

# R i a n n e S c h o u t e n

## Data scientist / Missing data specialist

### PERSONALIA



Rianne Schouten



[rienne.schouten@dpa.nl](mailto:rienne.schouten@dpa.nl)



11-5-1991



Leuvenplein 238  
3584 LM Utrecht



[/Rianne Schouten](#)

### COMMERCIEEL CONTACT



[gabor.fahmel@dpa.nl](mailto:gabor.fahmel@dpa.nl)



06 – 21 85 99 53

### VAARDIGHEDEN

Leervermogen



Vasthoudendheid



Analytisch



Durf



Resultaatgericht



Zelfstandigheid



### OPLEIDINGEN

#### Traineeship

2017 Excellence Program for Data Science in samenwerking met Big Data Republic

#### Education

2017 Research Master Methodology and Statistics for the Behavioural, Biomedical and Social Sciences, GPA: 4.0

### BRANCHES

Gezondheidszorg  
Energiesector  
Transport en Logistiek  
Overheid  
Telecom  
Verzekering  
Onderwijs

### PROFIEL

Vanwege een overvloed aan enthousiasme en doorzettingsvermogen en vanwege mijn leergierige karakter ben ik met name geschikt voor opdrachten gericht op de ontwikkeling van benodigde en geschikte analysemethoden. Op de werkvloer ben ik een gedreven, leergierige en enthousiaste collega. Ik ben direct, open en heb zelfspot. De omgang met mij is nooit saai en brengt altijd verrassing. Naast mijn werk als data scientist, verricht ik onderzoek naar de evaluatie van missing data methoden.

### SAMENVATTING WERKERVARING

#### Statistiek consultant

**Januari 2014 – Mei 2017**

Vele bachelor, master en PhD studenten hebben problemen met het uitvoeren van hun statistische analyses, vaak simpelweg omdat ze er niet goed in getraind zijn. Ik heb voor menig student gefunctioneerd als adviseur, leraar en gids.

#### Onderzoeksassistent bij Universiteit van Utrecht

**Mei 2015 - September 2016**

Als onderzoeksassistent bij diverse cursussen heb ik gezorgd dat de praktische organisatie op orde was. Tevens doceerde en assisteerde ik tijdens computer practica en zorgde dat alle studenten de analysemethoden begrepen en konden toepassen.

### SAMENVATTING ONDERZOEKSERVARING

#### PhD kandidaat Missing Data Research bij Universiteit van Utrecht

**Januari 2016 – Now**

Als PhD kandidaat verricht ik onderzoek naar methoden die omgaan met missing data. Ik stel vragen als: wanneer werkt een missing data methode goed en hoe weten we welke methode we in een specifieke situatie moeten gebruiken?

#### Staff Associate bij Columbia University in the City of New York

**November 2016 - December 2016**

Als onderdeel van de STAN-groep, een team van software developers onder leiding van prof. Andrew Gelman, deed ik divers werk variërend van het reviewen van literatuur tot het ontwikkelen van R-code en het doen van simulaties.

#### Intern at Dutch Institute for Space Research (SRON)

**September 2015 – December 2015**

Bij SRON heb ik me verdiept in Principal Component Analysis en zo de wetenschappers bij SRON geholpen om inzicht te verkrijgen in de voortgang van de ontwikkeling van ruimte-apparatuur.

---

## MEER OVER OPLEIDING

---

### Traineeship

**Excellence Program for Data Science**

#### Periode:

**April 2017 – Januari 2018**

Het doel van het traject is om te leren hoe ik mijn vaardigheden op het gebied van data science in kan zetten om bedrijfsprocessen te optimaliseren en verbeteren. Ik leer, lees en oefen onder andere met:

- Core IT skills: Asana, Python, Linux basics, Code versioning & sharing in GIT, Virtual machines, SQL
- Core data science skills: Pipeline, Statistics, Clustering, Feature engineering, Feature space visualisation
- Applying data science on use cases: Hands on categorization, clustering, regression, Kaggle
- Advanced data science skills: Ensemble, Neural networks, Evolutionary algorithms

Daarnaast ontwikkel ik mij tot een professionele consultant die bedrijven kan helpen in het vertalen van de businesswens naar data science oplossingen. Ik ontwikkel mezelf persoonlijk, als consultant en leer meer over hoe ik energie en tijd het beste kan verdelen.

---

## MEER OVER WERKERVARING

---

### Onderzoeksassistent

**Universiteit van Utrecht**

#### Periode:

**Mei 2015 – September 2016**

Ik gaf les in de volgende cursussen:

Survey Research: Design, Implementation and Data Processing

Survey Research: Statistical Analysis and Estimation

Introduction to Research Methods and Statistics for Social Scientists'

Een aanbeveling is geschreven door Professor Edith the Leeuw:

*Rianne was a first class assistant at our summer school courses. Not only was all material prepared extremely punctual and without errors, she also got very high student evaluations. I can wholeheartedly recommend Rianne!*

---

## MEER OVER ONDERZOEKSERVARING

---

### PhD kandidaat

**Universiteit van Utrecht**

#### Periode:

**Januari 2016 - Now**

Eenieder die wetenschappelijk onderzoek verricht krijgt te maken met missende data. Het ontbreken van gegevens is een probleem voor statistische analysemethoden, aangezien deze methoden meestal alleen werken met complete datasets. Daarom ontwikkelen we methoden die de geobserveerde data optimaal gebruiken om de missende waarden te schatten. Dit noemen we imputatie.

Voordat nieuwe imputatietechnieken gelanceerd kunnen worden, moeten we ze testen. Dit doen we door middel van een simulatiestudie. In zo'n studie simuleren we een complete dataset, vervolgens schieten we daar gaten in, en dan passen we de ontwikkelde missing data methode toe. Door de toepassing te vergelijken met de oorspronkelijke complete dataset, weten we of een imputatietechniek goed werkt.

Ik heb een methode ontwikkeld die op een valide en legitieme manier missende waarden creëert in een complete dataset. Deze multivariate amputatiemethode heb ik geïmplementeerd in een R-functie, welke inmiddels is getest en vrijgegeven is. Ik heb uitgebreide simulaties gedaan om te laten zien in welke situaties mijn functie beter werkt dan de huidige amputatietechnieken. Met de multivariate methode worden de missende waarden gecreëerd zoals je intuïtief verwacht en wenst.

Een onderzoeksvoorstel, mijn publicaties, en een overzicht van mijn huidige activiteiten kunnen gevonden worden op [rianneschouten.github.io](https://rianneschouten.github.io). De ontwikkeling van R en Python functions houd ik bij op mijn [github](#) pagina.

---

## SOFTWARE VAARDIGHEDEN

---

R, Rmarkdown, Package Development	Senior	2014 – 2017
Python	Medior	2015 – 2017
Javascript	Junior	2017 – 2017
SPSS	Senior	2012 – 2017
LateX	Senior	2014 – 2017
Shiny, Bokeh, Highcharts	Junior	2017 – 2017
HLM, Mplus, OpenBugs	Junior	2014 - 2016

---

## TALEN

---

Nederlands	Moedertaal
Engels	Academisch niveau