

Opleiding Data Science

De opleiding Data Science is ontwikkeld door:



rijksuniversiteit
 groningen

**IT Academy Noord-Nederland
zet zich in voor een krachtige
ICT-arbeidsmarkt in het Noorden.
Samen groeien.**



Vakgebied Data Science De toekomst is nu

Op 14 september 2020 start de negende editie van de opleiding Data Science. De vakken uit deze unieke en technische opleiding worden verzorgd door docenten van de Hanzehogeschool en Rijksuniversiteit Groningen, alsmede aangevuld en versterkt door experts van andere kennisinstellingen en bedrijven.

De opleiding Data Science bestaat uit 9 modules. Door middel van hoor-, werk- en gastcolleges wordt het vakgebied eigen gemaakt en wordt al het nodige over de verschillende facetten van Data Science bijgebracht.

Het vakgebied Data Science

Vandaag de dag is de waarde van data revolutionair in de moderne digitale economie. De statistische methoden en het ontdekken van patronen in grote datasets hebben een grote rol in het ontdekken van verborgen informatie en onbekende relaties tussen data. Op basis van deze nieuwe informatie kunnen bedrijven en organisaties met behulp van Data Science innoveren. Denk hierbij aan het ontwikkelen van nieuwe producten/diensten, managementinformatie en het verhogen van dienstverlening voor de klant.

Moderne organisaties zijn bezig om hun organisatie te transformeren naar een data gedreven organisatie om de business en operationele processen te optimaliseren. Deze

verandering vraagt om een type medewerker met kennis van data intensieve technologieën. Deze rol wordt gedefinieerd als de Data Scientist. Deze rol wordt in het onderzoek van het Horizon 2020 Edison Project als volgt omschreven:

De Data Scientist vindt en interpreteert data, management grote hoeveelheden data en voegt deze data samen. De Data Scientist zorgt voor consistente data en creëert visualisaties om anderen te helpen om de data te begrijpen. Daarnaast bouwt hij/zij wiskundige modellen, presenteert en communiceert data inzichten en bevindingen aan specialisten en collega's. Daarnaast adviseert een Data Scientist natuurlijk over oplossingen om de data toe te passen.

IT Academy Noord-Nederland

Wij zetten ons in voor een krachtige ICT-arbeidsmarkt in het Noorden. De IT Academy is onderdeel van de Hanzehogeschool Groningen en wordt mede mogelijk gemaakt door de Rijksuniversiteit Groningen, Noorderpoort, Samenwerking Noord en provincie & gemeente Groningen.

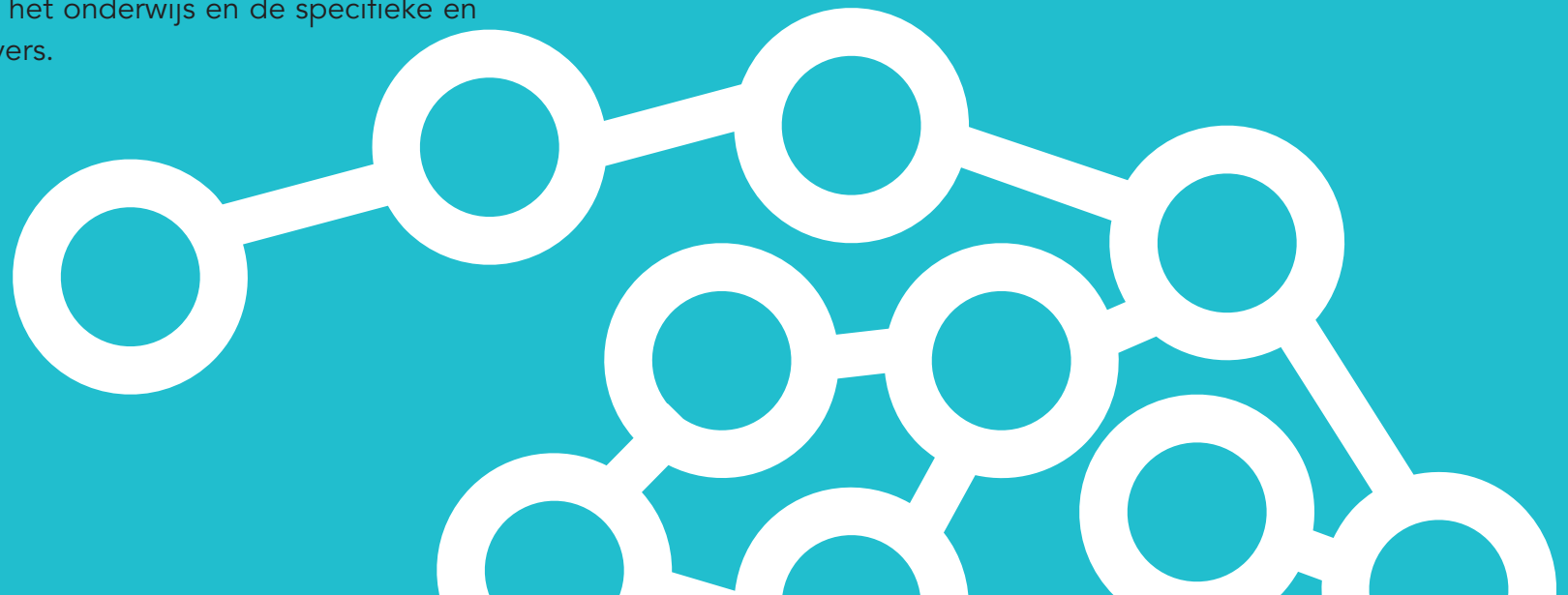
De IT Academy is een zichtbare en toonaangevende speler in de noordelijke regio. Onze ambities zijn:

Duurzame verbetering van de regionale arbeidsmarkt

We werken aan een duurzame verbetering van de regionale ICT-arbeidsmarkt door toekomstgericht professionals op te leiden. Hiermee willen we voorzien in de huidige en toekomstige om- en bijscholingsbehoefte. De IT Academy draagt zorg voor een optimale aansluiting van het onderwijsaanbod op de behoeften van werkgevers. Dit zijn niet alleen ICT-bedrijven, maar ook andere organisaties in Noord-Nederland, waarin ICT een belangrijke rol vervult in het primaire proces. Dit zorgt voor een cruciale 'brugfunctie' tussen het onderwijs en de specifieke en actuele wensen van werkgevers.

Versterking van de positie van het Noorden als ICT-regio

IT Academy Noord-Nederland helpt mee de positie van de regio als aantrekkelijke 'ICT-regio' te versterken, door het bedrijfsleven nadrukkelijk te betrekken bij ICT-onderwijs. Dit is een zichzelf versterkend proces. Hierdoor krijgen zowel de werkgelegenheid als de kennisontwikkeling in het Noorden een boost.



Opzet van het programma

De opleiding Data Science begint met een kick-off dag waar het vakgebied wordt geschetst en uitleg wordt gegeven over het opleidingsprogramma. Daarna wordt direct aan de slag gegaan met Design Thinking. Tijdens deze workshop wordt een beroep gedaan op creativiteit, een belangrijke eigenschap van iedere Data Scientist.

Het volledige programma bestaat uit een kick-off en negen modules. Iedere module bestaat uit drie collegedagen.

Tijdens deze collegedagen wordt de theorie afgewisseld met praktijkopdrachten.

De modules worden verzorgd door docenten van de Hanzehogeschool en Rijksuniversiteit Groningen.

Afhankelijk van de docent worden een aantal modules Engelstalig verzorgd. Elke module wordt afgesloten met een gastcollege. Het gastcollege staat los van de module en is bedoeld om vanuit een ander perspectief inzicht te geven in het vakgebied. Tijdens deze gastcolleges worden er bijvoorbeeld praktijkcases gepresenteerd of wordt er een college verzorgd over big data & recht of ethiek.

Keuzevarianten

Interesse in alleen het eerste deel (t/m module 4) van de opleiding Data Science? Kijk dan op <https://www.itacademy.nl/opleidingen/data-analyse>

Programma

Fase	Data
Kick-off: workshop Design Thinking	14 september 2020
Module 1: Databases	21, 28 september en 5 oktober 2020
Module 2: Data Visualisation	19, 26 oktober en 2 november 2020
Module 3: Descriptive Statistics	9, 16 en 23 november 2020
Module 4: Data Mining	30 november, 7 en 14 december 2020
Module 5: Python	4, 11 en 18 januari 2021
Module 6: Data Preparation/ pre-processing	25 januari, 1 en 8 februari 2021
Module 7: Predictive Analysis	15 februari, 1 en 8 maart 2021
Module 8: Machine Learning	15, 18 en 29 maart 2021
Module 9: Afsluitende opdracht	12, 26 april en 17 mei 2021

Doelgroep & Resultaat

Instroomniveau

De opleiding Data Science is speciaal ontwikkeld voor professionals. Om de opleiding succesvol af te sluiten, wordt van iedere deelnemer een bepaald kennisniveau verwacht. Voor de gehele opleiding gelden de volgende toelatingseisen:

- Afgeronde HBO/WO opleiding
- Basiskennis wiskunde en statistiek
- Basiskennis databases
- Basiskennis ICT-systemen

Programmeerervaring is niet vereist. Echter is programmeerervaring (bijvoorbeeld C/C++, Java, R en met name Python) wel een pré.

Resultaat na afronding

Na afloop van de gehele opleiding zijn deelnemers in staat om aan de slag te gaan met het toepassen van Data Science binnen de eigen organisatie.

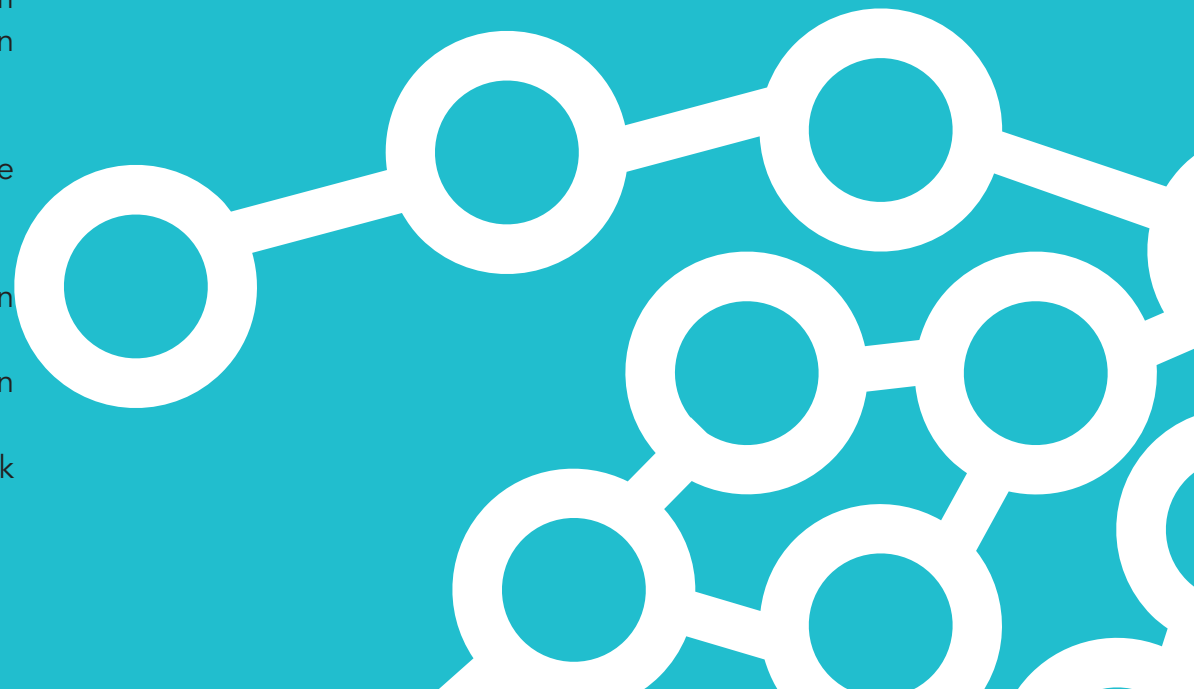
Na afloop beschikken deelnemers over de volgende competenties:

- Databases: Ontwerpen, modelleren bouwen en bevragen van databases.
- Data Visualisation: Weergeven van informatie toegesneden op de doelgroep.
- Descriptive Statistics: Toepassen van beschrijvende statistiek op dataverzamelingen.

- Data Mining: Het onderkennen van verschillende data mining concepten en fasen met focus op data analyse
- Python: Programmeren in Python.
- Data preparation/ pre processing: Het ontwikkelen van een pijplijn voor data preparatie en preprocessing.
- Predictive Analysis/ Forecasting: Toepassen van inference (predictive) statistiek.
- Machine Learning: Toepassen van Machine Learning.
- Afsluitende opdracht: Uitvoeren van een Data Science project

Diploma

Na het succesvol afsluiten van deze opleiding ontvang je een diploma van IT Academy Noord-Nederland.



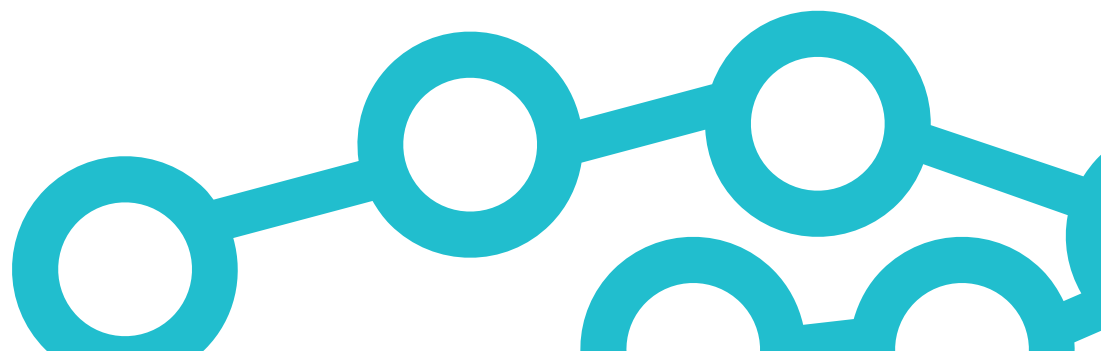


Naam	Design Thinking
Duur	1 dag
Algemene omschrijving	Design Thinking is een actiegerichte methode die creativiteit en analytisch denken combineert en empathie centraal stelt: 'Don't analyze, act!'. De wereld verandert snel. Snel leren van problemen en ervaringen wordt daarom steeds belangrijker. Een belangrijk kenmerk van Design Thinking is actie; vandaag beginnen, prototypes bouwen en direct testen. Je maakt kennis met de basiselementen en leert de methode toe te passen.
Inhoud	• Wie is de doelgroep? • Waar heeft de doelgroep behoefte aan? • Creatief ideeën genereren • Ideeën uitwerken • Pitch je plan bij de doelgroep
Afsluiting	n.v.t.
Leeruitkomsten	Aan het einde van de workshop ben je je ervan bewust dat data op vele manieren toegepast kan worden. Een Data Scientist denkt 'out of the box' en kijkt continue hoe data omgezet kan worden naar innovatie binnen het bedrijf.
Studiebelasting	n.v.t.
Docent	Immo Dijkma, MSc – Hanzehogeschool (De Ontwerpfabriek)

Naam	Databases
Duur	3 dagen
Algemene omschrijving	Tijdens de module Databases leer je hoe je data gestructureerd in een database zet, door een beschrijving te geven van het juiste datamodel, een database te bouwen en informatie hieruit te halen.
Inhoud	Modelleren tot de 3e normaalvorm • Definiëren van objecten in de database • Bevragen van de database middels SQL • Het manipuleren van data in de database
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Het modelleren en realiseren van relationele databases • Kennis van en beheersen van de derde normaalvorm • Kennis van en beheersen van Data definition language en SQL
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	HBO-ICT docent Hanzehogeschool Groningen *Docent onder voorbehoud

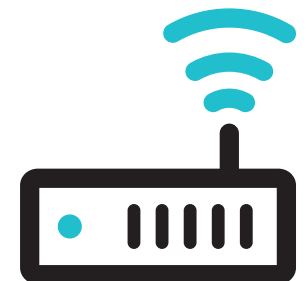
Naam	Data Visualisation
Duur	3 dagen
Algemene omschrijving	Het doel van data visualisatie is om gegevens inzichtelijk te maken door middel van afbeeldingen. Tijdens deze module leer je als data scientist om data via de juiste methode visueel vorm te geven zodat het inzichtelijk wordt voor verschillende doelgroepen.
Inhoud	• Visueel analyseren van data • Data visueel weergeven • Kiezen van de juiste weergave bij het analyseren van verschillende soorten data
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Het weergeven van informatie toegesneden op de doelgroep • Het kunnen hanteren van tooling i.c.m. verschillende soorten visualisaties zoals: staafdiagrammen, 3D diagrammen, etc.
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Andrey Tsyganov (Rijksuniversiteit Groningen) *Docent onder voorbehoud

Naam	Descriptive Statistics
Duur	3 dagen
Algemene omschrijving	Met behulp van beschrijvende statistiek maak je de ruwe data (alle scores van de eenheden op alle variabelen) overzichtelijker. Tijdens deze module leer je hoe je data karakteriseert met behulp van een beperkt aantal kengetallen.
Inhoud	• Beschrijvende statistiek met aandacht voor: • Normaalverdeling • Afwijkingen • Standaarddeviatie • Modus, Mediaan • Betrouwbaarheid
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Het kunnen toepassen van descriptieve statistiek op dataverzamelingen • Het kunnen bepalen van het gemiddelde, de modus, mediaan, spreiding, variantie, percentielen en standaarddeviatie
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Docent HBO-ICT Hanzehogeschool Groningen *Docent onder voorbehoud
Literatuur	Statistiek in 20 stappen (A. Buijs)



Naam	Data Mining
Duur	3 dagen
Algemene omschrijving	Data Mining is erop gericht om verbanden te vinden tussen grote datasets. Tijdens deze module leer je door data te "graven" om met behulp van tools op zoek te gaan naar (statistische) verbanden.
Inhoud	• Wat is data? • Data preparatie • Inhoudelijk weergeven van de data • Ontdekken van patronen met WEKA
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Het ontwerpen en inrichten van het kdd (knowledge discovery in databases) proces • Het onderkennen van verschillende datamining concepten en fases met een focus op data analyse
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Martijn Herber, MSc – Hanzehogeschool (Instituut voor Life Science & Technology) *Docent onder voorbehoud
Literatuur	Data Mining: Practical Machine Learning Tools and Techniques (Ian H. Witten)

Naam	Programmeren in Python
Duur	3 dagen
Algemene omschrijving	Tijdens deze module leren we je de basisprincipes van het programmeren in Python. Python is een scripting taal die veel gebruikt wordt door data scientists. Python wordt gebruikt om data te manipuleren en te analyseren met behulp van diverse tools.
Inhoud	• Basisprincipes van Python • Waarden, variabelen, functies en data structuren • Object georiënteerd programmeren
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent
Leeruitkomsten	Aan het einde van de module beheers je de basis principes van het programmeren in Python. Je kan dan je eerste functies programmeren in Python.
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Tsjerk Wassenaar - Hanzehogeschool Groningen (Instituut voor Life Science & Technology) *Docent onder voorbehoud
Literatuur	Think Python: How to Think Like a Computer Scientist (Allen B. Downey)



Naam	Data preparatie/ preprocessing
Duur	3 dagen
Algemene omschrijving	Data preparatie / preprocessing betreft het bewerken van de data om het passend bij de context te kunnen inzetten voor data science. Je gaat onder andere foutieve data verwijderen, ontbrekende data aanvullen, data combineren en de variabelen transformeren naar een bruikbare formaten.
Inhoud	• Dataset opzetten op basis van de gewenste toepassing • Het opstellen van een Directed Acyclic Graph (DAG) • Het transformeren van datasets naar een, voor data science toepasbaar, bruikbaar formaat
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Foutieve data uit de dataset halen en/of ontbrekende data aanvullen • Data variabelen omzetten naar bruikbaar formaat voor de toepassing • Datasets aggregeren en combineren
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Martijn Herber - Hanzehogeschool Groningen (Instituut voor Life Science & Technology) *Docent onder voorbehoud

Naam	Predictive Analysis
Duur	3 dagen
Algemene omschrijving	Naast het analyseren van data en het zoeken van verbanden hierin kun je op basis van data ook voorspellingen doen voor de toekomst. Tijdens deze module leer je de methodes en technieken om op basis van bestaande data voorspellingen te doen.
Inhoud	• Statistisch programmeren met R of Python voor predictive analysis • Werken met data en tree analysis, logistische regressie en lineaire regressie • Werken met data en k-means clustering • Werken met data en tekst analyse
Afsluiting	Aan het einde van de tweede dag krijgt je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	• Het kunnen toepassen van predictive/ inference statistiek • Het kiezen van de juiste statistische methode behorende bij de situatie • Het beheersen van de volgende statistische methodieken: k-means clustering, lineaire regressie, logistische regressie, tekst analyse en tree analysis
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	George Azzopardi (Rijksuniversiteit Groningen) *Docent onder voorbehoud
Literatuur	Python Machine Learning (Sebastian Raska)

Naam	Machine Learning
Duur	3 dagen
Algemene omschrijving	Met behulp van machine learning (ML) kan je programma's intelligent maken zodat het programma zelf input omzet in output. Hiervoor dient ook het programma getraind te worden. Tijdens deze module leer je wat machine learning is en hoe je het toepast.
Inhoud	- Machine learning toepassen voor zelflerende systemen - Opzetten van neural networks - Opzetten van support vector machines
Afsluiting	Aan het einde van de tweede dag krijg je een praktijkopdracht mee. Op de derde dag worden de praktijkopdrachten besproken en beoordeeld door de docent.
Leeruitkomsten	- Kennis van en toepassen van ML - Het kiezen van de juiste ML methode (neurale netwerken, vector machines etc.) - Het toepassen van Python om ML te creëren
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Herbert Kruitbosch (RUG - CIT) *Docent onder voorbehoud
Literatuur	Python Machine Learning (Sebastian Raska)

Naam	Afsluitende opdracht
Duur	3 dagen
Algemene omschrijving	In de afsluitende opdracht ga je aan de slag met een praktijkopdracht en laat je zien dat je Data Science kunt toepassen in de praktijk. Op de laatste dag presenteer je de onderzoeksresultaten.
Inhoud	Zelfstandig uitvoeren van een data science project.
Afsluiting	De opdracht wordt afgesloten met een presentatie.
Leeruitkomsten	Bij de afsluitende opdracht laat de deelnemer zien dat hij zelfstandig een Data Science project kan uitvoeren waarin alle facetten van de opleiding in terug komen.
Studiebelasting	Naast de drie contactdagen is de studiebelasting 16 uur (8 uur per week).
Docent	Talko Dijkhuis, MSc – Hanzehogeschool (Instituut voor communicatie, media & IT)



Meer werkplezier door zelfontwikkeling

"Het is heel erg boeiend om bezig te zijn met nieuwe ontwikkelingen en alle innovaties rondom Data Science (The Sexiest Job of the 21st Century). Dit geeft mij veel inspiratie voor mijn werk en laat me ook nadenken over de toekomst van mijn vak. Ik zie nieuwe kansen maar ook dat sommige aspecten van onderzoek op termijn zullen verdwijnen. Het is belangrijk om niet stil te staan, maar te blijven ontwikkelen. Zowel voor mijn eigen werkplezier als voor mijn werkgever."

Brede blik, praktijkgericht en goed voor je netwerk

"Buiten de inhoud van de lessen, is het ook heel leuk om deze opleiding te volgen met mensen van diverse bedrijven uit het Noorden. Mensen uit verschillende bedrijfstakken. Dit geeft een leuke dynamiek en laat je vaak ook even verder kijken naar de toepassing van de geleerde technieken. Het is waardevol om op deze manier te kunnen netwerken en gebruik te kunnen maken van elkaars kennis. Sommige connecties zullen in de toekomst nog wel eens van pas komen!"

Janneke Arntzen

Adviseur Statistiek & Onderzoek
bij Gemeente Leeuwarden



"De IT Academy speelt een belangrijke rol in het bij elkaar brengen van opleidingsvraag en -aanbod"

Een mooi totaalbeeld van de gehele "data-keten"

"De data science opleiding geeft een mooi totaalbeeld van de gehele "data-keten" waar een data scientist mee te maken gaat krijgen in zijn dagelijkse werkzaamheden: van het opslaan van de data in database systemen tot aan het ontwikkelen van verscheidene modellen om daadwerkelijk nuttige informatie uit de data te kunnen halen. Interessant is ook om het contrast te zien tussen de traditionele technieken die gebruikt worden en de recent ontwikkelde methodes."

Mogelijkheden voor de toekomst

"Voor het ontwikkelen van de gastransportmodellen maak ik op dit moment niet echt gebruik van de kennis die ik opdoe in de data science opleiding. Ik denk wel dat op termijn hier mogelijkheden liggen, bijvoorbeeld door het toepassen van zelflerende algoritmes, zoals neurale netwerken, om de best mogelijke gastransport routes door het gasnetwerk te kunnen bepalen. Aan de andere kant, voor mijn werk aan Analytics maak ik volop gebruik van de dingen die we leren bij de opleiding."

Erik Bloemsma

Senior Modelleerder bij Gasunie



Wilt u meer informatie?

Telefoon 050 - 595 4400
info@itacademy.nl
www.itacademy.nl
[@IT_Academy_NN](https://twitter.com/IT_Academy_NN)

Zernikeplein 11, 9747 AS,
Kamer D2.35, Groningen

Postbus 30030, 9700 RM
Groningen