Josh Day

Technical Computing Researcher/Developer

I am a PhD statistician with a passion for technical computing (particularly with the Julia programming language). My niche of computer science/math/statistics allows me to solve difficult problems, then quickly translate whiteboard solutions into efficient programs. I've developed several applications from the ground up to help non-programmers utilize machine learning and natural language processing; Building these tools has made me confident in both R&D and full-stack development roles. I am a strong proponent for open source software and enjoy giving back to the open source community (contributions available on GitHub).

@ hijoshday@gmail.com | @heyjoshday | joshday | (763)458-9630

</> — THINGS I'VE BUILT

— PRODUCTS —

One Click Tuner
Chromatic musical instrument tuner for iOS.

Trendspot.io
Market research via time series analysis on public datasets.

— OPEN SOURCE —

OnlineStats.jl
Single-pass algorithms for statistics and data visualization.

AverageShiftedHistograms.jl
Kernel density estimation for big data.

XKCD.jl
Retrieve metadata for the XKCD web comic.

Other
All other projects are hosted on GitHub.

— CLOSED SOURCE —

NLP API
REST API for Natural Language Processing.

Radar Error Analysis Platform
Toolkit for creating visualizations of sensor errors that took place during Air Force-Orange Flag large force events.

Time Series Dashboard
Front and back-end for time series analysis app based on NLP (sentiment, entities) of multiple news sources.

Healthcare Web App
back-end data preparation and aggregation in Julia.

— WORK HISTORY

Senior Research Scientist

Julia Computing | Oct 2017 - Aug 2019, Oct 2019 - Current

- Built custom Julia software for clients (healthcare and government sectors).
- Consulted with businesses on optimizing Julia code.
- Maintained and contributed to a variety of open source projects.

Data Scientist II

Valassis Digital | Aug 2019 - Oct 2019

- Researched use cases for streaming data models in Ad Tech.

Data Scientist

MaxPoint | May 2015 - Aug 2015

- Developed on-line algorithms for advertising retargeting (logistic and survival models).
- Worked with big data and associated technologies like Hadoop and Spark.
- Fit a lot of Scikit-learn models.

Statistical Development Tester

SAS/JMP | May 2013 - May 2014, May 2015 - May 2016

- Wrote test suites using JMP Scripting Language (JSL) for validating statistical results.
- Redesigned and implemented the JMP Starter user interface.
- Researched new methodologies being considered for JMP platforms.

— EDUCATION

PhD Statistics
NC State University | 2018 | Raleigh, NC

MS Statistics
NC State University | 2014 | Raleigh, NC

BS Math/Statistics
Winona State University | 2012 | Winona, MN

BA Economics/Music
Winona State University | 2009 | Winona, MN

— SELECTED TALKS AND LECTURES

Slides and other materials available at https://github.com/joshday/Talks.

Financial Modeling Using Julia on Large, Streaming Datasets
Julia Computing Webinar March 2020

Scalable Data Analysis with JuliaDB and OnlineStats
JuliaCon 2018

SparseRegression.jl: Linear Models with Sparse Coefficients
JuliaCon 2017

Sorting Algorithms
NC State, ST 758: Statistical Computing (Fall 2017)

Online MM Algorithms for Machine Learning
International Chinese Statistical Association Conference 2016

Julia for Modern Data Analysis
PyData Carolinas 2016

OnlineStats.jl: Online Algorithms for Big and Streaming Data
Joint Statistical Meetings 2016

Online Optimization
NC State, ST 790: Advanced Computing (Spring 2015)

— TEACHING

NC State: ST 312 - Intro to Statistics II
Spring 2017, Spring 2015

NC State: ST 311 - Intro to Statistics
Fall 2016, Fall 2014

NC State: ST 350 - Economic and Business Statistics
Fall 2012

Mentor for Summer Institute in Biostatistics
Summer 2014

— INTERESTS

Machine Learning | *Natural Language Processing* | *On-line Algorithms* | *Numerical Optimization* | *MM Algorithms* | *Penalized Regression* | *Julia* | *Python* | *Swift* | *R* | *Applied Math* | *Linear Algebra* | *Data Visualization* | *Pluto Notebooks* | *Jupyter Notebooks* | *Reproducible Research* | *Web Apps* | *Interactive Dashboards* | *MLOps* | *Signal Processing* | *Linear Models* | *Mixed Models* | *Design of Experiments* | *Time Series Analysis* | *DevOps Monitoring* | *SQL* | *NoSQL Databases*