Yifan (Tiffany) Hu

Los Angeles, CA | 310-905-1233 |eve9031bruin@g.ucla.edu

EDUCATION

University of California, Los Angeles **Los Angeles, CA**

Bachelor of Science in Statistics and Data Science *GPA: 3.98 / 4.0* *Sep. 2021 – Exp. Aug. 2025*

Relevant Coursework: Optimization, Regression Modeling, Linear Algebra, Text Mining, Statistical Learning, Causal Inference

WORK EXPERIENCES

UCLA Anderson School of Management **Los Angeles, CA**

App Developer *Dec. 2024 – Present*

- Design and implement a full-stack application for chronic disease data analysis, integrating statistical models like logistic regression and survival analysis, and machine learning algorithms such as random forests and gradient boosting to uncover actionable insights.
- Leverage advanced technologies including React for interactive interfaces, Flask for scalable API development, and AWS for secure data storage and real-time processing, ensuring high technical and academic standards.
- Collaborate with a multidisciplinary MBA team, translating complex healthcare datasets into strategic insights through dynamic visualizations and tailored reporting tools.
- Integrate rigorous software engineering methodologies, including iterative prototyping, code optimization, and scalable deployment, to build a robust application utilizing healthcare analytics.

Department of Economics, UCLA **Los Angeles, CA**

Data Research Team Lead *Oct. 2024 – Present*

- Led a team of research assistants and analyzed cognitive impairment trends across multiple countries, leveraging large-scale, multi-country datasets to uncover actionable insights for the International Long-Term Care Project.
- Conducted advanced data manipulation and exploratory data analysis (EDA) in Stata, identifying critical patterns within complex datasets and optimizing statistical models for the research paper.
- Enhanced statistical analysis efficiency by 33% through integrating R and Python into workflows, automating data processing, and streamlining complex data visualization tasks for robust reporting.
- Applied advanced data visualization techniques to transform raw statistical outputs into high-quality, publication-ready figures and tables in Excel, improving data clarity and interpretability for academic publications.

Department of Mathematics, UCLA **Los Angeles, CA**

Reader *Sep. 2022 – Dec. 2023*

- Automated and optimized data workflows using Python, streamlining grading processes and reducing manual workloads by 26%, ensuring efficient data management for 50+ students.
- Analyzed and visualized student performance data with Python and statistical techniques, generating insights that informed curriculum adjustments and improved educational outcomes.
- Developed scalable data solutions by designing data pipelines, ensuring accuracy through quality checks, and integrating computational tools like Python, R, and Excel for impactful reporting.

ACADEMIC PROJECTS

Text-based Financial Market Prediction Analysis Oct. 2024 – Dec. 2024

- Engineered a pipeline for scraping and preprocessing financial news and social media using NLP techniques like TF-IDF and BERT to derive textual features and sentiment scores for predictive modeling.
- Integrated stock price data with sentiment features, deploying machine learning methods such as LSTM networks and Random Forest classifiers to predict market trends based on textual and time-series financial data.
- Evaluated models using metrics like accuracy and F1 score, enhancing performance through hyperparameter tuning, ensemble stacking techniques, and feature importance decomposition.

Bayesian Analysis on Household Dietary Diversity Jun. 2024

- Conducted Bayesian analysis using Poisson and Negative Binomial regression to identify socio-economic factors influencing household dietary diversity in Tanzania, leveraging priors like Lasso for robust models.
- Performed exploratory data analysis and feature selection, optimizing model reliability and contributing valuable insights to nutritional epidemiology.

U.S. Presidential Election Voting Predictors Analysis Jul. 2023

- Analyzed voting behaviors using demographic and educational data, amalgamating k-nearest neighbors, Support Vector Machine, and boosted trees with XGboost and lightgbm engines.
- Emphasized the advantages of model stacking, notably enhanced tuning precision and diminished overfitting risks, attributed to the diverse set of base models utilized.

SKILLS

Technical Skills: Python, R, SQL, Stata, C++, JavaScript, Looker Studio, Advanced Excel

Machine Learning & Programming: A/B Testing, Regression, Time Series, Cloud Computing, Natural Language Processing