

BIOINFORMATICA

STATISTIEK IN R_

Week 3

ONDERWERPEN

Grafieken

- Plot
- Barplot
- Histogram

Plaatje vs. Grafiek

GRAFIEKEN

Grafieken maken met R kan erg gemakkelijk

Veel verschillende soorten ingebouwde grafieken in de package Graphics.

```
?library(help="graphics")
```

Vandaag gaan we werken met:

- Plot
- Barplot
- Histogram

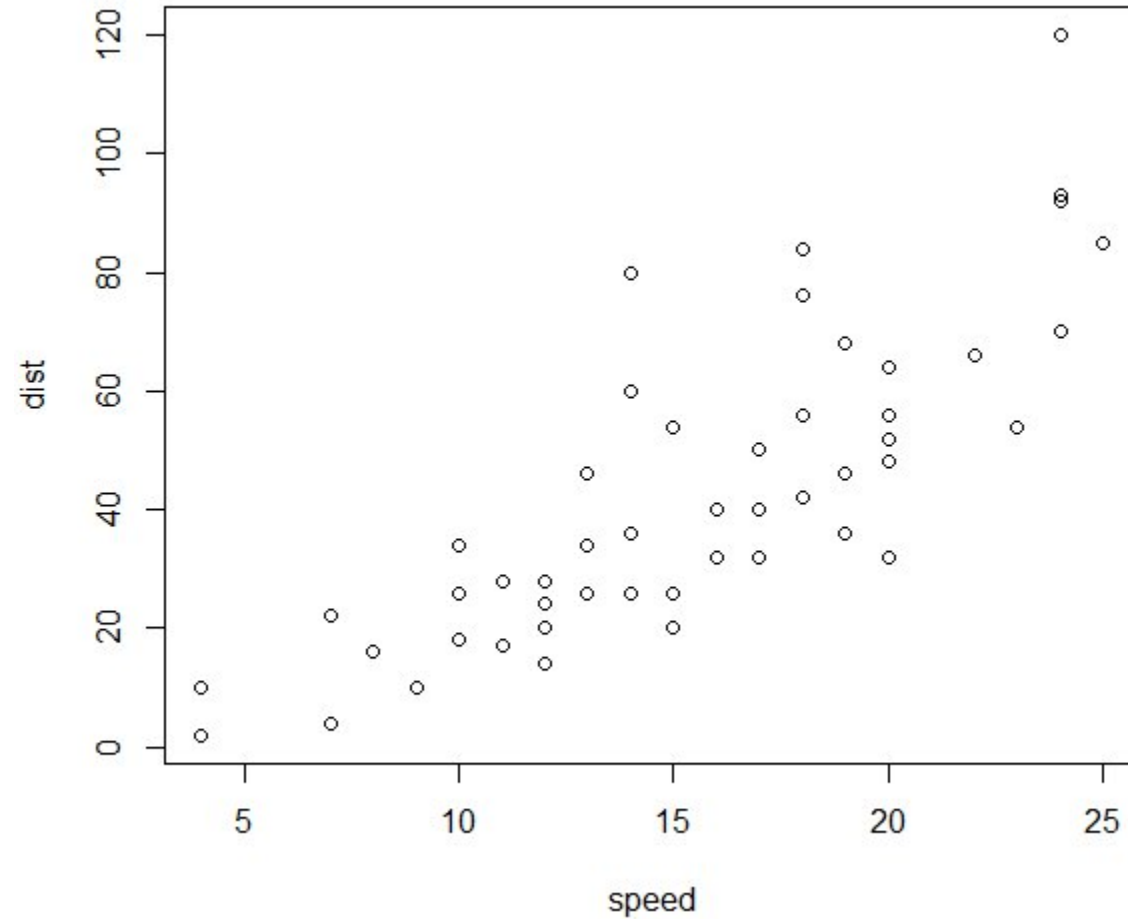
Volgende week gaan we werken met boxplots.

PLOT

Het gebruik van de functie plot is erg gemakkelijk:

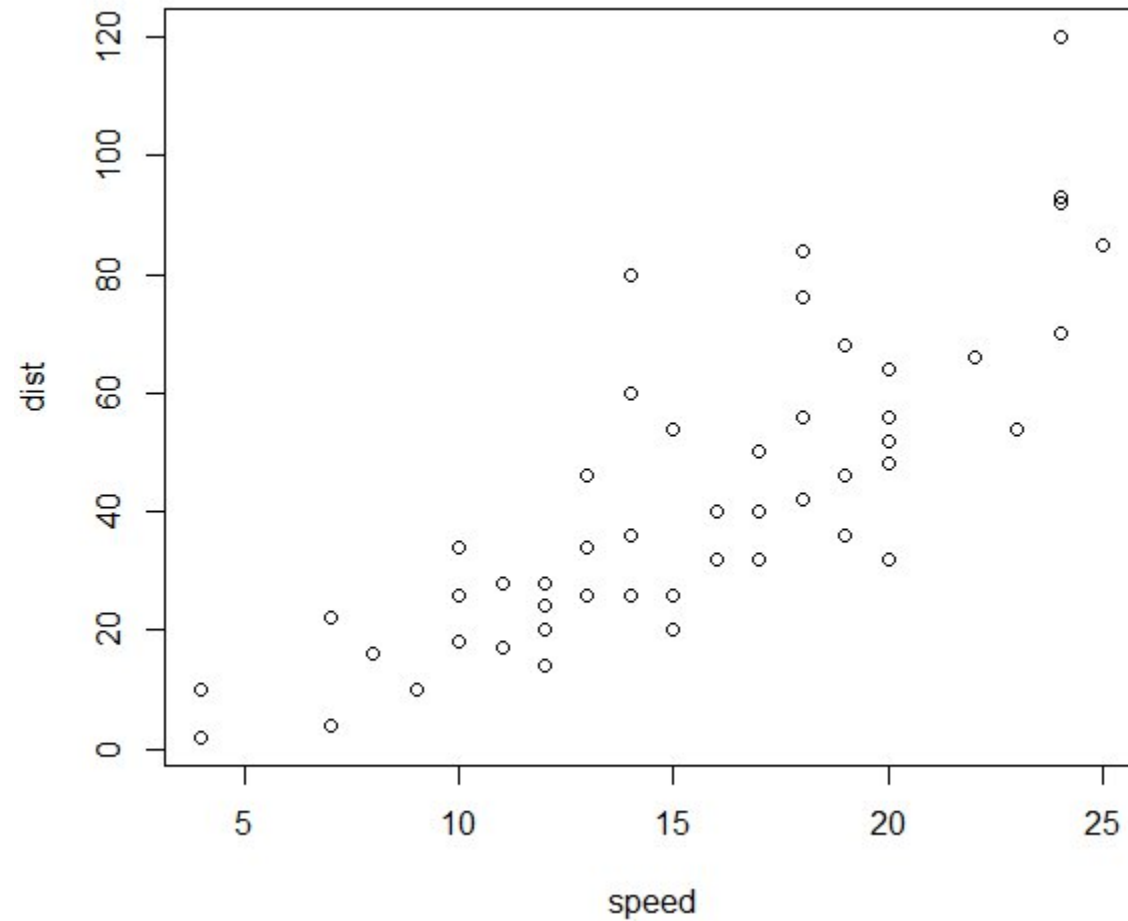
`example(plot)`

`plot(cars)`



PLOT(CARS)

Maar wat staat er in dit plaatje?



PLAATJE VS. GRAFIEK

?plot

?cars

cars {datasets}

R Documentation

Speed and Stopping Distances of Cars

Description

The data give the speed of cars and the distances taken to stop. Note that the data were recorded in the 1920s.

Usage

```
cars
```

Format

A data frame with 50 observations on 2 variables.

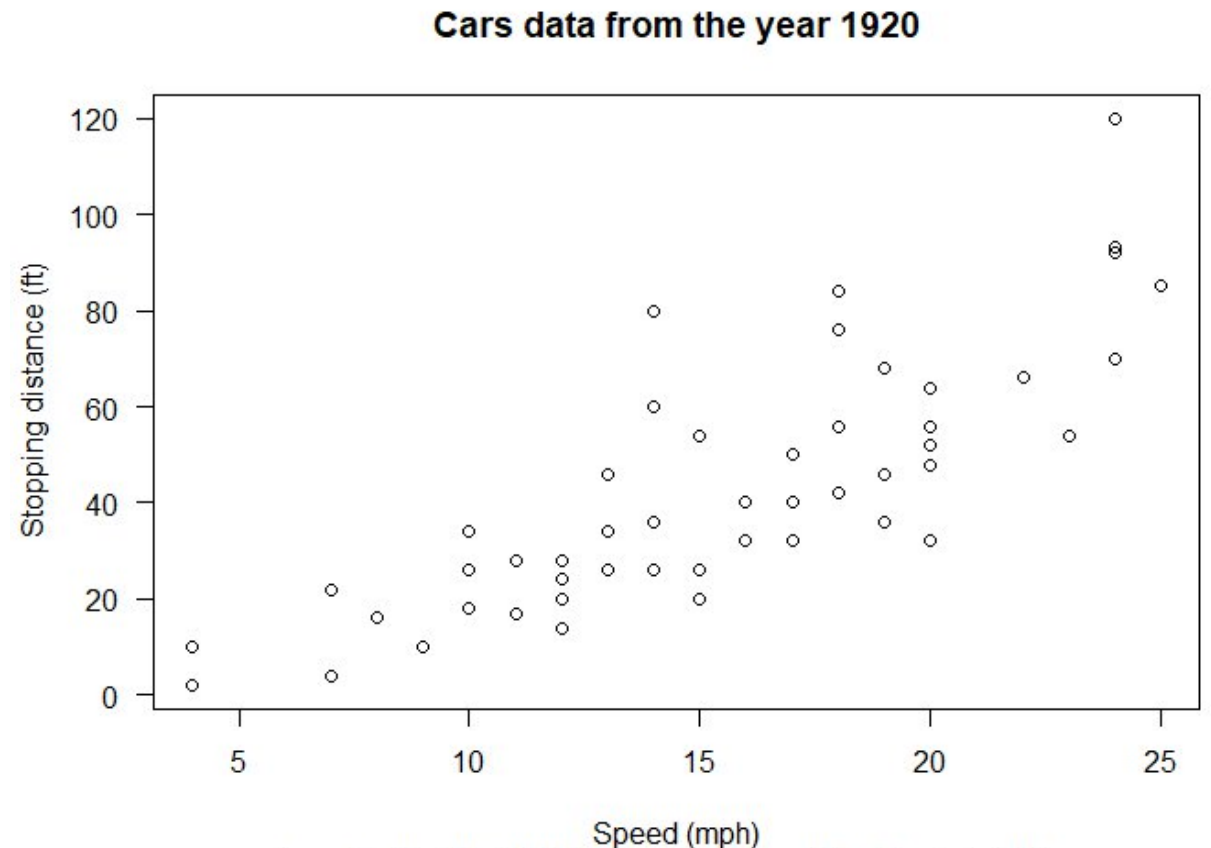
[,1] speed numeric Speed (mph)

[,2] dist numeric Stopping distance (ft)

PLAATJE VS. GRAFIEK

#Code voor grafiek, met alle extra informatie

```
plot(cars,  
     xlab = "Speed (mph)",  
     ylab = "Stopping distance (ft)",  
     sub = "bron: Ezekiel, M. (1930) Methods  
of Correlation Analysis. Wiley.",  
     main = "Cars data from the year 1920",  
     las=1)
```



bron: Ezekiel, M. (1930) Methods of Correlation Analysis. Wiley.

PLAATJE VS. GRAFIEK

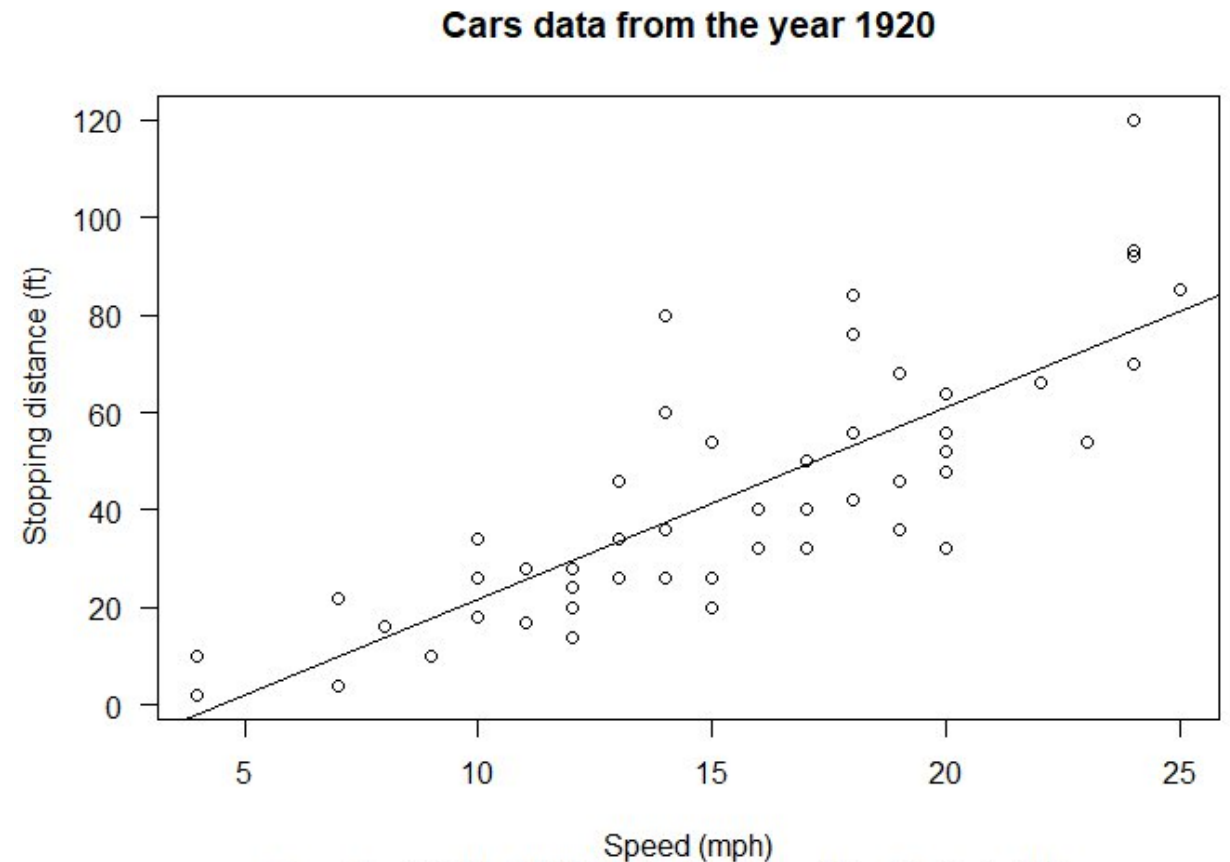
#Code voor grafiek, met alle extra informatie

```
plot(cars,  
     xlab = "Speed (mph)",  
     ylab = "Stopping distance (ft)",  
     sub = "bron: Ezekiel, M. (1930) Methods  
of Correlation Analysis. Wiley.",  
     main = "Cars data from the year 1920",  
     las=1)
```

#Toevoegen van lineaire trendlijn:

```
z <- lm(dist ~ speed, data = cars)
```

```
abline(z)
```



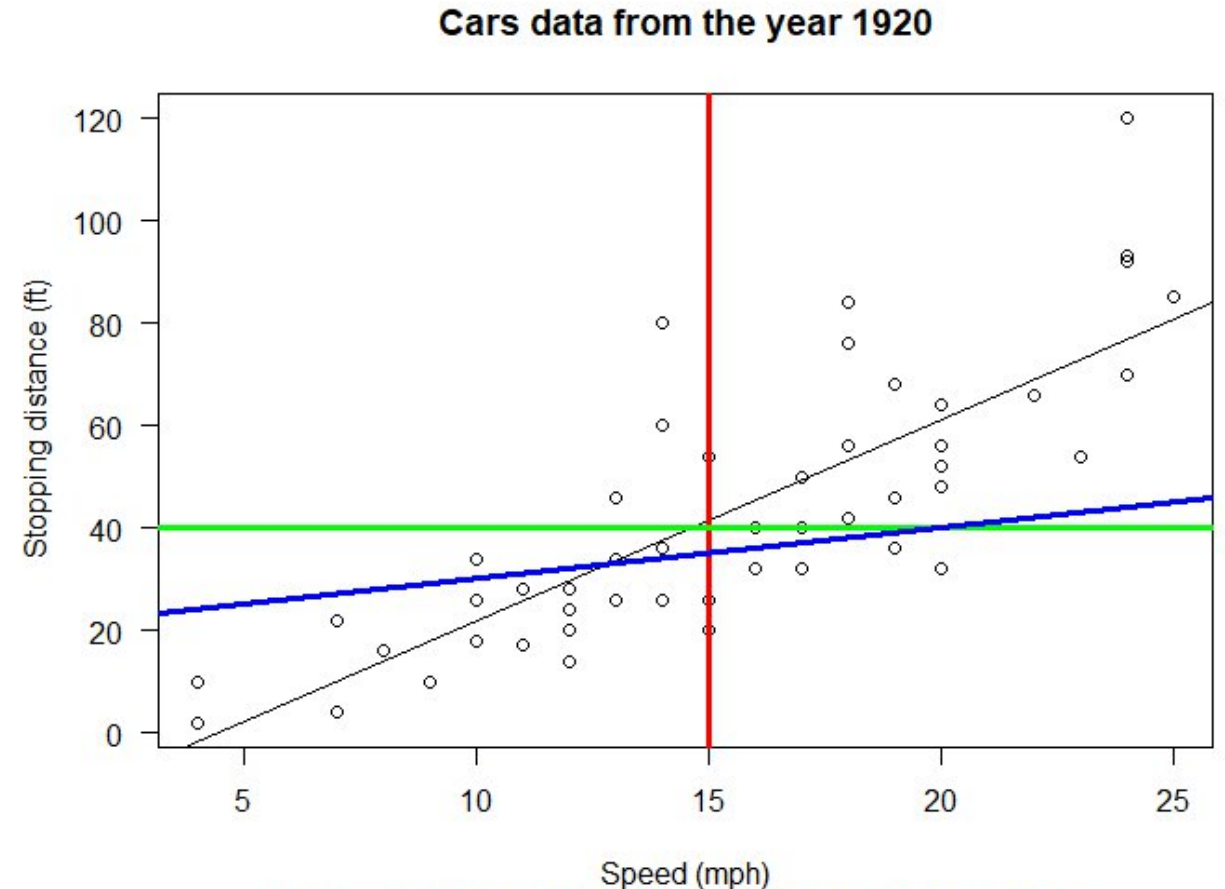
bron: Ezekiel, M. (1930) Methods of Correlation Analysis. Wiley.

?ABLINE

```
abline(v = 15, col = "red", lwd = 3)
```

```
abline(h = 40, col = "green", lwd = 3)
```

```
abline(20,1, col = "blue", lwd = 3)
```



bron: Ezekiel, M. (1930) Methods of Correlation Analysis. Wiley.

BARPLOT

?barplot

```
barplot(height, ...)
```

```
## Default S3 method:
```

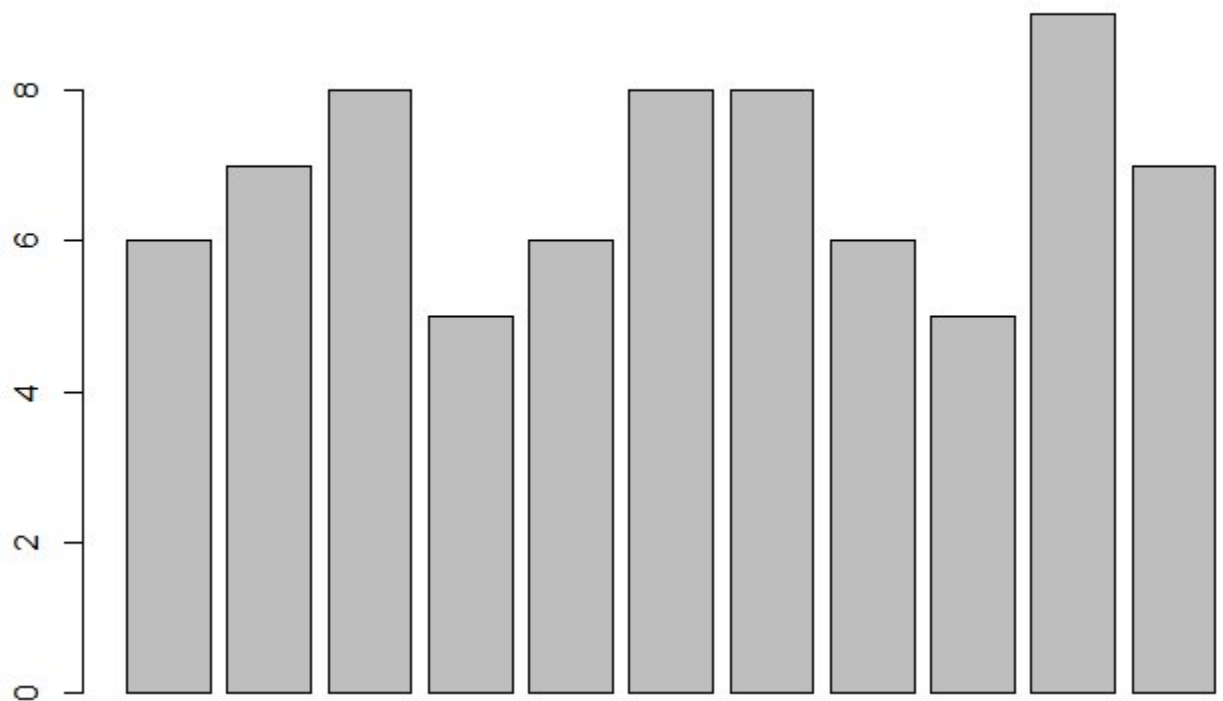
```
barplot(height, width = 1, space = NULL,  
        names.arg = NULL, legend.text = NULL, beside = FALSE,  
        horiz = FALSE, density = NULL, angle = 45,  
        col = NULL, border = par("fg"),  
        main = NULL, sub = NULL, xlab = NULL, ylab = NULL,  
        xlim = NULL, ylim = NULL, xpd = TRUE, log = "",  
        axes = TRUE, axisnames = TRUE,  
        cex.axis = par("cex.axis"), cex.names = par("cex.axis"),  
        inside = TRUE, plot = TRUE, axis.lty = 0, offset = 0,  
        add = FALSE, args.legend = NULL, ...)
```

BARPLOT

Maak een barplot van cijfers

```
cijfers <- c(6,7,8,5,6,8,8,6,5,9,7)
```

```
barplot(cijfers)
```

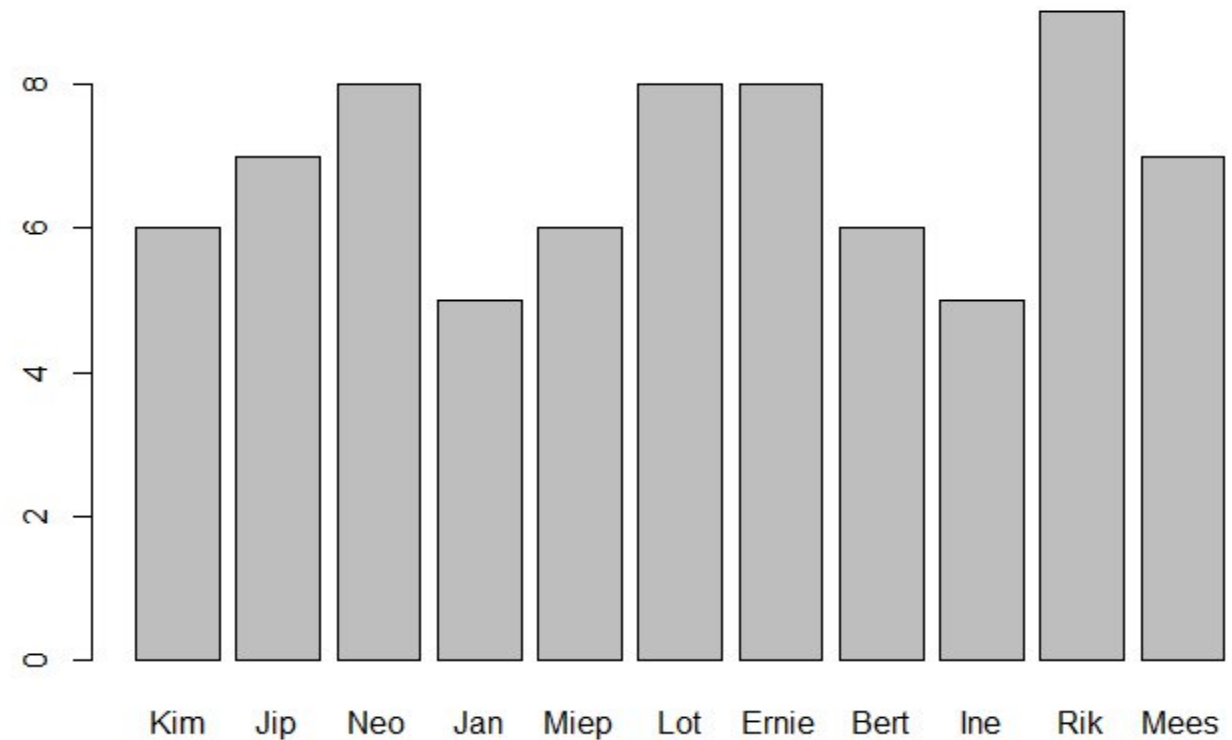


BARPLOT

Van wie is elk cijfer?

```
names(cijfers) <- c('Kim', 'Jip', 'Neo', 'Jan', 'Miep', 'Lot', 'Ernie', 'Bert', 'Ine', 'Rik', 'Mees')
```

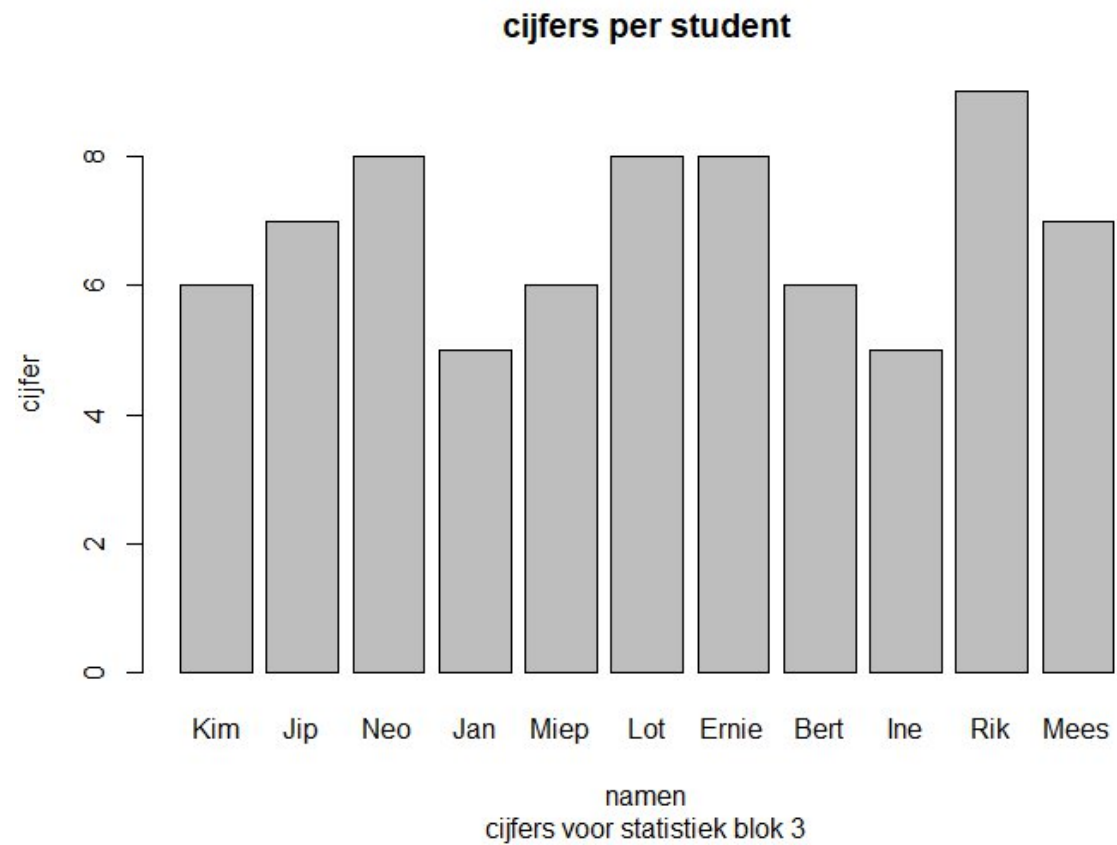
```
barplot(cijfers)
```



BARPLOT

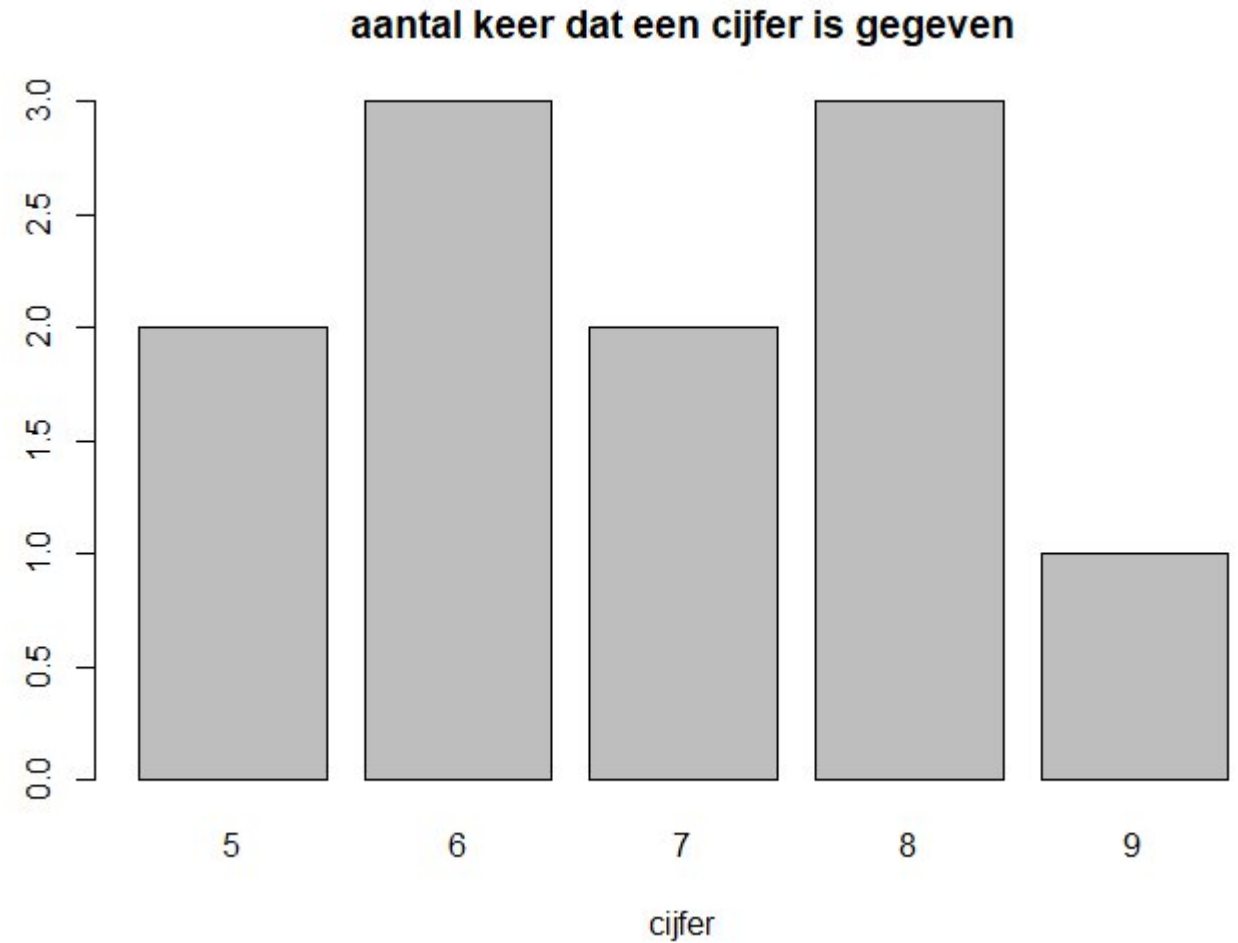
#maak een grafiek van het plaatje, voeg betekenis toe

```
barplot(cijfers, main = "cijfers per student",  
        xlab = "namen",  
        ylab = "cijfer")
```



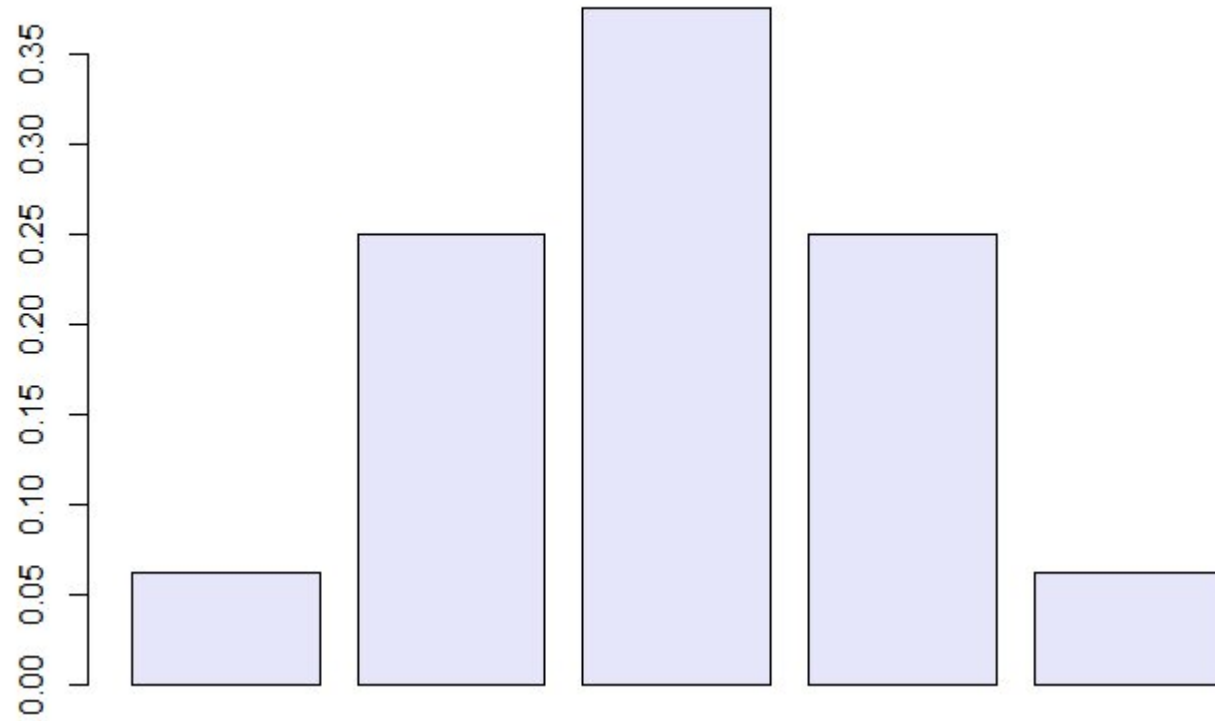
BARPLOT

`barplot(table(cijfers), main = "aantal keer dat een cijfer is gegeven")`



BARPLOT

```
barplot(dbinom(0:4,4,1/2),col="lavender")
```



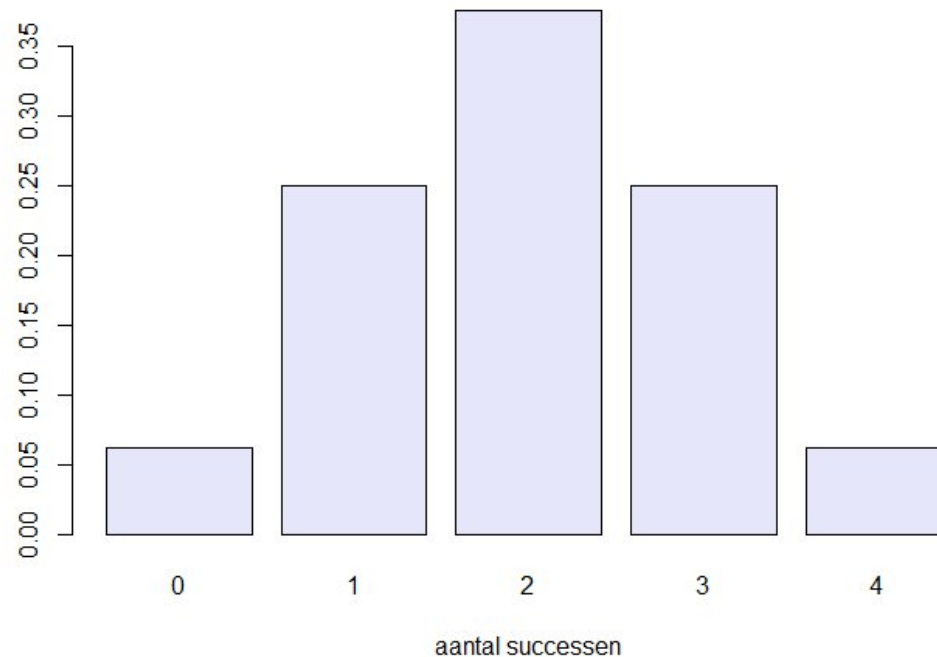
BARPLOT

Van je plaatje een grafiek maken:

```
binomialeverdeling<-dbinom(0:4,4,1/2)
```

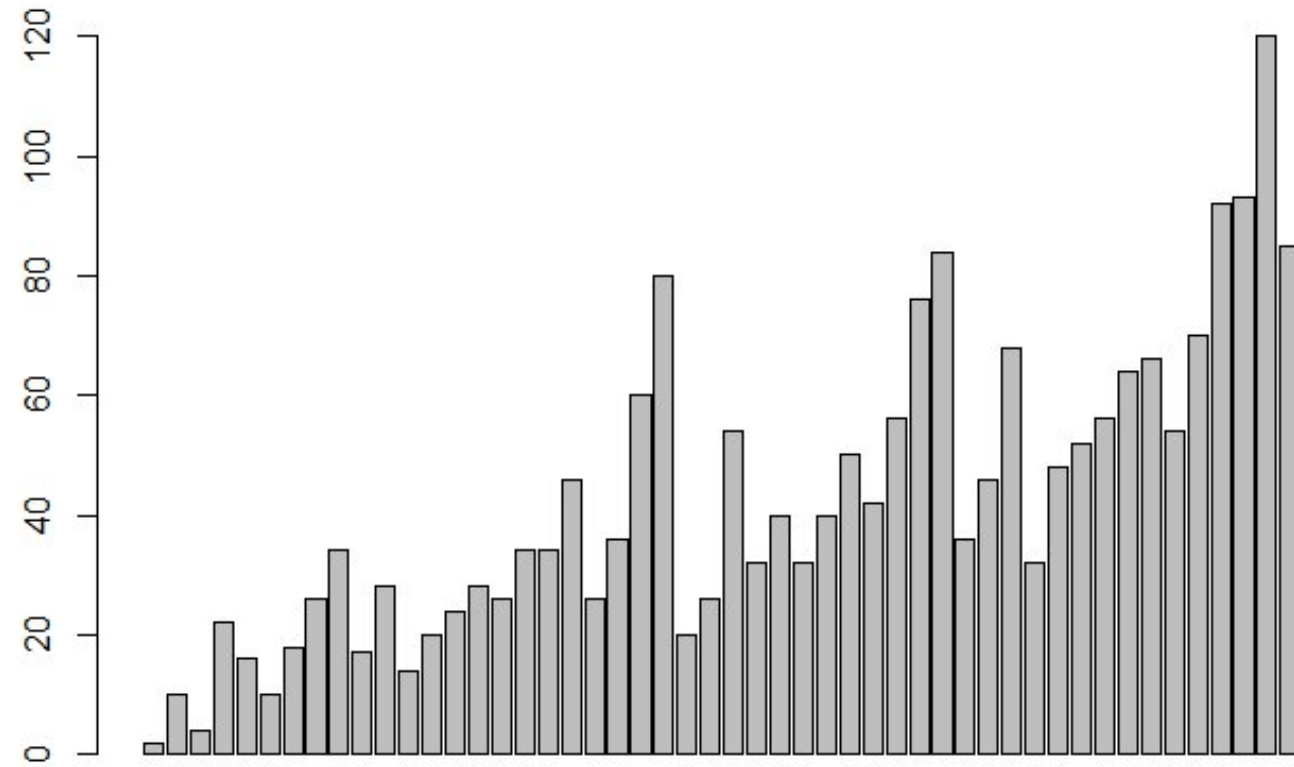
```
names(binomialeverdeling)<-c("0","1","2","3","4")
```

```
barplot(binomialeverdeling,col="lavender",xlab="aantal successen")
```



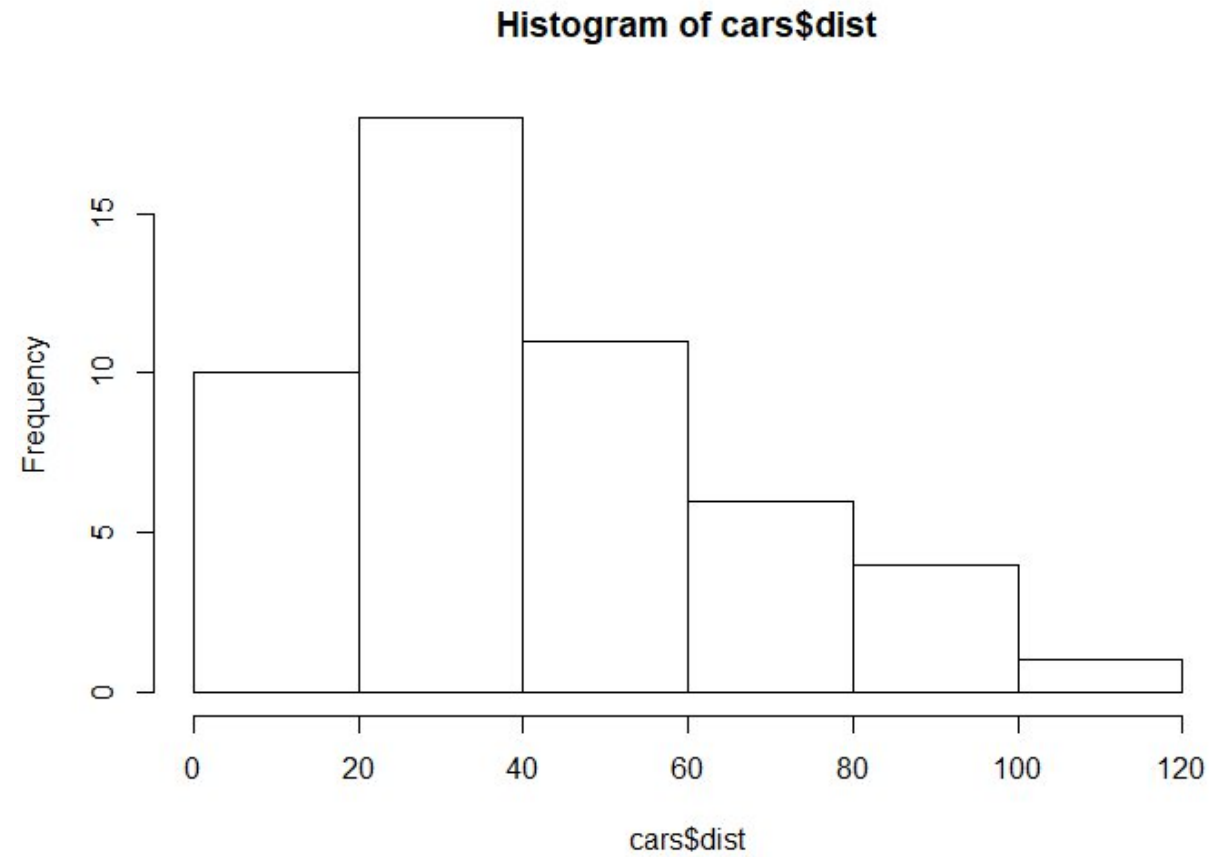
CARS\$DIST

```
barplot(cars$dist)
```



CARS\$DIST

```
hist(cars$dist)
```



HISTOGRAM

?hist

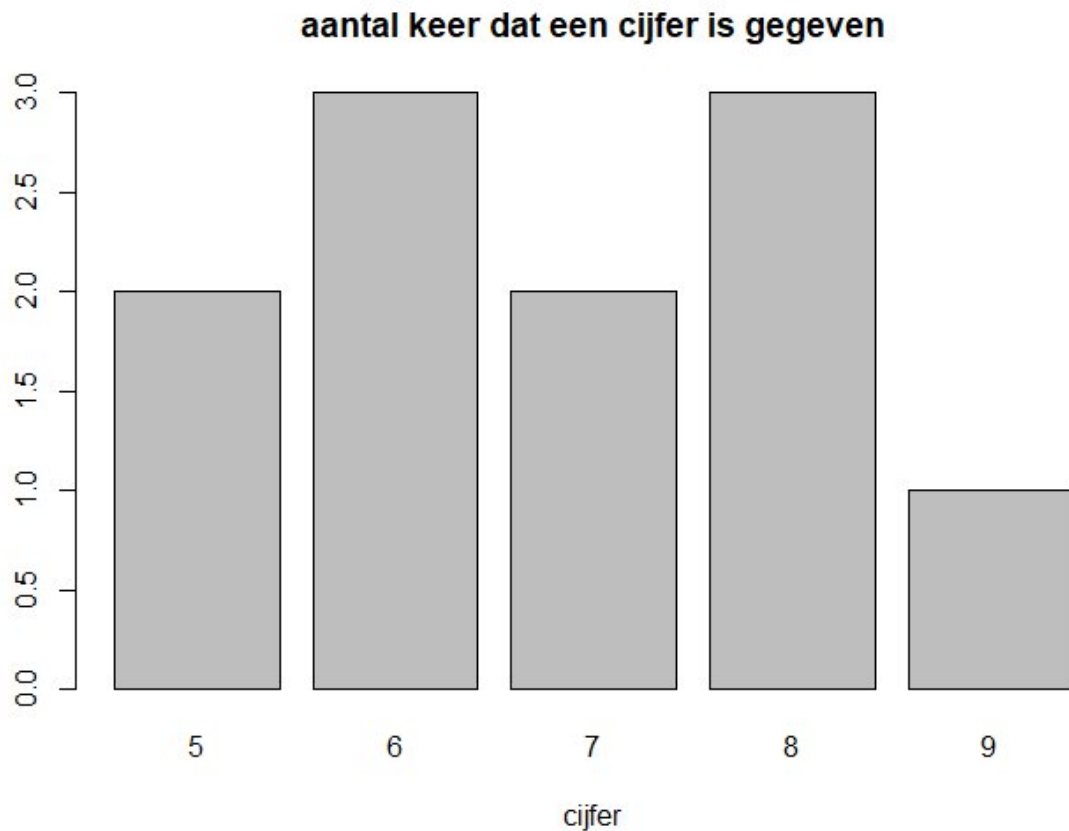
hist(x, ...)

Default S3 method:

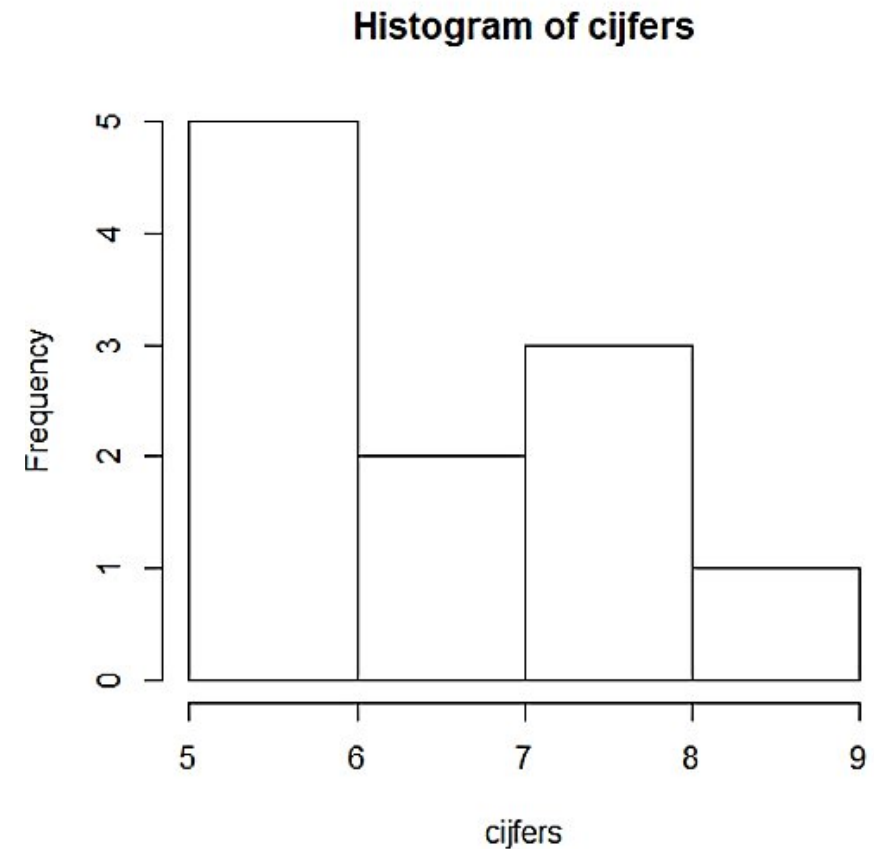
```
hist(x, breaks = "Sturges",  
     freq = NULL, probability = !freq,  
     include.lowest = TRUE, right = TRUE,  
     density = NULL, angle = 45, col = NULL, border = NULL,  
     main = paste("Histogram of" , xname),  
     xlim = range(breaks), ylim = NULL,  
     xlab = xname, ylab,  
     axes = TRUE, plot = TRUE, labels = FALSE,  
     nclass = NULL, warn.unused = TRUE, ...)
```

HISTOGRAM

`barplot(table(cijfers), main = "aantal keer dat een cijfer is gegeven")`



`hist(cijfers)`



HISTOGRAM

```
> x = hist(cijfers)
> x
$breaks
[1] 5 6 7 8 9

$counts
[1] 5 2 3 1

$density
[1] 0.45454545 0.18181818 0.27272727 0.09090909

$mids
[1] 5.5 6.5 7.5 8.5

$name
[1] "cijfers"

$equidist
[1] TRUE

attr(,"class")
[1] "histogram"
```

STACKED/GROUPED BARPLOT

Barplot met matrix zorgt voor een stacked barplot.

Matrix met beside=TRUE zorgt voor grouped barplot.

OPDRACHT

Bekijk de dataset Titanic

Maak mbv `margin.table()` een matrix van het aantal overlevenden per Classe

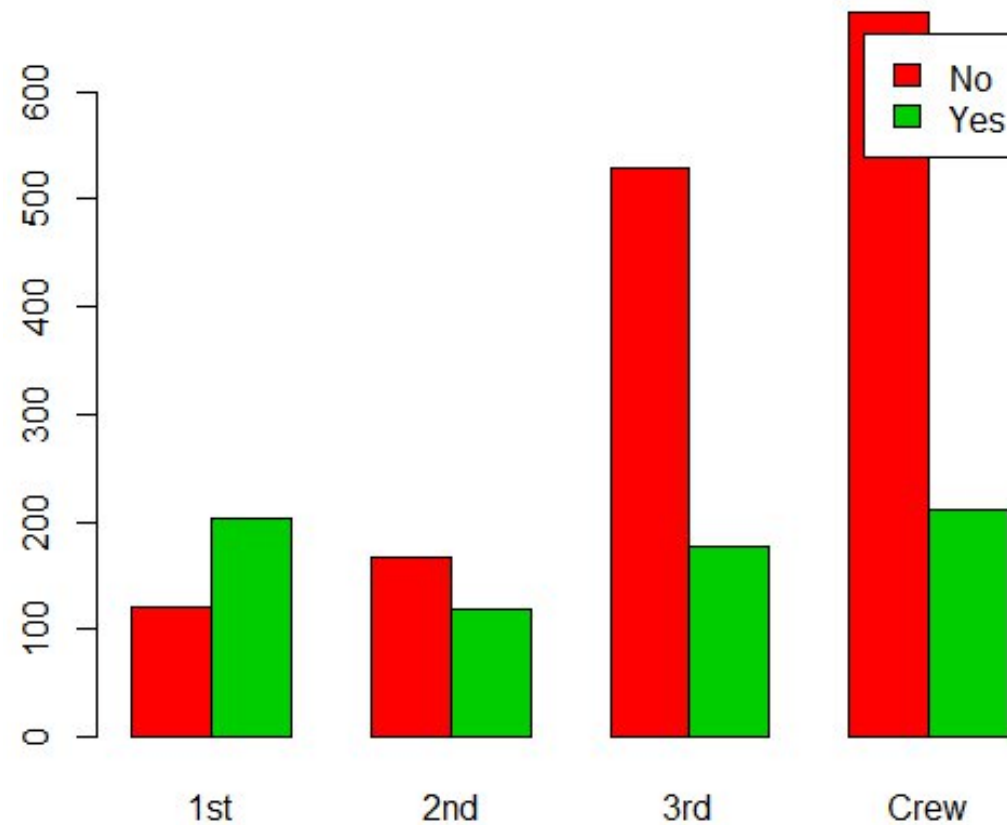
Maak hier een stacked barplot van

Maak hier een grouped barplot van

Vergelijk het verschil, waarmee kun je beter resultaten vergelijken?

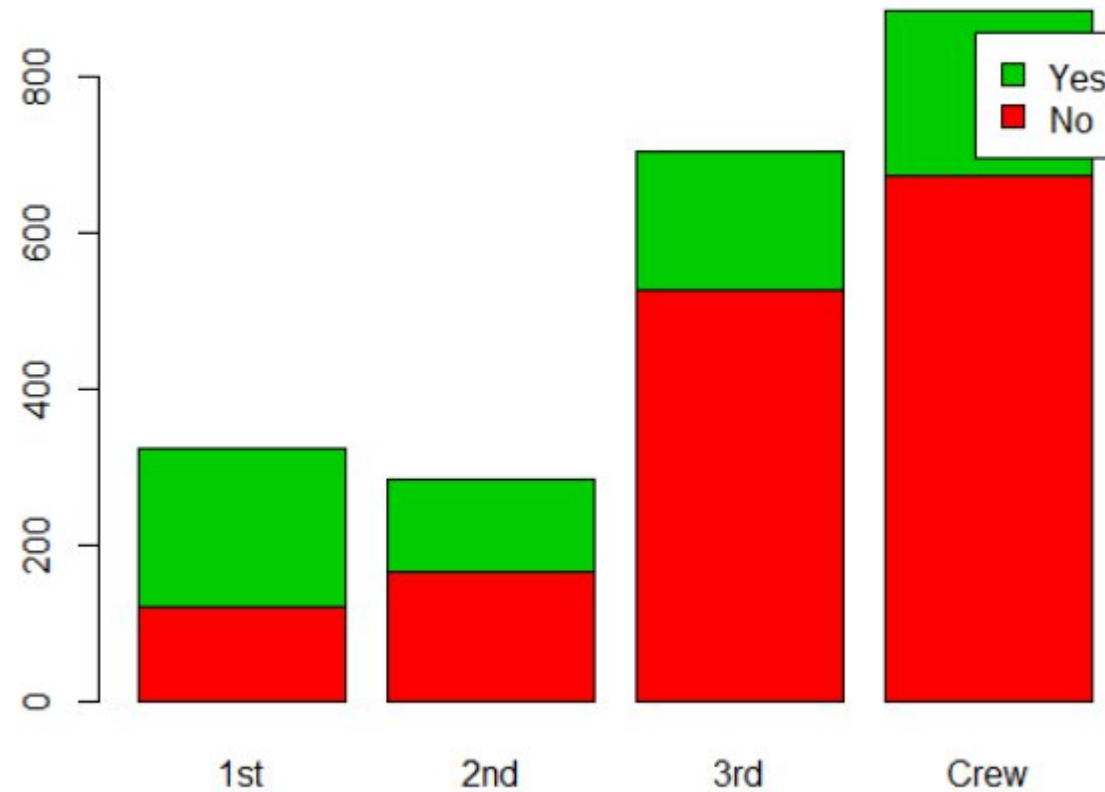
TITANIC

```
barplot(margin.table(Titanic,c(4,1)),legend.text = TRUE, beside = TRUE, col = c(2,3))
```



TITANIC

```
barplot(margin.table(Titanic,c(4,1)),legend.text = TRUE, col = c(2,3))
```



OPDRACHT

Visualiseer een mooie barplot van HairEyeColor

Bedenk een onderzoeksvraag bij de dataset Titanic en maak een grafiek om deze vraag te beantwoorden.

Plot de dataset Formaldehyde en voeg een trendlijn toe.

Bekijk zelf een dataset waar je een grafiek van maakt.

VERANTWOORDING

In deze uitgave is géén auteursrechtelijk beschermd werk opgenomen

Alle teksten © Gonny Henkes tenzij expliciet externe bronnen zijn aangegeven

Screenshots op basis van vernoemde websites

Eventuele images zijn opgenomen met vermelding van bron