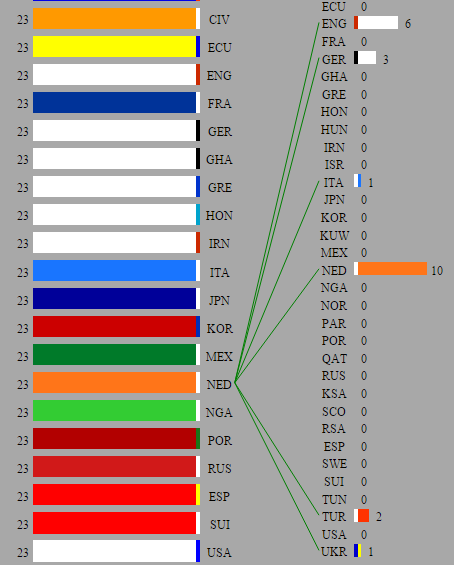
Eindproject: Competitieherkomst van spelers op het wereldkampioenschap

Lucas Zuurveld (10220380)

Doel: Het inzichtelijk maken van de competitieherkomst van aan de WK 2014 deelnemende spelers

Het doel van de visualisatie is om inzicht te verkrijgen in de competitieherkomst van de spelers die actief waren op het wereldkampioenschap voetbal 2014. Het is de bedoeling dat snel kan worden ingezien in welke competities er veel spelers spelen die op het WK 2014 actief waren en in welke juist weinig. Daarnaast moet het mogelijk zijn om snel te ontdekken hoe de spelers van een specifiek land over de verschillende competities verdeeld zijn en hoe spelers uit een specifieke competitie over de WK-landen zijn verdeeld.

Bovenstaande doel kan worden bereikt door een interactieve visualisatie, waarbij links staafdiagrammen worden getoond van de WK-landen en rechts de staafdiagrammen met de competitielanden. De staven geven aan hoeveel spelers er voor het betreffende WK-land, of in het betreffende competitieland spelen. Als er op een bepaalde bar van een WK-land geklikt wordt, moeten de staafdiagrammen van de competitielanden zodanig worden aangepast dat alleen de spelers zijn opgenomen van het betreffende land waarop geklikt is. Daarnaast moeten de namen van de spelers worden getoond op het scherm. Het omgekeerde geldt als er op een bar van een competitieland wordt geklikt. In Figuur 1 volgt een voorbeeld hiervan.



Figuur 1.

Als er op het WK-land Nederland (NED) wordt geklikt, wordt duidelijk dat er van de 23-koppige WK-selectie er 6 spelers in Engelse competitie spelen (ENG), 3 in de Duitse (GER), 1 in de Italiaanse (ITA), 10 in de Nederlandse (NED), 2 in de Turkse (TUR) en 1 in de Oekraïense (UKR).

De data is afkomstig van de officiële website van de FIFA (Fédération Internationale de Football Association). Via deze website is het mogelijk om alle spelersprofielen te scrapen. Een spelersprofiel bevat onder andere het WK-land waarvoor hij speelt, de club waarvoor hij speelt en de afkorting van het land van de betreffende competitie. Als de gehele dataset met spelersprofielen is gescrapt, kunnen die afkortingen worden omgezet in landsnamen middels een beschikbaar afkortinglijstje op Wikipedia. Via eenvoudige formules in Microsoft Excel is dit eenvoudig te bereiken.

Er kan een opsplitsing worden gemaakt in het scrapen van de data en visualiseren van de data. Bij het scrapen moet eerst van de “hoofdpagina” naar de pagina van een bepaald land worden gelinkt, om van daaruit door te linken naar de pagina van een betreffende speler. Van de pagina van een betreffende speler kunnen de juiste gegevens worden gescrapt.

De visualisatie kan worden onderverdeeld in vier onderdelen. De pagina zoals die getoond wordt als er nog geen interactie plaats heeft gevonden vormt de natuurlijk de basis. Daarnaast zijn er drie verschillende soorten interacties. Het is mogelijk om van de WK-landen of competitielanden alleen een bepaald continent te selecteren vanuit een dropmenu. Zodoende kan ingezien worden hoeveel spelers er van Zuid-Amerikaanse elftallen spelen in Europese competities en hoe deze verdeeld zijn. De tweede mogelijke interactie is het selecteren van een bepaald land, zoals in Figuur 1 duidelijk is gemaakt. De derde interactie is het selecteren van een spelersnaam. Als er een spelersnaam geselecteerd wordt, verschijnt er een soort toolbar met een het betreffende spelersprofiel. Er is bijvoorbeeld te zien bij welke club de speler speelt, hoeveel interlands hij al achter zijn naam heeft staan en op welke positie de speler speelt.

De visualisatie wordt gemaakt in Javascript. Van de D3 library is daarbij gretig gebruik gemaakt. De D3 library is zeer handig om DOM manipulaties uit te voeren. Ook staan er zeer bruikbare voorbeelden op het web voor vrijwel iedere soort visualisatie.

Er zijn geen problemen die op voorhand te definiëren zijn voor dit probleem. Met programmeren heb ik de ervaring dat problemen zich gewoon voordoen tijdens het proces.

In het onderzoek van Guy J. Abel zijn verscheidene visualisaties te zien die hetzelfde onderwerp betreffen als dit project. In de bijlage is het meest treffende voorbeeld opgenomen. Guy J. Abel gebruikt in zijn visualisatie een cirkel waarbij lijnen zijn verbonden om de uitwisseling van spelers van WK-landen en competitielanden aan te geven. In zijn visualisatie is bij de verbindingen moeilijk te zien welk land het WK-land voorstelt, en welk land het competitieland. Daarnaast worden er zo veel verbindingen gelegd in de visualisatie dat het moeilijk is om een inzicht te verkrijgen. Je ziet eigenlijk door de bomen het bos niet meer.

Mij lijkt het beter om de visualisatie weer te geven op een soortgelijke wijze als op de volgende pagina gedaan is:

<http://bl.ocks.org/NPashaP/9796212>

Bij dit voorbeeld is een interactie mogelijk waardoor ‘ingezoomd’ kan worden op een bepaald onderdeel, zodat de details daarvan beter zijn te bestuderen. De hoogtes van de bars geven bovendien een kwantiteit aan, terwijl de kleuren een onderscheid maken tussen nominale variabelen. Deze twee eigenschappen wil ik ook gaan gebruiken. Ik zal de kleuren gebruiken van het shirt van het land om onderscheid aan te brengen. De staafdiagrammen zal ik horizontaal willen weergeven, omdat het aantal landen zo hoog ligt, dat het anders te veel ruimte inneemt om dit verticaal te doen. Daarnaast zijn de hoogtes makkelijker te vergelijken als ze onder elkaar worden weergegeven horizontaal. In het genoemde voorbeeld zijn de hoogtes moeilijker te vergelijken.

Guy Abel onderzoek: <http://bl.ocks.org/NPashaP/9796212>

