JOSIAH DAVIS

DATA SCIENTIST

EDUCATION

M.A., Statistics – University of California, Berkeley

August 2016 – May 2017

- Project and course work in causal inference, machine learning, optimization and linear modeling.
- Managed workload of full-time program while working as an employee of Slalom.

Technical University of Denmark

August 2010 – December 2010

- Study abroad, coursework in robotics, java, and partial differential equations.

B.S., Mechanical Engineering – University of Maryland

August 2006 – Spring 2010

- Senior design team project was chosen to be the primary teaching example for the school's Mechanical Engineering design textbook.

EXPERIENCE

Slalom Consulting, Data Scientist — San Francisco, CA

May 2015 – Present

- Created a machine learning pipeline that included customer segmentation, time-series analysis, machine learning, variance propagation, and model validation (R – tidyr, purrr, earth, ggplot2, rpart, randomForest, data.table).
- Conducted an exploratory analysis that studied customer bias through the sentiments captured in yelp reviews (Python – NLTK, scikit-learn; R – tm, stringr, openNLP, syuzhet, plyr).
- Designed a new measure of company performance: came up with the idea of using the Gini coefficient to measure the concentration in workload and asset distribution in various business segments (Python – pandas).

Learning Data Science, Founder — San Francisco, CA

August 2016 – Present

- Provided corporate workshops on topics within Data Science including machine learning, linear modeling, and statistical programming (Python – pandas, scikit-learn, statsmodels, matplotlib, numpy).
- Created online e-mail mini-course with free instruction and guidance for breaking into the field of Data Science (learningdatascience.com).

General Assembly, Data Science Instructor — Washington, D.C.

October 2014 – May 2015

- Co-instructor for two iterations of the 66-hour course on Data Science covering the data science pipeline with a focus on supervised and unsupervised machine learning (Python – scikit-learn, pandas, numpy, matplotlib, statsmodels, nltk).
- Rated by students as one of the top 2 Data Science instructors nationwide.

Deloitte Consulting, Data Scientist — Washington, D.C.

February 2012 – May 2015

- Earned the outstanding performance award 2x for client work that included creating and presenting technical deliverables to director-level clients.
- Created tree-based models to predict the probability of rework in benefits-claims process (R – rpart).
- Derived and created a new estimate of latent process complexity (Python – pandas).
- Conducted a program evaluation of a multi-billion technology investment for the Federal government using survival analysis (R)

SlideRule, Data Science Expert Mentor — Washington, D.C.

February 2015 – February 2016

Lockheed Martin, Engineering Analyst — Fort Worth, TX

July 2011 – January 2012

Johnson and Johnson, Supply Chain Analyst (Co-Op) — Somerset, NJ

January 2011 – June 2011

ACCOLADES

What is most impressive about Josiah is his genuine intellectual curiosity and his ability to solve complex problems in clever ways…. He'll make you love what you do just because he does too.
– Dan Henebery, Data Scientist at Deloitte

His intellectual horsepower provided credibility to engage the key stakeholder who had initially voiced concern about the project... Josiah's collaboration quickly turned a doubter into an advocate and extended team member.
– Aaron Hardisty, Consultant at Slalom

I've been impressed at how Josiah continues to grow as a thought leader and partner in our practice. His commitment to the analytics space at Slalom while juggling a full-time masters program is not easy but has given him even more perspective.
– Kyle Roemer, Practice Leader at Slalom

SKILLS

Causal Inference

- Directed Acyclic Graphs
- Potential Outcomes
- Backdoor Criteria
- G-computation formula
- Super Learning
- Targeted Maximum Likelihood Estimation

Machine Learning

- Clustering
- Decision Trees
- Random Forests
- Adaboost
- Ensemble Learning
- Cross-validation
- Natural Language Processing

Statistics

- Linear Modeling
- Model Checking
- Regularization
- General Linear Modeling
- Hypothesis Testing
- Bootstrapping

Programming

- Python - pandas
- Python - scikit-learn
- Python – numpy
- Python - matplotlib
- R - dplyr/tidyr
- R - ggplot2
- R - rpart/randomForest
- R – earth
- R - devtools/roxygen2

josiahjdavis.com | 415-500-5958 | josiah.j.davis@gmail.com