

David Patrick Lundquist

Applied Scientist and Statistical Software Engineer

📍 Atlanta, Georgia, USA ✉ davidpatricklundquist@gmail.com

[in davidpatricklundquist](#) [Google Scholar](#) [🔊 lunddave](#) [🔗 Professional Website](#)

Citizenship: USA

Education

- PhD University of Illinois, Urbana-Champaign**, Statistics August 2019 - Present
- GPA: 3.79/4.0 ([Transcript](#) and [University Webpage](#))
 - **Interests (with hyperlinks to research):** stochastic processes, time series econometrics, supervised learning, reinforcement learning, forecasts, point/density/quantile forecasts, [forecast adjustment](#), forecast combination, model averaging, econometric analysis of shocks, panel data, volatility modeling, [equity valuation modeling](#), recession forecasting using ML, statistical software engineering, [simulations](#)
 - **Coursework:**
 - GLMs incl. linear, logistic, and Poisson regression; fixed, random, mixed effects
 - supervised learning: LASSO / Ridge, random forest, boosted trees, SVM
 - unsupervised learning: t-SNE, UMAP, K-means/medoids, hierarchical clustering e.g. HDBSCAN, Gaussian mixtures, self-organizing maps, NN matrix factorization
 - Deep Learning: PyTorch, TensorFlow, Keras, JAX, recommendation systems
 - Nonconvex optimization, gradient descent, linear and mixed integer programming
 - Natural Language Processing (NLP) including Word2Vec, Doc2Vec, LDA, Transformers, RAG and fine-tuning, Text2SQL
- MS Rutgers University**, Statistics Jan 2016 - May 2019
- GPA: 3.4/4.0 ([Transcript](#))
 - **Coursework:**
 - ARIMA/GARCH/LSTM/GRU models, forecasting, changepoint detection
 - multivariate analysis including PCA, factor models, canonical correlation analysis
 - analysis of algorithms, data structures, complexity theory; cryptography (RSA, ECC)
 - biostatistics / causal inference, A/B testing, survival models, synthetic control, DiD, matching, propensity score matching, inverse propensity score weighting, switch-back/crossover designs
- BA American University**, Philosophy, *cum laude* Aug 2005 - Dec 2007
- GPA: 3.67/4.0 ([Transcript](#))

Experience

- Amazon**, Alexa Information Analytics, Data Scientist II Intern Santa Barbara, California
Jul 2024 - Oct 2024
- Wrote end-to-end, production-ready statistical software and supporting dashboard hosted on an AWS EC2 instance to identify drivers of topline metrics and test for mix-shifts, largely automating the creation of callouts for Weekly Business Reports
 - Used Amazon Bedrock for prompt-engineering for LLM-generated SQL to improve Claude-powered Text2SQL; employed mixed effect modeling to identify points for improvement
- Amazon**, Payments, Data Scientist II Intern Seattle, USA
Aug 2023 - Dec 2023
- Reduced churn from Amazon Currency Converter using parametric modeling (Cox proportional hazards (CPH) with time-varying covariates), accelerated failure time (AFT), multi-state models, as well as packages PySurvival and Scikit-Survival; provided richer business-

actionable information compared to classification approaches to churn, including (1) probabilities of progression to the early warning signs of churn, (2) seller disbursement dollar amounts at risk, and (3) causal inference on variables driving churn, calling out the troubling features of each seller.

Google, Developer Intelligence (DevIntel), Data Science Intern

Sunnyvale, California
May 2023 - Aug 2023

- Built production-ready metric-monitoring system using Python targeting changepoints and trends by augmenting the [PELT](#) algorithm, with twin goals of spotting unusual activity in developer productivity and supporting dashboards for manual explorations of productivity. Validated system via simulations and used asymmetric Jaccard index to compare inferred changepoints and trend with the ground truth; asymmetry reflects greater danger of false negatives.
- Assisted with A/B tests exploring efficacy of developer tools on productivity, including checks on covariate balance and network effects.

Google, Core Compute Analytics (CCA), Google Cloud, Data Science Intern

Sunnyvale, California
May 2022 - Aug 2022

- Led GCROC project aimed at classifying Borg compute jobs suitable for spatial shifting, with goal of reducing carbon and energy expenditure while limiting transmission costs, yielding a classifier that correctly recalls over 98% of shiftable Google Compute Units (GCU). Furnished ML pipeline in Plex for real-time classification of previously-witnessed Borg jobs as well as novel jobs. GCROC covered [here in The Economist](#) and [here in Bloomberg](#).

Point72, Market Intelligence Intern

New York, New York
Jun 2021 - Aug 2021

- Analyzed streaming providers (Netflix, Disney+, etc) using Python, PySpark, and 30TB of alternative to develop leading indicators, metrics, KPIs, and graphics for use in dashboards available to market analysts and portfolio managers, preparing traders for Netflix's 75% stock price decline from its peak in 2021 to its trough in 2022.

Tsinghua University, Lecturer of Western Philosophy

Beijing, P.R. China
Sep 2010 - Jun 2012

- Instructed students in seminars and composition at pre-eminent Chinese institution.

Publications

Volatility Forecasting Using Similarity-based Parameter Correction and Aggregated Shock Information David Lundquist, Daniel J. Eck [arXiv:2406.08738](#)

June 2024

Presentations

Jane Street's Symposium, January 2021, Presented *Jump Diffusion with a Heterogeneous Poisson Process*, a method for modeling jumps in asset price series.

2024 NBER-NSF Time Series Conference at Penn, September 2024, presented poster for *Volatility Forecasting Using Similarity-based Parameter Correction and Aggregated Shock Information*

Additional Experience And Awards

University of Illinois Center for Innovation in Teaching & Learning

Spring 2020 LIST OF TEACHERS RANKED AS EXCELLENT BY THEIR STUDENTS [↗](#)

Fall 2020 LIST OF TEACHERS RANKED AS EXCELLENT BY THEIR STUDENTS [↗](#)

Technological Skills

Languages: Python, R, C++, Java, SQL, Shiny

Software: AWS and GCP tech stacks, SLURM, Git, Unix, PyTorch, TensorFlow, Keras, JAX, PySpark, pandas, polars, numpy, matplotlib, dash, ggplot

Miscellaneous

Languages: Mandarin (proficient), Hindi (conversational), bahasa Indonesia (intermediate), French (intermediate), limited proficiency in each of Uyghur, Korean, Italian, Spanish, German, Arabic, Persian

Leisurely interests: international travel and development, global affairs and news, macroeconomics and social science, mentoring and tutoring youth, reading, the outdoors, foreign film