

## Education

**Georgetown University**- Washington, D.C.

May 2018

*Masters of Science in Analytics Concentration in Data Science*

**Cornell College**- Mount Vernon, IA

May 2016

*Bachelor of Arts in Economics and Business, and Mathematics, magna cum laude*

## Skills and Tools

- Software and Programming Languages: Python (numpy, scipy, pandas, NLTK), R (dplyr, ggplot2), Java, SQL, Stata, SAS
- Machine Learning: classification, regression, clustering, Bayesian Networks, support vector machines, tensor flow, K-Nearest Neighbor, and random forest
- Statistical Methods: time series, regression models, hypothesis testing and confidence intervals, principal component analysis and dimensionality reduction
- Massive Data Tools: MySQL, Spark, Amazon Web Services, MapReduce, Hadoop
- Data Visualization Tools: GGplot2, Tableau, Plotly, Bokeh, Leaflet, NetworkD3, Shiny
- Development tools: Eclipse, JetBrains PyCharm, Spyder, Jupyter Notebook, RStudio

## Work Experience

**Data Analytics Consultant, World Bank Group**

Oct 2018-Present

- Review, maintain and develop Tableau Dashboards on Climate Co-Benefits and corporate climate change commitments
- Maintain and handle datasets from multiple data sources, including OPCS lending projections and sector/theme coding
- Support regular reports on Climate Co-Benefits such as trend analysis, projections, ad hoc-data requests and pipeline analytics

**Data Analyst, American Urological Association (AUA)**

Jul 2018 – Oct 2018

- Support healthcare study projects with standard data cleaning, data analysis using SPSS and R
- Help maintain the urology quality registry and AUA's Census data
- Assist other departments with ad-hoc analysis of other research projects, data sets and data organizations

**Research and Technology Assistant, Georgetown University**

Jul 2017-May 2018

- Collaborated with the University Information Services' engineering team to grow the Expandable Open Source (EOS) database, a Big Data archive of open source media used in universities researches
- Assisted researchers from different universities to build new robots to crawl designated websites
- Queried using SQL to create monthly reports to present to the director of EOS
- Helped users create regular expression strings to execute searches

**Research Assistant, University of Iowa**

Jun 2016-Aug 2016

- Queried and analyzed 1.5 billion patient visits data from Healthcare Cost and Utilization Project
- Coded in R to manipulate data and connect patients across time and different hospital visits
- Cooperated with medical professionals to identify potential misdiagnosis
- Applied logistic regression and association rules to predict the likelihood of potential misdiagnosis
- Presented the analysis at University of Iowa Hospital and Cornell College's Annual Symposium

**Quantitative Consultant, Quantitative Reasoning Studio, Cornell College**

Sep 2014-May 2016

- Aided students with different types of mathematical and statistical needs
- Collaborated with faculty to deliver statistics software training (R, SAS, Stata, SPSS, Tableau, Excel)
- Trained students and other peer consultants on R, Tableau, and Excel

## Projects

**Forecasting Energy Emission- Tools: R**

Dec 2016

- Built a machine learning model using Support Vector Machine that predicted the energy production with root mean square error of 19%

**Sentiment and Stock Analysis during 2008 Financial Crisis- Tools: Python**

Dec 2016

- Scraped Yahoo Finance, New York Times and Fox News for relevant data
- Analyzed trends of different industries using linear regression
- Performed sentiment analysis on how media reacted to the Financial Crisis