CHARLIE HOOSE

(805) 245-6324 — charliehoose@g.ucla.edu — Los Angeles, California, 90024 — linkedin.com/in/charliehoose/

EDUCATION

University of California, Los Angeles — Los Angeles, CA
B.S. in Statistics and Data Science, Minor in Data Science Engineering — June 2026
- GPA: 3.9 / 4.0, 5x Dean's Honor List
- **Relevant Coursework:** Data Structures and Algorithms, Machine Learning, Statistical Modeling and Data Mining, Data Analysis and Regression, Linear Models, Statistical Optimization, Experimental Design, Linear Algebra, Probability
- **Club Involvement:** Data Science Union (roles include Data Scientist, Director of Membership, Executive Advisor, Project Mentor), Statistics Club (General Member)

EXPERIENCE

First Citizens Bank — Pasadena, CA
Data Analyst Intern — June 2025 – Present
- Utilized SQL and Netezza to update philanthropy deposit dashboards by incorporating new data and field calculations, strengthening reporting accuracy
- Executed complex data joins and transformations in Excel using RegEx to match borrower IDs with geocodes, streamlining Low-to-Moderate Income rate analysis and saving $250,000 in loan payments
- Led an intern team to produce data visualizations in Excel and Tableau of historic giving and modeled philanthropy donations, contributing to a holistic bank-wide philanthropic strategy and presented in front of 60+ employees
- Reconciled mismatching donation payments between online tracking systems in Excel using ETL and anomaly detection with 97% accuracy
- Built an automated reconciliation workflow in Excel, saving 93% of manual processing time and improving data integrity.

B. John Garrick Institute for the Risk Sciences at UCLA — Los Angeles, CA
Data Science Research Intern — April 2024 – July 2025
- Debugged and optimized an AI-powered wildfire evacuation prediction application in VSCode to enable faster emergency planning and prediction
- Engineered an NLP pipeline to extract keywords and contextual snippets from wildfire research PDFs, producing feature importances to improve human evacuation model accuracy
- Built an automation workflow to produce structured summary tables of research papers, reducing research synthesis time by over 80%
- Authored three research papers on human and traffic evacuation models, including data analysis, research, and design processes
- Presented a data visualization–driven analysis of 5+ wildfire evacuation software competitors using matplotlib to provide insights on model performance gaps

PROJECTS

Data Science Union Chatbot Project — March 2025
- Designed a RAG-based chatbot to automate recruitment Q&A for DSU, improving accessibility of information for 500+ applicants per cycle
- Engineered a retrieval pipeline with LangChain, converting DSU website content into vector embeddings and ranking results via cosine similarity and max marginal relevance
- Constructed an interactive Streamlit interface enabling real-time user queries and contextualized answer generation using a LLM

Automated Resume Screener Project — April 2024
- Built a K-Nearest Neighbor (KNN) model with NLP-based features to screen resumes with a 15% acceptance rate
- Developed a Python function to parse candidates' job history, extract dates, compute total tenure, and calculate z-scores
- Conducted exploratory data analysis with matplotlib and seaborn to evaluate tenure distributions and feature patterns

Wildfire Acreage Prediction Project — October 2023
- Developed a linear regression model in scikit-learn to predict final wildfire size achieving a mean squared error of 2.8
- Interviewed wildfire and GIS specialists to source datasets used by national wildfire response teams to validate data
- Performed data cleaning in pandas and data visualization in R (ggplot2) to identify and analyze key trends influencing wildfire spread

SKILLS

- **Languages/Domains:** R (ggplot2, tidyverse, dplyr, caret, R Markdown, Quarto), Python (Plotly, Pandas, NumPy, Matplotlib, scikit-learn, seaborn), SQL, C++, Tableau, Excel, LaTeX, Kadince, Netezza, Git
- **Statistical Tools:** Data Cleaning, Data Visualization, Data Modeling, Machine Learning, Statistical Testing, NLP
- **Business & Communication:** Public Speaking, Business Reporting, Storytelling with Data, Communication, Intermediate Spanish, No Limit Texas Hold 'Em Poker