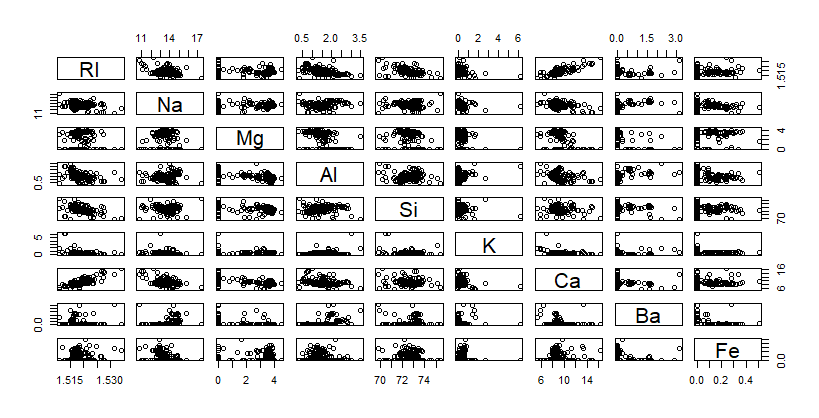
> data("Glass")

> myds <- Glass

(2)

str(myds)

pairs(myds[-10])



(3)

cor(myds[-10])

> cor(myds[-10])

RI Na Mg Al Si

RI 1.0000000000 -0.19188538 -0.122274039 -0.40732603 -0.54205220

Na -0.1918853790 1.00000000 -0.273731961 0.15679367 -0.06980881

Mg -0.1222740393 -0.27373196 1.000000000 -0.48179851 -0.16592672

Al -0.4073260341 0.15679367 -0.481798509 1.00000000 -0.00552372

Si -0.5420521997 -0.06980881 -0.165926723 -0.00552372 1.00000000

K -0.2898327111 -0.26608650 0.005395667 0.32595845 -0.19333085

Ca 0.8104026963 -0.27544249 -0.443750026 -0.25959201 -0.20873215

Ba -0.0003860189 0.32660288 -0.492262118 0.47940390 -0.10215131

Fe 0.1430096093 -0.24134641 0.083059529 -0.07440215 -0.09420073

K Ca Ba Fe

RI -0.289832711 0.8104027 -0.0003860189 0.143009609

Na -0.266086504 -0.2754425 0.3266028795 -0.241346411

Mg 0.005395667 -0.4437500 -0.4922621178 0.083059529

Al 0.325958446 -0.2595920 0.4794039017 -0.074402151

Si -0.193330854 -0.2087322 -0.1021513105 -0.094200731

K 1.000000000 -0.3178362 -0.0426180594 -0.007719049

Ca -0.317836155 1.0000000 -0.1128409671 0.124968219

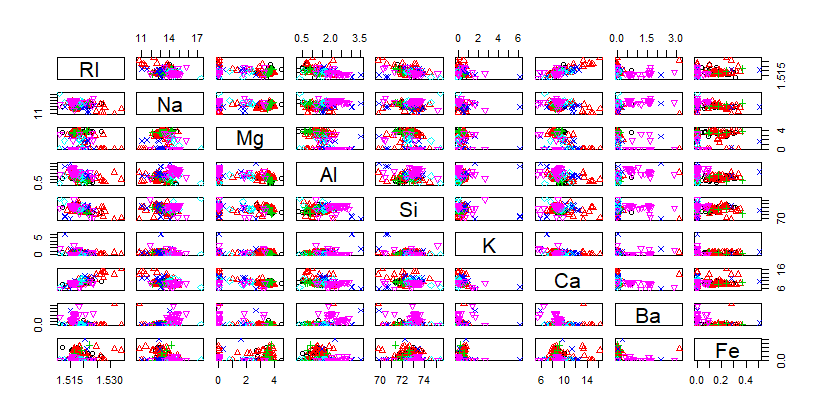
Ba -0.042618059 -0.1128410 1.0000000000 -0.058691755

Fe -0.007719049 0.1249682 -0.0586917554 1.000000000

(4)

point <- as.integer(myds$Type)

pairs(myds[-10], pch=point, col=point)



# 13.

(1)

> library(mlbench)

> data("Ionosphere")

> myds <- Ionosphere[,3:12]

(2)

(s <- sort(abs(cor(myds)), decreasing = T)[-(1:10)])

index <- which(cor(myds) == s[1] | cor(myds) == s[3])

index

# v5, v7

v5 <- myds[3]

v7 <- myds[5]

plot(v5, v7)

cor(v5, v7)

# v9, v11

v9 <- myds[7]

v11 <- myds[9]

plot(v9, v11)

cor(v9, v11)

> (s <- sort(abs(cor(myds)), decreasing = T)[-(1:10)])

[1] 0.670812770 0.670812770 0.597075076 0.597075076 0.476586950

[6] 0.476586950 0.471613756 0.471613756 0.461888764 0.461888764

[11] 0.450454390 0.450454390 0.449828915 0.449828915 0.441505091

[16] 0.441505091 0.440254375 0.440254375 0.429145526 0.429145526

[21] 0.412875920 0.412875920 0.373567050 0.373567050 0.364002810

[26] 0.364002810 0.337374391 0.337374391 0.336013342 0.336013342

[31] 0.325015504 0.325015504 0.315877278 0.315877278 0.302317088

[36] 0.302317088 0.291447328 0.291447328 0.274747389 0.274747389

[41] 0.255002962 0.255002962 0.251949534 0.251949534 0.217968180

[46] 0.217968180 0.207697324 0.207697324 0.200079919 0.200079919

[51] 0.190307607 0.190307607 0.190089786 0.190089786 0.169981091

[56] 0.169981091 0.167705319 0.167705319 0.163933239 0.163933239

[61] 0.150423664 0.150423664 0.143364804 0.143364804 0.120711999

[66] 0.120711999 0.090268334 0.090268334 0.054029528 0.054029528

[71] 0.047916265 0.047916265 0.042896223 0.042896223 0.038323122

[76] 0.038323122 0.034236002 0.034236002 0.029793759 0.029793759

[81] 0.025767514 0.025767514 0.020395361 0.020395361 0.010226919

[86] 0.010226919 0.008717135 0.008717135 0.001151850 0.001151850

> index <- which(cor(myds) == s[1] | cor(myds) == s[3])

> index

[1] 25 43 69 87

> # v5, v7

> v5 <- myds[3]

> v7 <- myds[5]

> cor(v5, v7)

V7

V5 0.5970751

> # v9, v11

> v9 <- myds[7]

> v11 <- myds[9]

> cor(v9, v11)

V11

V9 0.6708128

# 14.

data(LakeHuron)

str(LakeHuron)

years <- 1875:1972

hlevel <- as.vector(LakeHuron)

plot(years, hlevel, type = 'l', col='blue',

main='LakeHuron', ylab='수위')

> data(LakeHuron)

> str(LakeHuron)

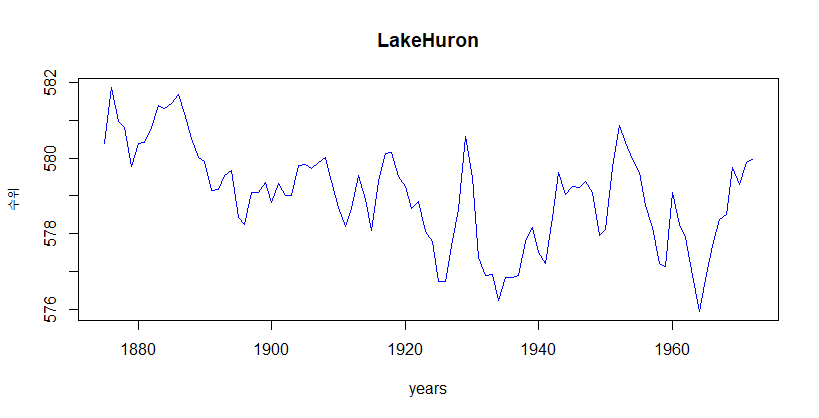
Time-Series [1:98] from 1875 to 1972: 580 582 581 581 580 ...

> years <- 1875:1972

> hlevel <- as.vector(LakeHuron)

> plot(years, hlevel, type = 'l', col='blue',

+ main='LakeHuron', ylab='수위')



# 16.

data("USAccDeaths")

str(USAccDeaths)

mon3 <- USAccDeaths[1:12]

mon5 <- USAccDeaths[25:36]

mon7 <- USAccDeaths[49:60]

month <- 1:12

plot(month, mon3, type = 'l', col='red', ylim = c(6000, 13000))

axis(side = 1, at = 1:12)

lines(month, mon5, type = 'l', col='blue', ylim = c(6000,13000))

lines(month, mon7, type = 'l', col='violet', ylim = c(6000,13000))

> data("USAccDeaths")

> str(USAccDeaths)

Time-Series [1:72] from 1973 to 1979: 9007 8106 8928 9137 10017 ...

> mon3 <- USAccDeaths[1:12]

> mon5 <- USAccDeaths[25:36]

> mon7 <- USAccDeaths[49:60]

> month <- 1:12

> plot(month, mon3, type = 'l', col='red', ylim = c(6000, 13000))

> axis(side = 1, at = 1:12)

> lines(month, mon5, type = 'l', col='blue', ylim = c(6000,13000))

> lines(month, mon7, type = 'l', col='violet', ylim = c(6000,13000))

