Akash V. Iyer

Berkeley, California | 408-607-3557 | akash.v.iyer@gmail.com | LinkedIn.com/akash-v-iyer | akashviyer.github.io

EDUCATION

University of California, Berkeley Graduation: May 2025

B.A. Data Science | Concentration in Business and Industrial Analytics (GPA 3.5)

Relevant Courses: *Principles and Techniques for Data Science, Data Science Foundations, Essential Tools For Data Science, Data Structures, Structure and Interpretation of Computer Programs, Probability & Statistics, Integral and Differential Calculus, Linear Algebra, Urban Data Analytics, Linux System Administration*

Certificates & Awards: *Stanford University: Relational Databases and SQL, Live Oak Varsity Badminton*

WORK EXPERIENCE

Data Science Intern — *California State Water Resources Control Board* Aug 2023 – Present

- Build cross-regional, interpretable machine learning model with socio-economic and parcel data to identify and map lead pipe water service lines, assisting utilities in their removal
- Used AWS based APIs and web scraping to conduct ETL on data, designing final deliverable machine learning models, visualizations, and interactive web application for residents across five U.S. states

Data Science Consultant — *Roots of Success* Aug 2023 – Present

- Mined and analyzed partner organization data to assess education course delivery and effectiveness, delivering impactful visualizations for the organization's website, decreasing user dropoff by 56%
- Analyzed survey data to communicate positive impact of programs on incarcerate behavior and employment to potential partners and donors, increasing average donations by 30%

Data Science Researcher, Vice President of Social — *DataGood @ Berkeley* Jan 2023 – Present

- Conducted an in-depth end-to-end data analysis study on education inequality, with data scraped from various sources
- Implemented a multiple linear regression model to predict the average SAT score of schools based on data aggregated on all schools in California, and presented findings to 80+ club members

Data Analyst — *Data Science Society @ Berkeley* Mar 2022 – May 2022

- Visualized heatmaps, histograms, scatter plots, and conducted hypothesis tests, investigating relationships between meat consumption, wealth, BMI, and life expectancy, across 190+ countries
- Developed 2D geographical visualizations with GeoPandas package and cleaned data with Pandas
- Aggregated data exploration and analysis process into presentation, communicating results to 150+ student symposium

PROJECT EXPERIENCE

Asteroid Classification Jun 2023 – Jul 2023

- Executed end-to-end machine learning project with CI/CD pipelines, modular & organized code, logging, error handling, and interactive Tableau dashboard on joined data from NASA's JPL Database and Planetary Data System
- Deployed trained models on AWS cloud computing infrastructure, integrated with Flask web application framework
- Classified asteroids based on 8 spectral types with variety of ML models including XGBoost, CatBoost, Decision Trees, Random Forest, kNN, and artificial neural networks
- Experimented with techniques such as SMOTE, tuning class-weights, and ensemble techniques under stratified k-fold cross validation to handle class imbalance and improve ROC AUC metric

Measles, Electricity, and Life Expectancy Mar 2022 – May 2022

- Produced a 13-page research paper with box, scatter, and histogram plots, with A/B tests programmed in R, to analyze measles immunization and electricity access's impact on life expectancy for 120+ countries
- Cleaned, statistically summarized, and analyzed World Health Organization data in RStudio, and visualized with the ggplot2 package

Shifting Work and Commuting Patterns Jul 2023 – Aug 2023

- Wrote in-depth research paper on the impact of COVID-19 on San Francisco and Contra Costa County commuters
- Visualized and analyzed data with Excel, R, and Python, and created detailed, interactive maps with CARTO
- Found specific, causal evidence for heightened social inequality during lockdowns from data trends in public transit, traffic, and ACS data

TECHNICAL SKILLS

Programming Languages: Python, R, mySQL & noSQL, Java, Javascript
Machine Learning: Deep Learning, Artificial Intelligence, NLP, LLMs, Data Visualization, Data Cleaning, Statistical Analysis, Statistical Modeling, A/B Testing
Tools & Frameworks: Pandas, NumPy, Tensorflow, Keras, Microsoft Excel, Tableau, Scikit-Learn, Git, Amazon Web Services, Docker, Bash, Seaborn, Matplotlib, RStudio, ggplot, Plotly, Jupyter Notebook, SciPy, Flask, Optuna