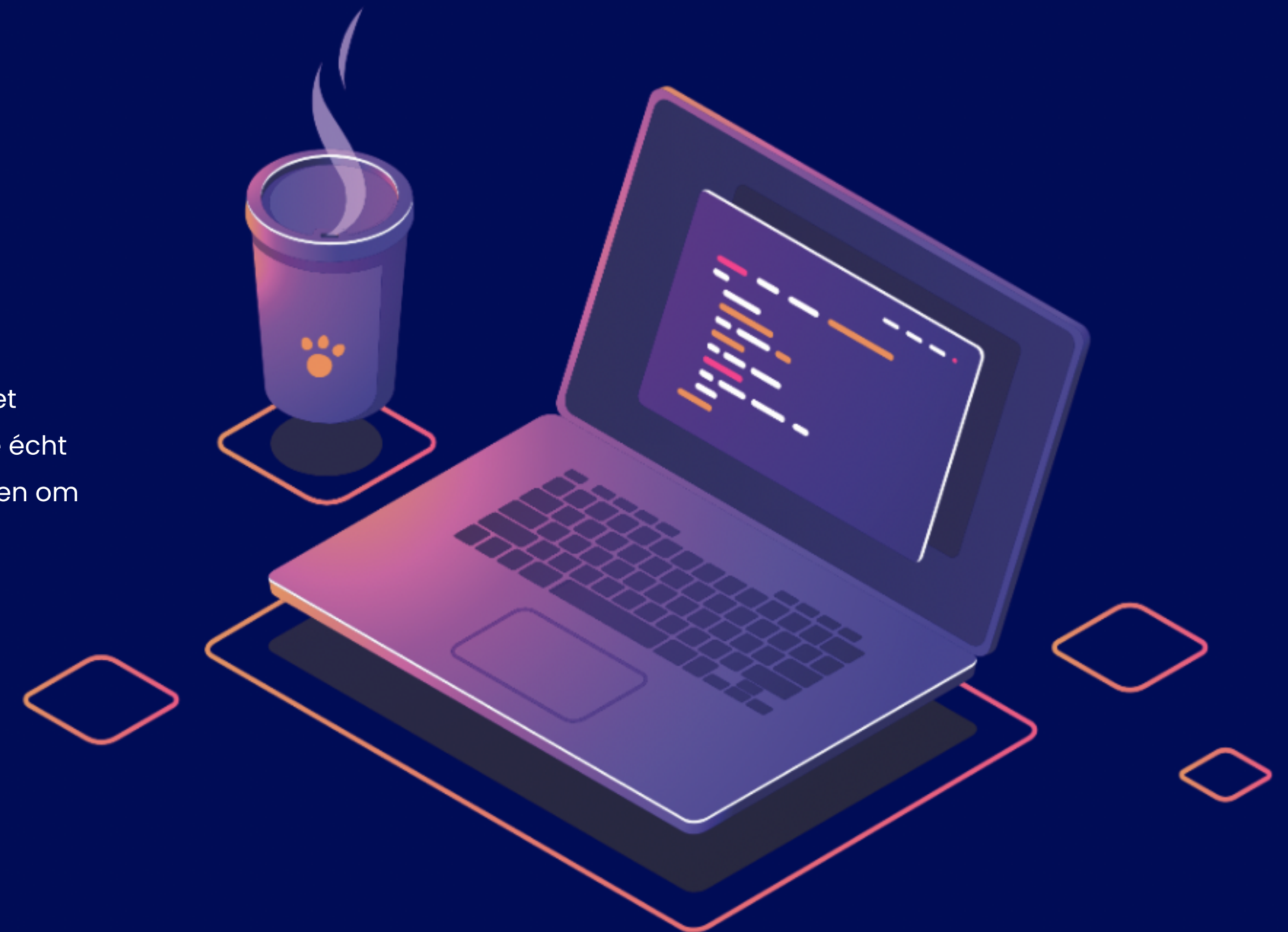


Junior data engineer track

Tijdens de opleiding ga je verschillende modules door via ons online leerplatform Jarvis. Door het maken van oefeningen en het reviewen van de uitwerkingen van medestudenten leer je code je écht eigen te maken. Een aantal modules zijn verplicht en een aantal zijn optioneel. Mocht het je lukken om deze optionele modules ook te maken binnen de tijd dan raden wij dat zeker aan.



Wat ga je precies leren?

week 01

Introductie terminal & databases

In de eerste week maak je kennis met **databases en de terminal**. De terminal is een programma wat standaard op elke Windows/Apple computer is geïnstalleerd. Hierin leer je hoe je met een aantal commando's via je terminal door je mappen kunt navigeren en deze kunt aanpassen. Ook wordt er in deze week **SQL** geïntroduceerd, dé programmeertaal waarmee je (relationele) databases kan beheren. Je leert onder andere hoe je een database kan aanmaken, data kan toevoegen, updaten en verwijderen. Deze skills zijn enorm belangrijk bij vrijwel elke baan die te maken heeft met data.



week
02

Introductie Python

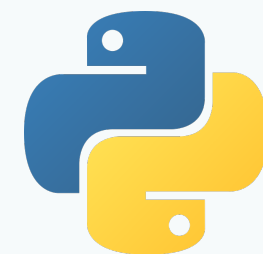
Tijdens deze week maak je kennis met Python. De verschillende data types zullen behandeld worden. Je leert **basistechnieken** die binnen Python, maar eigenlijk binnen vrijwel elke programmeertaal, belangrijk zijn. Denk hierbij aan if/else statements, loops, en functies. Je zal onder andere een eigen rekenmachine gaan programmeren. Na deze week kan je de output van je code afhankelijk laten zijn van de input die een gebruiker geeft.



week
03

Deepdive Python

In week 2 heb je een introductie gehad van Python. Tijdens week 3 wordt hierop voortgeborduurd en wordt duidelijk waarom de basis belangrijk is om goed onder de knie te hebben. Bouw een vereenvoudigde versie van roulette in Python waarin loops en if-else statements een grote rol spelen. Leer ook hoe je virtual environments kan aanmaken zodat jouw collega's of andere geïnteresseerden jouw code ook kunnen uitvoeren op hun computer. Als kers op de taart leer je verbinding te maken met een **MySQL database** vanuit Python.



week

04

&

05

Visualiseren

In week 4 en 5 leer te werken met **Jupyter Notebooks**, zodat je jouw code via de browser kan uitvoeren. Dit is momenteel de meest gebruikte open source tool voor Data Engineers. Verder leer je ook wat je kunt doen met **Pandas**, een bibliotheek met data manipulatie functies die je kunt gebruiken met Python. Tot slot leer je hoe verschillende grafieken kunt plotten en stijlen. Denk aan een histogram, lijnplot, piecharts en meer.



week

06

&

07

Verzamelen

In week 6 en 7 worden technieken behandeld waar een Data Engineer veel aan heeft. Denk aan slimme manieren om in één keer veel csv bestanden te verwerken. Bestanden opschonen met behulp van reguliere expressies (**regex**). Ook leer je hoe je informatie uit websites geautomatiseerd kunt verzamelen, ook wel **scrapen** genoemd.

Aan het einde van deze weken leer je ook een connectie te maken met de **NS API**, zodat je via deze API informatie over treinstations in Nederland op kan halen. Hiermee kan je bijvoorbeeld zelf een route plannen of de NS-app namaken.



week
08

Manipuleren

In week 8 leer je hoe je data kunt manipuleren. Denk aan methodes die je kunt gebruiken om met missende waarden of met uitschieters in je dataset om te gaan. Ook leer je hoe je gemiddelden, medianen en sommen kan berekenen. Verder kun je na deze week ook tabellen filteren, omgaan met tijden, kolommen sorteren, kolommen toevoegen en nog veel meer.



week

09

&

10

Het eindproject

In deze laatste weken staat alles in het teken van je eindopdracht. Het eindonderwerp kies je zelf en alle opgedane kennis en vaardigheden kun je toepassen om zelfbedachte vragen te beantwoorden met behulp van **Python** en/of **SQL code**.



bit
academy

Bij ons krijg je een vliegende start!