

Dataset verkopen - Voorbeeld opdracht [jouw naam]

Dit is een voorbeeld van hoe je aan de slag zou kunnen gaan met R chunks in een Notebook. Hierin ga je:

- opmaak toepassen
- functies uit Datacamp toepassen op een "eigen" dataset
- oefenen met “knitten”

Er is met zogenaamde *Markdown* allerlei opmaak toegepast op de tekst rondom de chunks. Dit is anders dan hoe het in Word gaat (daar is het meer **What You See Is What You Get**), het is wel vergelijkbaar met de opmaak van pagina's in Wikipedia.

Importeren libraries en dataset

Hieronder worden libraries geladen waar functies in zitten die we verderop in het Notebook gaan gebruiken: *readxl* en *tidyverse*.

Daarnaast wordt hier de dataset geïmporteerd zodat we daar vragen aan kunnen stellen. De dataset verkopen gaat over verkopers uit verschillende delen van de wereld, die iPhones en iPads verkopen.

```
# laadt readxl
library(readxl)

# laadt tidyverse
library(tidyverse)

# importeer de dataset
iVerkopen <- read_excel("dataset_verkopen.xlsx")
```

Als je bovenstaande code uitvoert door op het groene pijltje rechtsbovenin de chunk te drukken, gebeurt er nog vrij weinig. Als het goed is verschijnt er alleen even een groen balkje naast de code regels. Tenminste, als je het Excel bestand op de juiste plek hebt geüpload.

Het laden van de libraries hoeven we maar 1 keer te doen en het is wel zo overzichtelijk om dat allemaal aan het begin van je document te doen voor je de inhoud in duikt.

Welke verkoper komt van het Scandinavische continent?

Voor het beantwoorden van deze vraag heb je niet perse R nodig als je naar de data kijkt, maar we gebruiken het nu toch even.

```
## # A tibble: 1 x 9
##   Naam Geslacht Land Continent Producten Opbrengst Dagen_geleden_1...
##   <chr> <chr>    <chr> <chr>    <chr>         <dbl>         <dbl>
## 1 Knut   m        Noor... Scandina... iPhone         6700          34
## # ... with 2 more variables: Klanttevredenheid <dbl>, Aantal_besteld <dbl>
```

Knut is de verkoper die van het Scandinavische continent komt. Merk op dat Scandinavië eigenlijk geen continent is.

Een paar vragen over de opbrengst

Let op: Omdat jullie functies *min* en *max* waarschijnlijk nog niet gehad hebben, lossen we dit op een andere manier op... namelijk door te zorgen dat het antwoord op onze vraag bovenaan in de tabel komt.

En psst, zie je dat het kopje hieronder net iets kleiner is dan die hierboven?

Welke verkoper heeft de laagste opbrengst?

Om deze vraag te beantwoorden sorteren we de data zodat de verkoper met de laagste opbrengst bovenaan komt.

```
## # A tibble: 9 x 9
##   Naam Geslacht Land Continent Producten Opbrengst Dagen_geleden_1...
##   <chr> <chr>    <chr> <chr>    <chr>         <dbl>         <dbl>
## 1 Ling   v        China Azie     iPhone         1000          312
## 2 John   m        Vere... Noord-Am... iPhone         2000          23
## 3 Ariko   v        Japan Azie     iPad           2550          512
## 4 Wim     m        Nede... Europa    iPhone         3000           5
## 5 Ling   v        Taiw... Azie     iPad           3150          259
## 6 Ther... v        Fran... Europa    iPad           3450           15
## 7 Mitch  m        Cana... Noord-Am... iPhone         3500           8
## 8 Fabr... m        Span... Europa    iPad           4100           8
## 9 Knut   m        Noor... Scandina... iPhone         6700          34
## # ... with 2 more variables: Klanttevredenheid <dbl>, Aantal_besteld <dbl>
```

Mevrouw Ling staat bovenaan, dus zij heeft de laagste opbrengst. Het is best lang geleden dat de laatste bestelling bij haar is geplaatst, dus misschien heeft ze maar heel even als verkoper hier gewerkt.

Welke verkoper uit Azie heeft de hoogste opbrengst?

Om deze vraag te beantwoorden filteren we op verkopers uit Azie en sorteren we de data zodat de verkoper met de hoogste opbrengst bovenaan komt.

```
## # A tibble: 3 x 9
##   Naam  Geslacht Land  Continent Producten Opbrengst Dagen_geleden_1...
##   <chr> <chr>   <chr> <chr>      <chr>      <dbl>      <dbl>
## 1 Ling  v       Taiw... Azie      iPad        3150        259
## 2 Ariko v       Japan  Azie      iPad        2550        512
## 3 Ling  v       China  Azie      iPhone      1000        312
## # ... with 2 more variables: Klanttevredenheid <dbl>, Aantal_besteld <dbl>
```

Mevrouw Ling staat bovenaan (en onderaan). Zou dat dezelfde dame zijn? Met iPads heeft ze het dan beter gedaan dan met iPhones.

Bedenk zelf nog een paar vragen die je aan deze dataset zou kunnen stellen. Misschien kun je ze op dit moment nog niet allemaal beantwoorden met je R skills, maar dat komt vanzelf. Experimenteren mag!