

Alejandro Rodas

Berkeley, CA • (510) 973-9134 • alex9102017@berkeley.edu •

Linkedin: <https://www.linkedin.com/in/alejandro-rodas/>

EDUCATION

University of California, Berkeley, CA (CDSS)

Aug 2021-May 2025

B.A. in Data Science with Domain Emphasis in Industrial Analytics

RELEVANT COURSEWORK: Data Structures and Algorithms (**DATA STRUCTURES/GITHUB/GRAPHS**); Foundations of Data Science (**EDA, PANDAS, NUMPY**); Principles and Techniques of Data Science (**PIPELINES, SQL**); Probability for Data Science (**STATISTICAL MODELING**); Data Engineering (**ETL, MONGODB, OLAP/DATABASES**); Data Mining and Analytics (**ML ALGS**); Data Inference and Decisions (**ADVANCED STATISTICAL MODELING - GLMS, NONPARAM MODELS**)

SKILLS: Hypothesis testing, logistic regression, IPW, propensity score matching, DAGs, bootstrapping, confidence intervals, statistical inference, clustering, cross-validation, maximum likelihood estimation, data wrangling, feature engineering, causal inference, EDA, pandas, NumPy, seaborn, matplotlib, SQL, Python, A/B testing, geopandas, Tableau, etc.

EXPERIENCE

CDSS Discovery Research Program

Jan 2024 - May 2024

Lead Researcher

- Led a team analyzing federal financial datasets, structuring unorganized records from 350+ centers all over the country.
- Designed and implemented A/B tests to measure impact on key performance metrics.
- Conducted hypothesis testing, fitted both generalized linear models (GLMs) and parametric models to assess performance.
- Presented findings to policy and research teams unfamiliar with statistical and data science terms, guiding decision-making through clear visuals. Relying on libraries like matplotlib seaborn and platforms like Sigma.

U.S. Health Department

May 2024 - Aug 2024

Data Science Intern

- Developed optimization models and A/B tests to evaluate resource allocation, applying causal inference methods, logistic regression, maximum likelihood estimation, and cross-validation to assess treatment effects and model fit.
- Conducted statistical analysis comparing pre and post pandemic data using hypothesis testing, confidence intervals, and bootstrapping to assess shifts in service delivery and risk exposure.
- Created dashboards in BigQuery & Tableau to present risk insights to non-technical decision-makers.
- Collaborated with executive directors to release a federal report coming in late 2025.

UC Berkeley's Financial Aid & Scholarships

Mar 2022 - May 2025

Data Science Intern

- Analyzed multi-year financial aid data using logistic regression, missing data imputation, hypothesis testing, and bootstrapping to uncover patterns in student support usage and aid distribution equity.
 - Applied causal inference methods (e.g., IPW, matching) to evaluate the impact of aid disbursement on student outcomes.
 - Built dashboards in AWS/GCP to visualize trends and statistical summaries for financial officers, aiding policy decisions.
-

HIGHLIGHTED PROJECTS

Research Project: Causal Inference, ML for U.S. Primary Elections

Mar 2025 - May 2025

- Conducted causal analysis using Directed Acyclic Graphs (DAGs), Inverse Probability Weighting (IPW), and Propensity Score Matching to estimate endorsement effects.
- Built and evaluated logistic regression and Random Forest models to predict election outcomes; performed EDA on 2,600+ candidates across race, gender, and incumbency. Presented findings in a research paper.

Predicting Streaming Service Churn with Machine Learning

Oct 2024 - Nov 2024

- Built machine learning models (NN's, Random Forests, Decision Trees) to predict churn rates with 85% accuracy.
 - Used logistic regression to model churn probability; evaluated ROC curves and confusion matrices.
-

LEADERSHIP

Latin American Leadership Society at Berkeley

Jan 2022 - Jun 2022

Forum Committee Vice President

- Organized events for 100+ attendees using data-driven tools to optimize scheduling and participation. Streamlined event logistics through technology platforms, enhancing communication while addressing complex organizational challenges.