

Opdracht_3.Rmd

Lisa Hu

12-5-2021

Opdracht 2

Om de working directory te veranderen kan de volgende command worden gebruikt door de path mee te geven als argument.

```
setwd()
```

Opdracht 3

```
read_table <- read.table(file="Week3_opdracht2.csv", sep=",")
read_table
```

##	V1	V2	V3	V4	V5
## 1	id	naam	leeftijd	woonplaats	Inkomen
## 2	1	Harrie	18	Groningen	16344
## 3	2	Klaas	23	Drachten	17000
## 4	3	Kees	44	Leeuwarden	45899
## 5	4	Shanita	36	Assen	34500
## 6	5	Charlotta	23	Emmen	23456
## 7	6	Nanine	21	Groningen	12876
## 8	7	Lema	67	Groningen	34899
## 9	8	Jan	33	Drachten	66799

Opdracht 4

Factor data slaat nominale data op als vector en geeft het een telbare waarde mee.

Opdracht 5

```
read_csv <- read.csv("Week3_opdracht2.csv")
read_csv
```

##	id	naam	leeftijd	woonplaats	Inkomen
## 1	1	Harrie	18	Groningen	16344
## 2	2	Klaas	23	Drachten	17000
## 3	3	Kees	44	Leeuwarden	45899
## 4	4	Shanita	36	Assen	34500
## 5	5	Charlotta	23	Emmen	23456
## 6	6	Nanine	21	Groningen	12876
## 7	7	Lema	67	Groningen	34899
## 8	8	Jan	33	Drachten	66799

Opdracht 6

```
read_2csv <- read.csv("Week3_opdracht6(1).csv", na.strings=c(""))
read_2csv
```

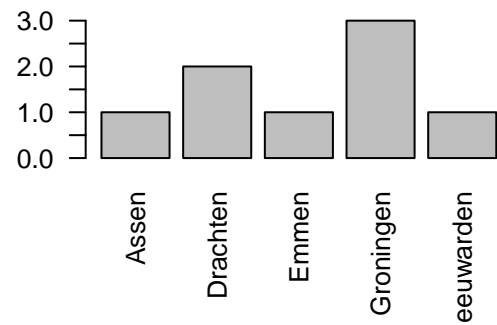
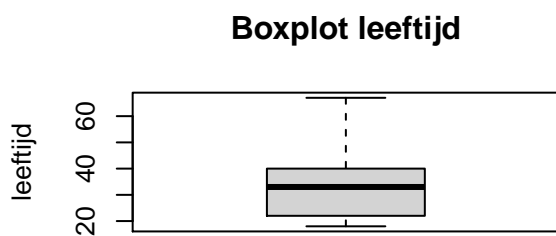
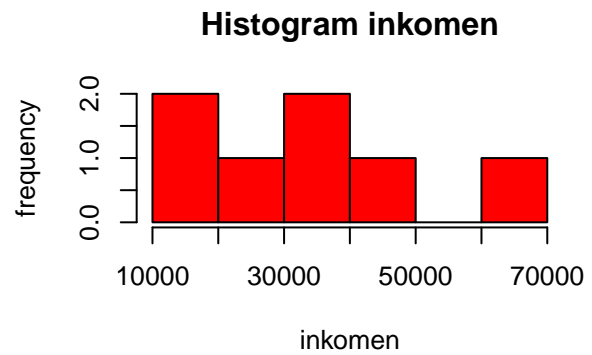
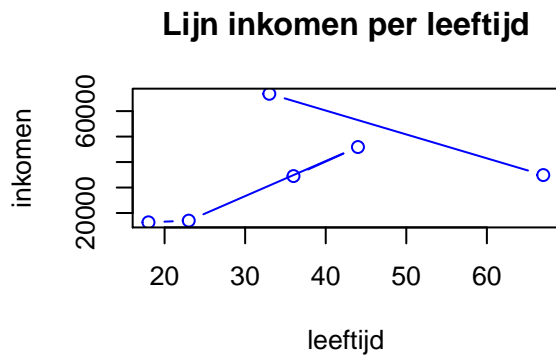
##	id	naam	leeftijd	woonplaats	Inkomen
## 1	1	Harrie	18	Groningen	16344
## 2	2	Klaas	23	Drachten	17000
## 3	3	<NA>	44	Leeuwarden	45899
## 4	4	Shanita	36	Assen	34500
## 5	5	Charlotta	NA	Emmen	23456
## 6	6	Nanine	21	Groningen	NA
## 7	7	Lema	67	Groningen	34899
## 8	8	Jan	33	Drachten	66799

De data van deze file is incompleet:

- Nummer 3 mist zijn naam;
- Nummer 5 mist haar leeftijd;
- Nummer 6 mist haar inkomen.

Opdracht 7

```
read_2csv <- read.csv("Week3_opdracht6(1).csv", na.strings=c(""))
par(mfrow=c(2,2))
plot(x=read_2csv$leeftijd, y=read_2csv$Inkomen, type="b", col="blue1",
     main="Lijn inkomen per leeftijd", xlab="leeftijd", ylab="inkomen")
hist(read_2csv$Inkomen, col="red",
     main="Histogram inkomen", xlab="inkomen", ylab="frequency")
boxplot(read_2csv$leeftijd, main="Boxplot leeftijd", ylab="leeftijd")
plot(factor(read_2csv$woonplaats), las=2)
```



Opdracht 8

8 obs. zijn de rijen. 5 variables zijn de kolommen.

Opdracht 9

```
min(read_2csv$leeftijd, na.rm=TRUE)
```

```
## [1] 18
```

```
max(read_2csv$leeftijd, na.rm=TRUE)
```

```
## [1] 67
```

```
mean(read_2csv$leeftijd, na.rm=TRUE)
```

```
## [1] 34.57143
```

```
min(read_2csv$Inkomen, na.rm=TRUE)
```

```
## [1] 16344
```

```
max(read_2csv$Inkomen, na.rm=TRUE)
```

```
## [1] 66799
```

```
mean(read_2csv$Inkomen, na.rm=TRUE)
```

```
## [1] 34128.14
```

Opdracht 10

```
summary(read_2csv)
```

```
##           id           naam           leeftijd           woonplaats
##  Min.      :1.00   Length:8      Min.      :18.00   Length:8
##  1st Qu.:2.75   Class :character  1st Qu.:22.00   Class :character
##  Median :4.50   Mode  :character  Median :33.00   Mode  :character
##  Mean    :4.50           Mean    :34.57
##  3rd Qu.:6.25           3rd Qu.:40.00
##  Max.    :8.00           Max.    :67.00
##                                     NA's    :1
##           Inkomen
##  Min.      :16344
##  1st Qu.:20228
##  Median :34500
##  Mean    :34128
##  3rd Qu.:40399
##  Max.    :66799
##  NA's    :1
```