# Wat is Data Science?

Data science is het vergaren van gegevens en deze omzetten naar informatie en kennis om op basis hiervan voorspellingen te maken voor de toekomst. Het voorspellen van deze gegevens kan op basis van verschillende methoden gaan. Dit begint bij het uitzoeken wat voor type analyse er plaats moet gaan vinden. Dit kan 1 uit 3 analyses zijn:

* Univariate analysis
* Bivariate analysis
* Multivariate analysis

Op basis van de gekozen analyse kan aan de hand van het soort data een type analyse of plot uitwijzen. Op basis hiervan kunnen de gegevens omgezet worden tot informatie en eventueel voorspellingen.

Een van de meest gebruikte talen voor data science is python. Deze taal heeft veel verschillende libraries die helpen bij het verwerken en analyseren van data. De top 3 meest gebruikte libraries binnen python, voor het analyseren en verwerken van data, zijn:

* NumPy
* Pandas
* Matplotlib

Kortom, data science maakt data en gegevens inzichtelijk zodat hier informatie uit gehaald kan worden. Vervolgens worden op basis van deze gegevens beslissingen genomen of schattingen gemaakt voor in de toekomst. Python en de bovenstaande libraries vormen een cruciale rol in dit proces aangezien dit de meest gebruikte tools zijn door data scientists.