



## CAREER PROFILE

Data scientist with a versatile background and a strong interest in social and behavioural sciences, finance and artificial intelligence. Experience with project management, data dashboards, machine learning and academic writing. Currently in the last year of my MSc. Following courses in mathematical statistics, multivariate statistics, data mining and statistical learning.



## WORK EXPERIENCE

### Research trainee

Jun 2019 - Aug 2019

Cortical.io, Vienna

Part of academic team investigating the use of Cortical's 'Semantic Folding' in predicting stock and commodity price volatility using close to one billion news articles going back to 2016. Responsibilities include:

- Building and maintaining infrastructure using Python, PostgreSQL, Docker and Google Cloud Services to process historical and current news articles using Cortical's semantic folding algorithm.
- Aid in collecting and preprocessing training data to create custom models using Cortical's semantic folding algorithm.
- Design and implement downstream tests to assess the usefulness of custom models, using e.g. monte carlo simulations, clustering methods and convolutional/recurrent neural networks.
- Provide support to other researchers in terms of statistical expertise and by creating custom docker applications that simplify the access to the custom models.

### Data Scientist

Oct 2018 - current

Leiden University Center for Innovation, The Hague

### Research intern

Jun 2018 - Aug 2018

Cortical.io, Vienna

Investigated the application of Cortical.io's core technologies to finance & investing. Focused on:

- Gaining an in-depth understanding of Cortical.io's core technology and its application to large, textual datasets.
- Develop methods to process, store and query news articles using Cortical.io's semantic folding technology
- Use these methods to develop a prototype for visualizing and downloading historical and live data representing the nature and quantity of news relevant to user-selected stocks, commodities and portfolios.
- Serving as liaison between the research team led by Prof. David Stolin and Cortical.

### Data Analyst

Mar 2015 - Feb 2017

Leiden University Center for Innovation, The Hague

The Center for Innovation facilitates the development of Massive Open Online Courses (MOOCs) at Leiden University. I was responsible for



# Jasper Ginn

Data Scientist / Student



Download resume as pdf



jasperginn@gmail.com



jasperginn.io



jasperginn



JasperHG90



jasper\_ginn



@jhginn

## EDUCATION

### MSc, Methodology & Statistics

Utrecht University  
2018 - present

### BSc, Liberal Arts & Sciences

Leiden University  
2011 - 2014

## PROFESSIONAL DEVELOPMENT

### Practical Time Series Analysis

Coursera (Univ. of New York)  
August 2019 - present

### Deep Learning Specialization

Coursera (deeplearning.ai)  
August 2018 - January 2019

### UT.5.05x: LAFF: Linear Algebra - Foundations to Frontiers

edX (University of Texas)  
May 2018

### DAT201x: Querying with Transact-SQL

edX (Microsoft)  
Apr. 2017

### Machine Learning

Coursera (Stanford University)  
Nov. 2016

### Machine Learning With Big Data

Coursera (UC San Diego)  
Jun. 2016

retrieving, storing and analysing MOOC data with the aim of improving courses and providing feedback to content creators & academics. Co-published several academic articles with researchers and published open-source tools to process and analyze MOOC data.

### Assistant Strategic Analyst (Internship Program)

Sep 2014 - Feb 2015

The Hague Centre for Strategic Studies, The Hague

The HCSS helps governments, non-governmental organisations and the private sector to anticipate the challenges of the future with practical policy solutions and advice. While at the HCSS, I worked on several projects in areas of Big Data, security studies, and international development. Further set up an event-monitoring server that scraped and processed roughly 20.000-50.000 news articles each day, and created a package in R to analyze assertive behaviour among 'great powers' such as Russia, the United States and China.



## PROJECTS

**Phonorm** - Phonetic text normalization using Recurrent Neural Networks.

**NNet** - Implementation of a multi-layered vanilla neural network from scratch in Julia.

**blm** - R library that contains an implementation of bayesian linear regression in R and Julia, as well as many other core features of bayesian analysis such as posterior predictive checks, multilevel bayesian models, MCMC sampling and model evaluation.

**pararius** - Docker application that scrapes the website pararius every couple of minutes to check for new apartment listings. Written in R and Python.

**qualtRics** - Toolkit to retrieve Qualtrics surveys using the API. Original author and creator of the library, which is now managed by Julia Silge.

**FinTxt** - Back-end and front-end infrastructure for analyzing the stock price impact of news articles

**sfutils** - R library that ports and extends Cortical's semantic fingerprinting API to R.



## PUBLICATIONS

Evaluating retrieval practice in a MOOC: how writing and reading summaries of videos affects student learning.

*van der Zee, Tim, et al.*

Proceedings of the 8th International Conference on Learning Analytics and Knowledge. ACM, 2018.

qualtRics: retrieve survey data using the Qualtrics API

*Ginn, Jasper*

Journal of Open Source Software, 3(24), 690,  
<https://doi.org/10.21105/joss.00690> (2018)

Learning about Learning at Scale: Methodological Challenges and Recommendations

*Van der Sluis, Frans, Tim Van der Zee, and Jasper Ginn*

In Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale.

### Introduction to Big Data Analytics

Coursera (UC San Diego)  
Apr. 2016

### Hadoop Platform and Application Framework

Coursera (UC San Diego)  
Nov. 2015

## LANGUAGES

English (Native)

Dutch (Native)

German (Elementary)

ACM, 2017.

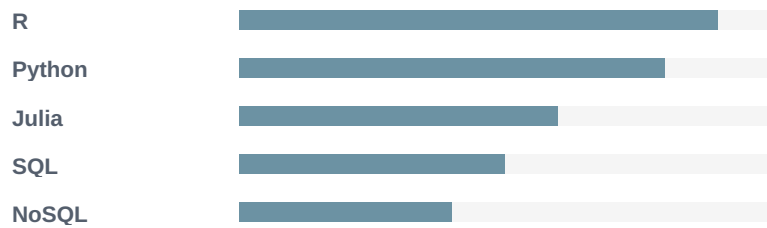
Explaining Student Behavior at Scale: The influence of video complexity on student dwelling time

*Van der Sluis, F., Ginn, J., & Van der Zee, T.*

In Proceedings of the Third (2016) ACM Conference on Learning@ Scale (pp. 51-60). ACM. 2016



## SKILLS & PROFICIENCY



**OTHER:** Shiny, Markdown, Plumber, Docker, Pandas, Keras, Numpy, Pytorch, Linux, Git, Spark, Jekyll, Hugo, MLM, Mplus, SPSS,