***Opdrachten bij de dataset airquality***

Informatie over de dataset airquality

Airquality{datasets}

A data frame with 154 observations on 6 variables:

1. Ozone numeric ozone (ppb)
2. Solar.R cumeric Solar R(lang)
3. Wind numberic Wind(mph)
4. Temp numeric Temperature (degrees F)
5. Month numberic Month (1-12)
6. Day numeric Day

Airquality$Wind

Attach(airquality) 🡪 namen van kolommen zal apart laten zien.

Attach(ariquality)

Cor.test(Wind.Ozone)

Of

Cor.test(airquality$Wind, airquality$Ozone)

Selecteer de subset, waarvoor de temperatuur onder de 70F en laat alleen rijnummer, Ozone en Temp zien.

**subset (airquality, Temp<70, select = c(Ozone, Temp))**

Selecteer de subset, waarvoor Wind>12,0 en laat alleen rijnummer, Ozone en Wind zien.

**Subset(airquality, Wind>12, select = c(Ozone, Wind))**

Selecteer de subset, waarvoor Day=1 en laat alleen rijnummer, Ozone en Temp en Day zien.

**Subset(airquality, Day=1, select = c(Ozone, Temp, Day**))

Selecteer de subset, waarvoor Day=1 en laat de kolom Temp weg.

**subset(airquality, Day==1, select= -Temp)**

Selecteer de kolommen Ozone tot en met Wind

**Subset(airquality, select = Ozone:Wind)**

Selecteer regel 8 tot en met 14 (week 2) uit het dataframe

**subset(airquality[8:14,])**

Selecteer de eerste tien zondagen uit het dataframe

**Subset(airquality[c(6,13,20,27,34,41,48,55,62,69)])**

**of**

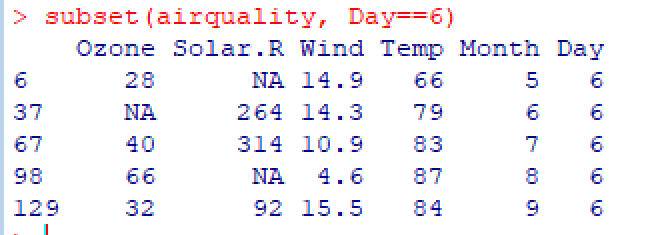
**Subset(airquality[seq(6,nrow(airquality),7),])**

Selecteer alle dagen uit het dataframe waarop Ozone de waarde Na heeft

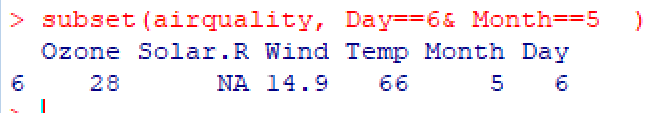
**Subset(airquality,is.na(Ozone))**

Laat het weer in NewYork zien in 19… op elke 6e dag van de maand.

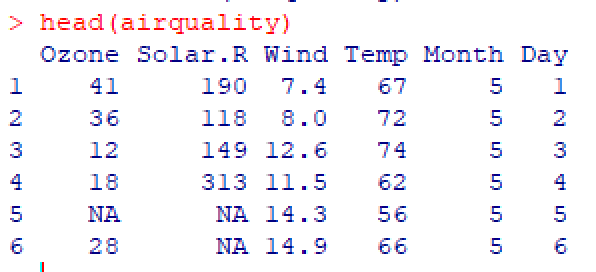
Subset(airquality, Day==6)



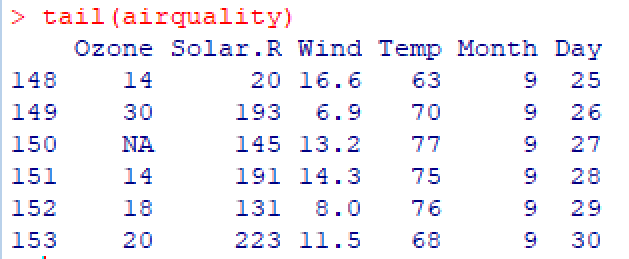
Laat het weer in NewYork zien in 19.. op de 6e dag van de 5e maand.



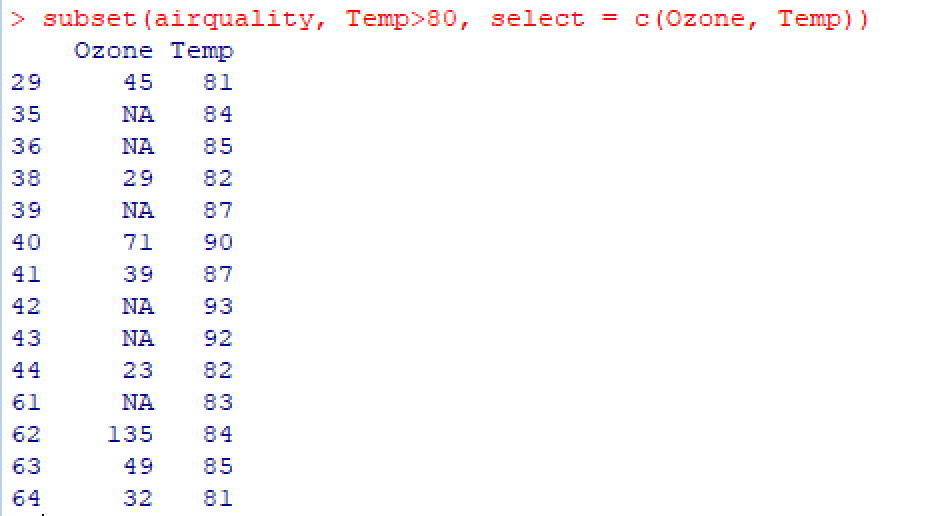
Head(airquality)



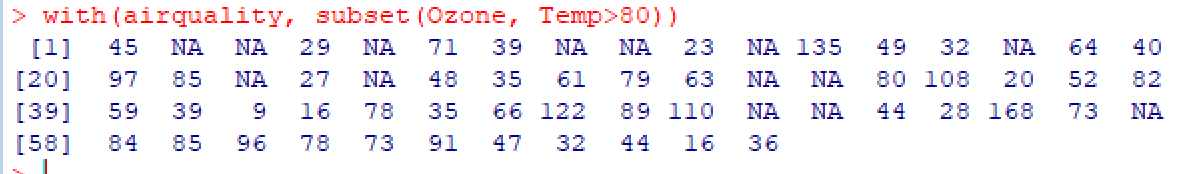
Tail(airquality)



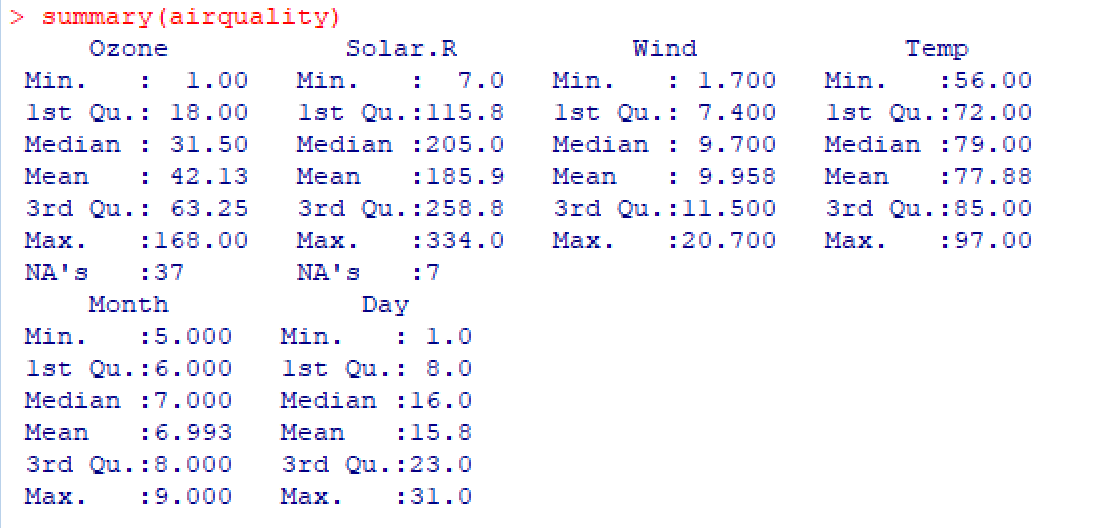
Subset(airquality, Temp>80, select = c(Ozone, Temp))



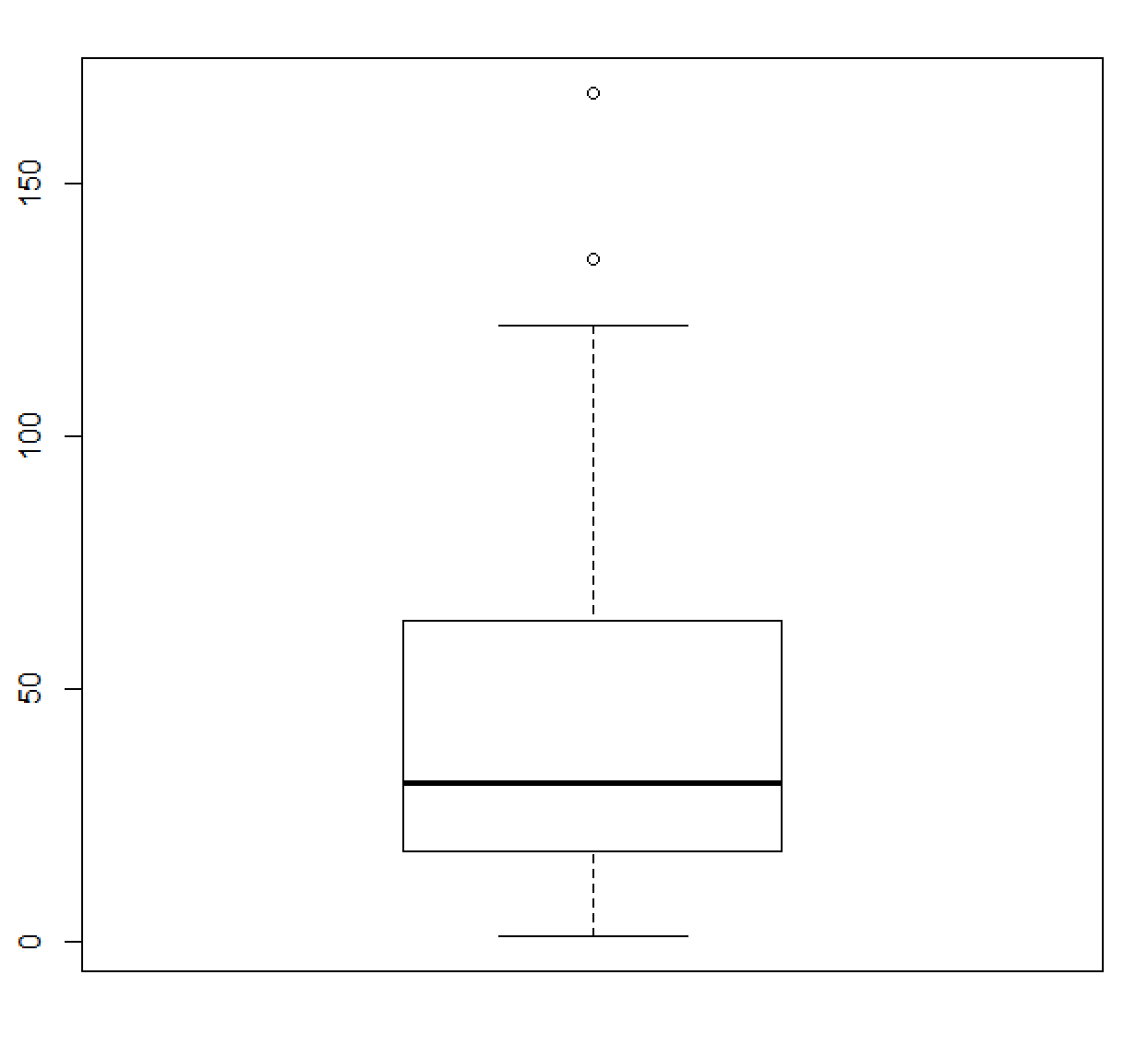
With(airquality, subset(Ozone, Temp>80))



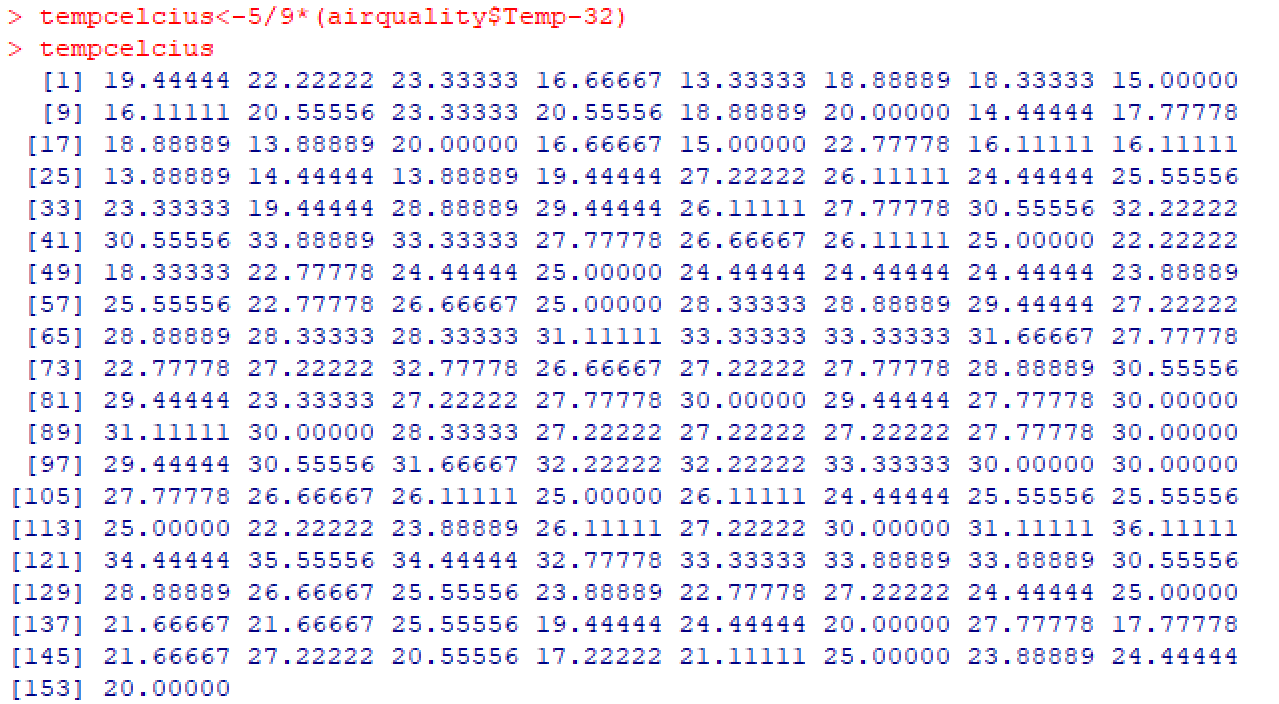
Summary(airquality)



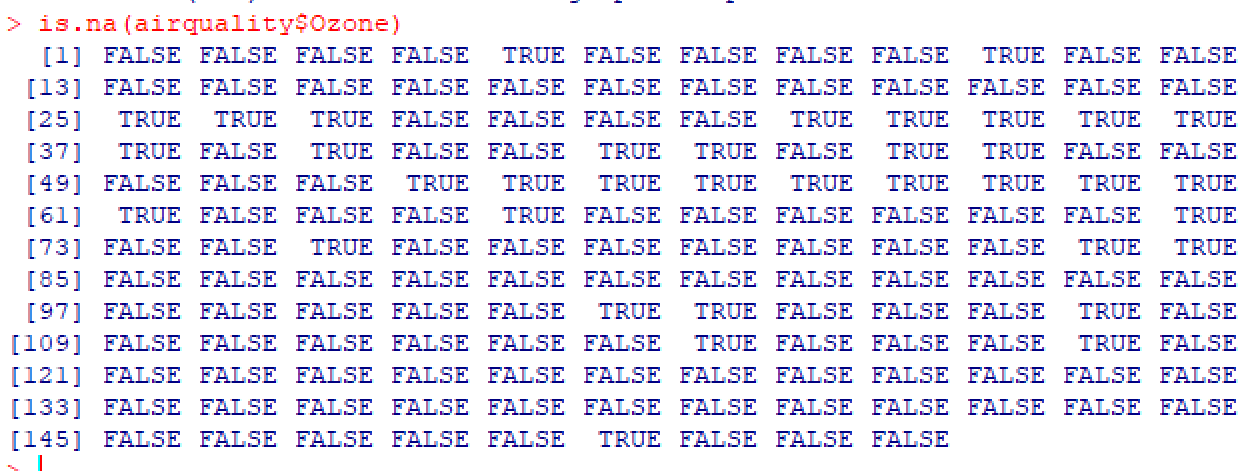
Boxplot(airquality$Ozone)



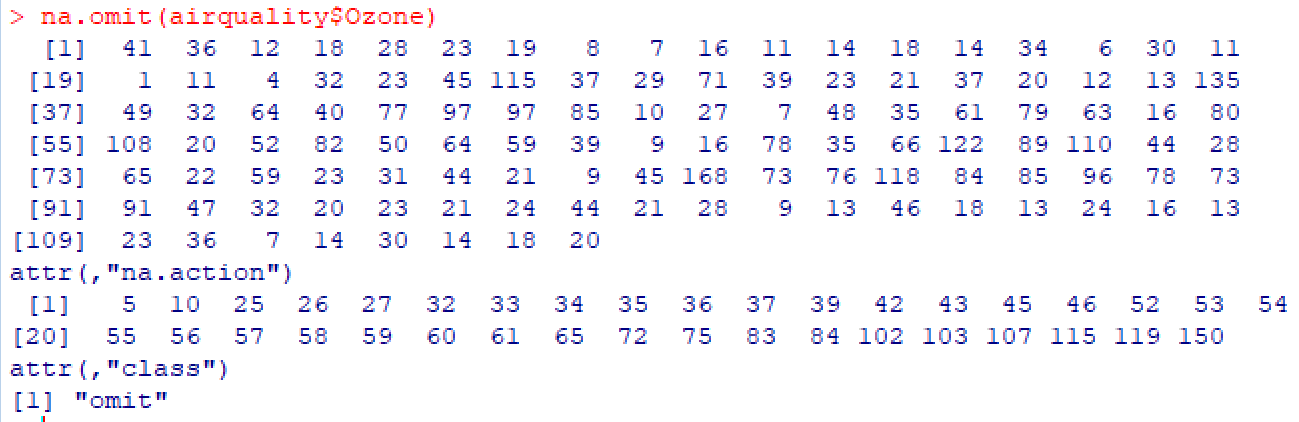
Tempcelcius<-5/9\*(airquality$Temp-32)



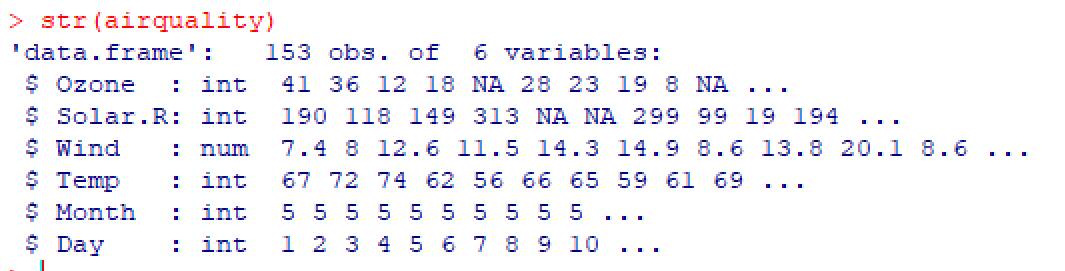
Is.na(airquality$Ozone)



Na.omit(airquality$Ozone)



Str(airquality)



Library(help=’stats’)

***?binomial***

dbinom(x, size, prob, log =FALSE)

dbinom(x=3, size=20, prob=1/6) = 0,2378866

dbinom(0:4,4,1/2)

sum(dbinom(0:4,4,1/2))

plot(dbinom(0:4,4,1/2),col="red")

> plot(dbinom(0:4,4,1/2),col="red", Xlim = 0:4)

pbinom(q, size, prob, lower.tail = TRUE, log.p=FALSE)

dbinom(3,20,1/6) = 0,2378866

qbinom(p,size, prob, lower.tail=TRUE, log.p=FALSE)

rbinom(n, size, prob)

***Binomial test example 1***

# Under (the assumption of) simple Mendelian inheritance, a cross between plants of two particular genotypes produces progeny 1/4 of which are "dwarf" and 3/4 of which are "giant“.

In an experiment to determine if this assumption is reasonable, a cross results in progeny

having 243 dwarf and 682 giant plants. If "giant" is taken as success, the null hypothesis is that

p = 3/4 and the alternative that p ≠ 3/4.

**binom.test(682,682+243, p = 3/4 )**

# or

**binom.test(c(682, 243), p = 3/4)**

# x number of successes, or a vector of length 2 giving the numbers of successes and failures, respectively. n number of trials; ignored if x has length 2.

**binom.test(c(682, 243), p = 3/4)** # Output:

Exact binomial test

data: c(682, 243)

number of successes = 682, number of trials = 925,

**p-value = 0.3825**

alternative hypothesis: true probability of success is not equal to 0.75

95 percent confidence interval: 0.7076683 0.7654066

sample estimates: probability of success 0.7372973

***?chisq.test***

chisq.test(x, y = NULL, correct = TRUE, p = rep(1/length(x), length(x)), + rescale.p = FALSE, imulate.p.value = FALSE, B = 2000)

**Arguments**

x a numeric vector or matrix. x and y can also both be factors.

y a numeric vector; ignored if x is a matrix. If x is a factor, y should be

a factor of the same length.

***Chisq.test example 1  
# goodness of fit***

**x <- c(A = 20, B = 15, C = 25)** # hier staan de geobserveerde waarden

**chisq.test(x)**

Chi-squared test for given probabilities

data: x X-squared = 2.5, df = 2, p-value = 0.2865

***Chisq.test example 2***

***#goodness of fit***

**x <- c(89,37,30,28,2)** # hier staan de geobserveerde waarden

**p <- c(0.30,0.20,0.20,0.19,0.11)** # is samen 1

# hier staan de verwachte verhoudingen

**chisq.test(x, p = p)**

Chi-squared test for given probabilities

data: x

X-squared = 39.328, df = 4, p-value = 5.96e-08

***Chisq.test example 3***

***#goodness of fit***

**x <- c(89,37,30,28,2)**

**p <- c(40,20,20,15,5)** # is samen 100 (%)

**chisq.test(x, p = p, rescale.p = TRUE)**

Chi-squared test for given probabilities

data: x

X-squared = 9.9901, df = 4, p-value = 0.04059