Data Science Portfolio

**101**

Ik heb niks aan te merken op de uitleg in assignment over de betekenis van Data Science.

In opdracht 12 analyseren we eerst de gegevens uit de Pokemon dataset. Ik zou in dit geval graag willen weten hoeveel de HP van een Pokemon per GENERATION is. Ik bekijk eerst alle gegevens per GENERATION op gemiddelde, daarna maak ik een barplot waarin de HP per GENERATION word weergeven. Hierin kan ik makkelijk het verschil van HP tussen Pokemon zien per GENERATION

**102**

Een data scientist kijkt naar de dataset waar ie mee gaat werken. Hij bedenkt en gebruikt algoritmen voor de nuttigste berekeningen om uit te voeren op deze dataset. De berekeningen worden uitgevoerd, de resultaten worden in een boxplot of histogram geladen, de data scientist kijkt naar deze resultaten en kan aan de hand hiervan bedenken wat de uitkomst van de specifieke gegevens binnen de dataset gaat zijn. Vervolgens word met bijvoorbeeld Jupyter een voorspelling gedaan. De data scientist kijkt nu of deze voorspelling overeenkomt met wat het daadwerkelijk had kunnen zijn. In opdracht 16 wil ik bijvoorbeeld weten welk type Pokemon het hoogste HP heeft. Ik maak hier een decision tree, hierin is te zien wat de voorspelling is, en daaronder kan ik zien wat de daadwerkelijke waarde is. De HP van Water-type Pokemon zit het dichtst bij de hoogst mogelijke HP. Vergeleken met een Normal-type is Water-type hoger en de voorspelde waarde komt ook overeen met de waarde van Water-types.