Uitleg R script

**Uitleg van de volgende segmenten:**

[**Dataframe** 2](#_Toc61646857)

[**Objecten** 2](#_Toc61646858)

[**For loop om de overwinningen van AOpen te tellen** 3](#_Toc61646859)

[**Loop om de wins van Novak Djokovic te vergelijken met Roger Federer** 4](#_Toc61646860)

# **Dataframe**

Allereerst heb je in problem1.txt het dataframe meegekregen (dubbelklik op het icoontje om te openen):

In het eerste commando wordt er aangegeven dat de eerste kolom optellend is van de values 2004 tot en met 2016. Dus er komt sws in het datafame een lijst te staan vanaf 2004 tot en met 2016.

Code is als volgt:

AOpen <- data.frame(2004:2016,c(**HIERIN STAAN DE TENNISERS**)

colnames(AOpen) <- c("Year","Champion")

In het commando erna wordt “colnames” uitgevoerd, dit is een afkorting voor columnnames. Nu je dat weet kan je zien dat de eerste kolom “Year” gaat heten en de tweede kolom “Champion”.

Voor visualisatie komt het er zo uit te zien in een Excel bestand:

|  |  |
| --- | --- |
| Year | Champion |
| 2004 | Roger Federer |
| 2005 | Marat Safin |
| 2006 | Roger Federer |
| 2007 | Roger Federer |
| 2008 | Novak Djokovic |
| 2009 | Rafael Nadal |
| 2010 | Roger Federer |
| 2011 | Novak Djokovic |
| 2012 | Novak Djokovic |
| 2013 | Novak Djokovic |
| 2014 | Stan Wawrinka |
| 2015 | Novak Djokovic |
| 2016 | Novak Djokovic |

Zoals je kan zien zijn er meerder overwinningen voor Roger Federer & Novak Djokovic.

Roger heeft er 4 en Djokovic heeft er 6. We gaan dit tellen met behulp van een scriptje.

# **Objecten**

Om de overwinningen te tellen van deze 2 tennissers worden er 2 objecten aangemaakt. Voor Roger Federer en voor Novak Djokovic.

Dit doen we zo:

Wins\_ND <- 0

Wins\_RF <- 0

De benamingen spreken voor zich, we zetten beiden objecten op 0 omdat we door het bovenste dataframe heen willen zoeken naar voorkomingen van winsten.

# **For loop om de overwinningen van AOpen te tellen**

Om de hoeveelheids overwinningen van de AOpen te tellen gaan we een for loop gebruiken. En for loop kan je gebruiken om door elk item in een lijst/dataframe heen te gaan.

De code ziet er als volgt uit:

**For(y in AOpen[,2]) {**

**if(y == “Novak Djokovic”) {**

**Wins\_ND <- Wins\_ND + 1**

**}**

**if(y == “Roger Federer”){**

**Wins\_RF <- Wins\_RF + 1**

**}**

**}**

In de eerste zin geven we aan dat we een FOR loop gebruiken. Y kan hier elke letter zijn. Het dient om dus aan te geven dat 1 object uit het dataframe voortaan y heet. Daarna geven we aan dat we willen werken met het DataFrame AOpen, daarna kan je de door de [,2] zien dat we uit de 2e kolom gaan werken. De dubbele == staat voor **IS GELIJK AAN.** We gaan de for loop doorlopen en beginnen met de eerste waarde van column 2 (zie tabel vorige pagina):

1. Hier heeft y als eerst de waarde gekregen “Roger Federer”
2. Dan gaat hij door de loop, if (als) y hier gelijk aan “Novak Djokovic” tel 1 bij Wins\_ND op, onze waarde is niet gelijk aan “Novak Djokovic” dus gaat het door naar de volgende if statement.
3. Dan gaat het door naar de volgende if (als) y hier gelijk is aan “Roger Federer” tel dan 1 op bij Wins\_RF. Onze waarde van y is hier gelijk aan “Roger Federer”
4. En zo gaat het de hele lijst door van boven naar beneden.

Als dit voltooid is zullen Wins\_ND en Wins\_RF corresponderen met de data in de tabel die ik gemaakt heb. Dat betekent dat Wins\_ND nu een waarde heeft van 6 en Wins\_RF een waarde van 4.

Deze waardes zijn nu bijgewerkt en kunnen we verder mee gaan werken.

# **Loop om de wins van Novak Djokovic te vergelijken met Roger Federer**

In het laatste code segment gaan we de 2 waardes met elkaar vergelijken door middel van een if (als) loop. De code van dit segment ziet er zo uit:

**if (Wins\_ND > Wins\_RF {**

**print(“Djokovic has more wins”)**

**} else{**

**if (Wins\_ND == Wins\_RF) {**

**print(“Both have the same number of wins”)**

**} else{**

**print(“Federer has more wins”)**

**}**

**}**

De code werkt als volgt:

1. Eerst gaan we kijken of Wins\_ND > (groter is dan) Wins\_RF als dit het geval is printen we “Djokovic has more wins”. Print betekent output deze tekst (ook string genoemd) in dit geval.
2. Als dit niet het geval is gaan we door naar de volgende if, dit wordt aangegeven door else. Als dit niet klopt doe dan het volgende
3. Bij het 2e segment gaan we kijken of de Wins\_ND == (gelijk is aan) Wins\_RF als dit het geval is printen we “Both have the same number of wins”
4. Als dit niet het geval is gaan we door naar de volgende opdracht
5. Bij het 3e segment is er geen vergelijking meer nodig dus printen we in het geval dat allebei de bovenstaande condities niet zijn voldaan. Hier printen we “Federer has more wins”.

Ik hoop dat je wat snapt van m’n uitleg, als je iets niet snapt kan je me altijd appen/bellen/snappen en dan zal ik m’n best doen om het zo goed mogelijk uit te leggen.😉😘