

VanderBinckes

Load the necessary libraries to be used in the code

```
library(tidyr)
library(tidyverse)
library(dplyr)
library(ggplot2)
library(readxl)           # Voor het inlezen van data uit Excel
library(lubridate)        # Voor het interpreteren van datums, tijden en periodes
```

Load data into the database and do some initial cleaning

```
# Lees de gegevens
dfData <- read_excel('~\VanderBinckes.xlsx')
```

Vul het geslacht 'onbekend' in als het niet is ingevuld

```
dfData <- dfData %>%
  mutate(Geslacht = replace_na(Geslacht, 'Onbekend'))
```

Corrigeer de types voor alle datums in de tabel

```
dfData <- dfData %>%
  mutate(Created = as.Date(Created, format="%Y-%m-%d", zone="CET")) %>%
  mutate(GebDat = as.Date(GebDat, format="%Y-%m-%d", zone="CET")) %>%
  mutate(dtVertrek = as.Date(dtVertrek, format="%Y-%m-%d", zone="CET")) %>%
  mutate(dtTerug = as.Date(dtTerug, format="%Y-%m-%d", zone="CET")) %>%
  mutate(dtAankoop = as.Date(dtAankoop, format="%Y-%m-%d", zone="CET"))
```

Deel klanten in naar leeftijd in groepen

```
agebreaks <- c(0,5,13,18,30,50,67,80,500)
agelabels <- c("0 -< 5", "5-< 13", "13 -< 17", "18 -< 30", "30 -< 50", "50 -< 67", "67 -< 80", "> 80")

# Verwerk de verdeling in het dataframe
dfData <- dfData %>%
  mutate(Leeftijd = floor(as.numeric(interval(GebDat, today()) %/% years(1)))) %>%
  mutate(AgeGroup = cut(Leeftijd, breaks = agebreaks, right = FALSE, labels = agelabels))
```

Bepaal hoe lang een klant al bekend is (in maanden en in jaren)

```
dfData <- dfData %>%
  mutate(mndKlant = as.numeric(interval(Created, today()) %/% months(1))) %>%
  mutate(jrnKlant = as.numeric(interval(Created, today()) %/% years(1)))
```

Bereken de huurtijd van de fiets in uren

```
dfData <- dfData %>%  
  mutate(tdVertrek = paste(dtVertrek, tdVertrek)) %>%  
  mutate(tdTerug = paste(dtTerug, tdTerug)) %>%  
  mutate(HuurTijd = difftime(tdTerug, tdVertrek, units="hours"))
```

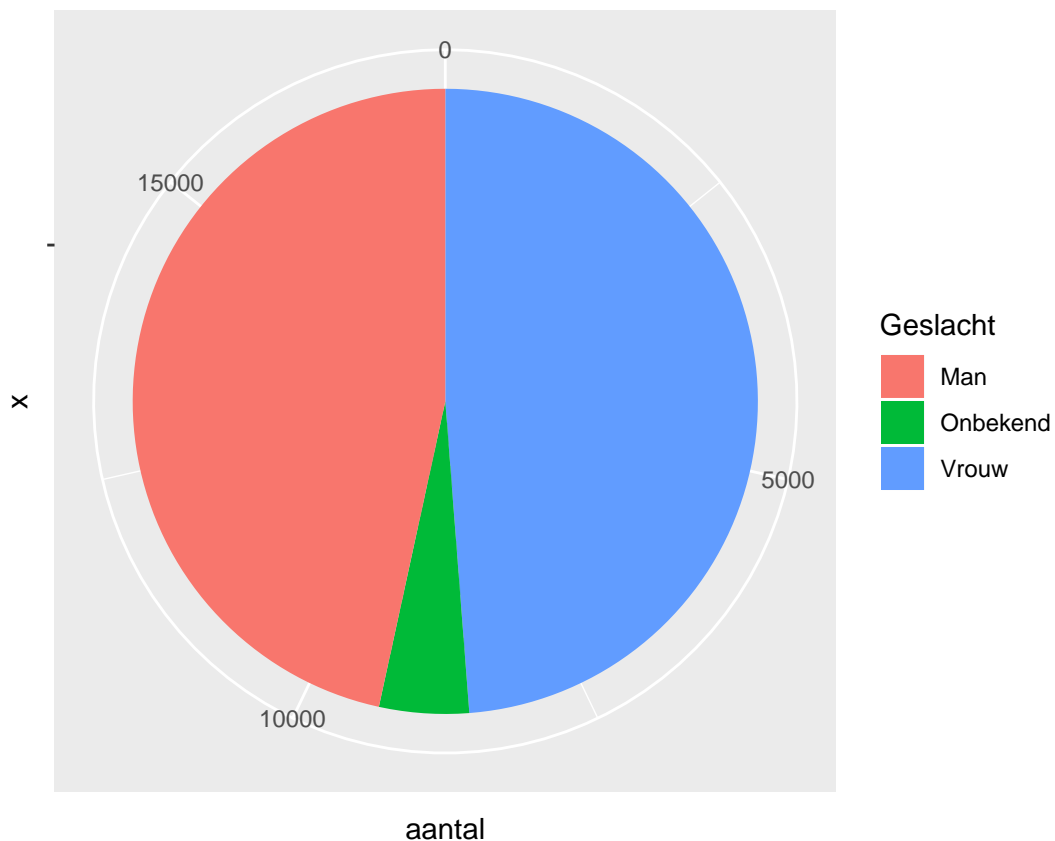
Gebruik tidyverse om de data te groeperen naar aantallen per geslacht (zie ook de chunk)

```
frameGeslacht <- dfData %>%  
  group_by(Geslacht) %>%  
  summarise(aantal = n())  
frameGeslacht
```

```
## # A tibble: 3 x 2  
##   Geslacht aantal  
##   <chr>      <int>  
## 1 Man        8153  
## 2 Onbekend    811  
## 3 Vrouw      8535
```

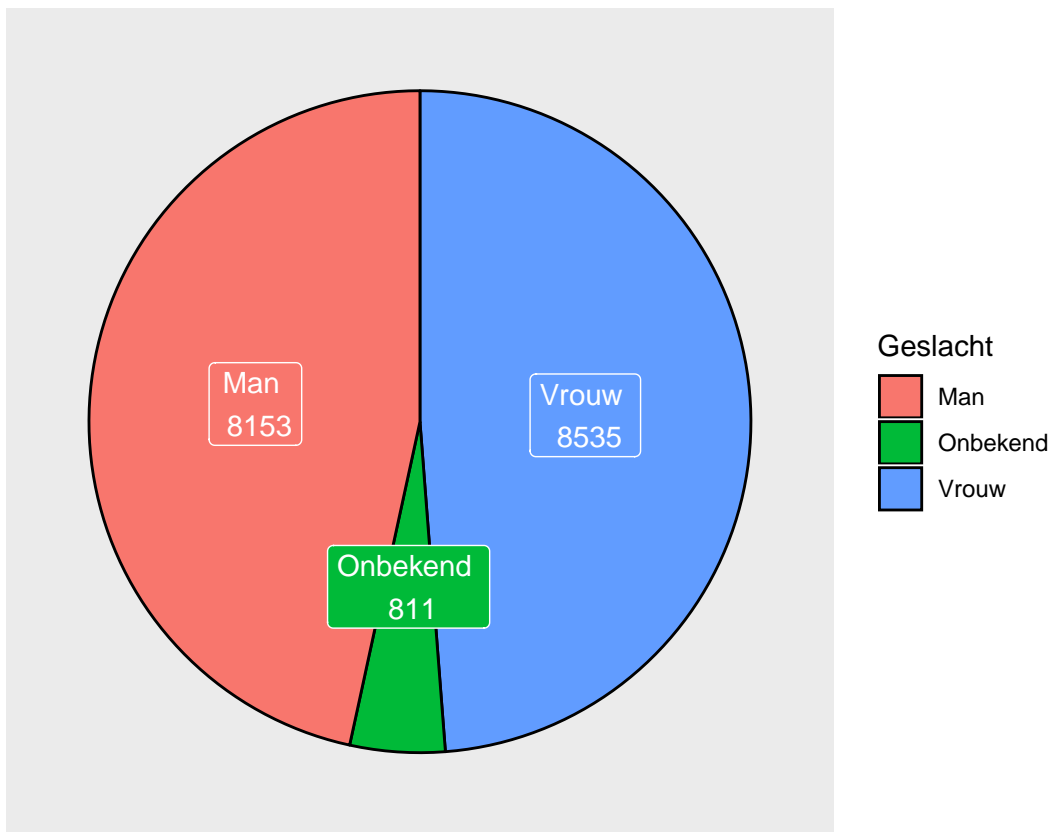
Gebruik <https://www.data-to-viz.com/> voor het maken van een grafiek

```
ggplot(frameGeslacht, aes(x="", y=aantal, fill=Geslacht)) +  
  geom_bar(stat="identity", width=1) +  
  coord_polar("y", start=0)
```



Met ggplot2 kun je alle onderdelen toevoegen, zodat het er nog mooier uit ziet

```
ggplot(frameGeslacht, aes(x="", y=aantal, fill=Geslacht)) +  
  geom_col(color = "black") +  
  theme(legend.position="right",  
        axis.text = element_blank(),  
        axis.ticks = element_blank(),  
        axis.title = element_blank(),  
        panel.grid = element_blank()) +  
  geom_text(aes(label = aantal),  
            position = position_stack(vjust = 0.5)) +  
  geom_label(aes(label = paste(Geslacht, '\n', aantal)),  
            color = "white",  
            position = position_stack(vjust = 0.5),  
            show.legend = FALSE) +  
  coord_polar(theta = "y")
```



#Hier volgen jouw chunks voor het berekenen van de KPIs