

Jonathan Min

jmin111301@berkeley.edu | 925-989-0603 | <https://www.linkedin.com/in/jonathan-min/> | <https://jmin111301.github.io/>

EDUCATION

University of California, Berkeley

Bachelor of Arts in Statistics (Emphasis in Economics) and Applied Mathematics

Cumulative GPA: 3.968

Graduation: May 2023

Selected Coursework: Linear Models (Grad Level) · Time Series Analysis · Statistical Inference · Causal Inference · Techniques of Data Science · Stochastic Processes · Linear Algebra · Real/Numerical Analysis · Financial Economics

EXECUTIVE SUMMARY

Collaborative and multi-disciplined analyst with industry experience working with predictive and mathematical models, statistical inference, and research. A fast learner with adept problem-solving skills, a strong-work ethic, and an amicable attitude with passion for improving the quality of insights and decisions from complex, real-world data.

TECHNICAL SKILLS

- Logistic/Linear Regression, Nonparametric Models, Stochastic Models, Optimization, A/B Testing, Statistical and Causal Inference, Econometrics, Big Data Querying and Visualization, Experimental Design, Bayesian Inference, NLP
- Python (Numpy, TensorFlow, Seaborn), R, SQL, Stata, Tableau, MATLAB, Julia, JavaScript, Microsoft Office

WORK EXPERIENCE

University of California, San Francisco

Data Analyst and Researcher

San Francisco, CA

Mar 2022 – Present

- Optimizing an ensemble model in R to predict patient *C.Diff* risk, cleaning large electronic health records and improving model performance to over 85% precision with techniques like bagging, feature transformation, and cross validation.
- Researching methods model stability and generalizability to create adaptive data subsets, applying these methods by stratifying the training data into different groups to compare model performance across spatial and temporal features.
- Built and analyzed multiple logistic regression model on UCSF clinical COVID-19 data to breakdown the infection timeline into different stages to find associations between social determinants and disease outcomes.

University of California, Berkeley

Undergraduate Student Instructor (Macroeconomic Analysis)

Berkeley, CA

Dec 2021 – Jan 2023

- Served as a student instructor for a class of 500+ students in Berkeley's upper-division macroeconomics course, committing 20+ hours per week to facilitate sections, grade assignments, and refined discussion worksheets for the class.
- Facilitated 2 sections (each with 30 students), devising a curriculum that effectively builds both quantitative and qualitative skills to boost their confidence with macroeconomic concepts, Excel, and analytical techniques.
- Hosted office hours for 5+ hours per week to offer tailored assistance for students, helping students build their economic intuition as well as general skills like quantitative analysis, critical thinking, and Microsoft Excel.

University of California, Berkeley

Research Assistant

Berkeley, CA

May 2021 – May 2022

- Simulating the standard model of life-cycle consumption with hyperbolic preferences (as opposed to the outdated asymptotic preferences model), utilizing R to recreate the model, run trials with different parameters, and compare results.
- Utilized the Method of Simulated Moments to estimate the optimal parameters of the life-cycle consumption model by minimizing the distance from the simulated data moments and model moments.

PORTFOLIO

Project: *Analyzing the Causal Effect of the 1973 and 2020 Designated Hitter Rule (R)*

Mar 2023 – Present

- Utilizing regression discontinuity analysis to analyze the causal effect of the designated hitter rule on batter performance.

Project: *imputation Package (R)*

Jan 2023 – Present

- Publishing a R package for public use to help individuals utilize methods to impute missing data in large and complex datasets, accelerating data processing and potentially improving the qualities of models trained on imputed data.
- KNNimpute: created a function that utilizes the K-nearest neighbors algorithm to identify and aggregate similar data by weighted mean or median to impute missing values.

Project: *UC Berkeley Gym Crowd Predictor (Python)*

Aug 2022 – Dec 2022

- Web-scraped UC Berkeley Recreational Sports Facility and weather data to train a Bayesian ensemble model to predict how full the gym is at some time, optimizing model performance by comparing predictions with updated values and CV.