

Akash V. Iyer

Berkeley, California | 408-607-3557 | akash.v.iyer@gmail.com | [LinkedIn.com/akash-v-iyer](https://www.linkedin.com/in/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 machine learning model with socio-economic and parcel data to identify and map the presence of lead pipe water service lines, assisting utilities in their removal
- Compile and feature engineered data, design final deliverable machine learning models, interactive visualizations and maps, and findings report for showcase

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 70+ 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 100+ students

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 final paper with box, scatter, and histogram plots, with A/B tests all programmed in R, analyzing measles immunization vs electricity access impact on life expectancy more for 120 countries
- Performed data cleaning, statistically summarized, and analyzed World Health Organization data in RStudio and visualized plots with ggplot2

Shifting Work and Commuting Patterns

Jul 2023 – Aug 2023

- Conducted in-depth group research project on the impact of COVID-19 on San Francisco and Contra Costa County from 2019 to 2021
- Visualized data with Excel, R, and Python, and created detailed, interactive maps with CARTO
- Highlighted data trends across traffic, public transit, and ACS data, supported by previous literature

Question Sincerity NLP

Jul 2023 – Aug 2023

- Classified questions from the forum Quora as sincere or insincere using natural language processing and transfer learning techniques
- Fine tuned pre-trained large language model BERT with Tensorflow and deep learning techniques, achieved validation accuracy of 96.35%

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, 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, SciPy, Tableau, Keras, Scikit-Learn, Git, Amazon Web Services, Microsoft Excel, Seaborn, Matplotlib, RStudio, ggplot2, Plotly, Jupyter Notebook, Deepnote, Flask API, Optuna