Olivia Nguyen

+1 (925) 316-5058 | olivianguyen@ucla.edu | www.linkedin.com/in/olivianguyen324 | olivianguyen.netlify.app

EDUCATION

University of California, Los Angeles (UCLA) — **Sept 2023 - June 2027**
B.S. Statistics and Data Science, Minor in Data Science Engineering — GPA: 3.7
Coursework: Data Science, Data Analysis & Regression, Machine Learning, Optimization,

EXPERIENCE

Joshi Lab - David Geffen School of Medicine — **Oct 2024 - Present**
Data Undergraduate Researcher — *Los Angeles, CA*
- Analyzing IBS dietary, microbiome, and clinical data in R (tidyverse, glmnet, ggplot2) using linear modeling, multivariate analysis, and exploratory data visualization.

EQ.app — **Oct 2025 - Dec 2025**
Data Engineer - Contract — *Remote*
- Designed a cloud-based ETL pipeline using Airflow to process data from AI chatbots, phone bots, and email bots.
- Built a scalable data infrastructure to enable real-time analytics and performance insights for recruiting agents.

AGENTS HQ (AGENTS.inc) — **June 2025 - Aug 2025**
AI Engineer Intern — *Berlin, Germany*
- Developed an AI agent that automated the finding of 35K+ company domains by designing rule-based semantic matching logic and integrating Beautiful Soup, ValueSERP, and Gemini LLM, achieving ~95% accuracy.
- Reduced runtime from 15 minutes to <1 minute using asynchronous processing.

Mobility Lab - Henry Samueli School of Engineering — **Apr 2025 - June 2025**
Machine Learning Undergraduate Researcher — *Los Angeles, CA*
- Detected ~15K anomalies in 3M+ GPS staypoints in Tokyo using python, pandas, and scikit-learn.
- Built interactive geospatial visualizations with Folium and Plotly to explore patterns in urban transportation strategies.

Zeitios — **Apr 2025 - June 2025**
Software Developer - Contract — *Remote*
- Collaborated on a full-stack AI tutoring platform using React and Node.js to generate interactive courses and videos.
- Integrated OpenAI, ElevenLabs, and MagicSlides APIs to create AI-generated content with voiceovers and slides.

PROJECTS

Skin-tuition: Unmasking Skincare Secrets | *DataResolutions - DataBlog Project*
- Analyzed 1,400+ skincare products to evaluate ingredient functionality, customer feedback, and profitability.
- Built a Scrapy web scraper to collect and data on 1,100+ ingredients and engineered percentile-based features.
- Improved matching of 45,000+ messy ingredient records by 80% using TF-IDF vectorization and cosine similarity.

NBA Game Prediction | *Statistical Modeling and Data Mining Final Project*
- Predicted the outcomes of NBA games in 2023 and 2024 based on past game results with ~70% accuracy using Python
- Engineered features using weighted averages of past game stats, applied L1-regularized Logistic Regression for feature selection, and evaluated models including Logistic Regression, SVM, LDA/QDA, KNN, and Random Forest

How to Win Every Round of Mario Kart | *DataResolutions - DataBlog Project*
- Analyzed 190+ Mario Kart 8 Deluxe world-record builds across 40+ character–vehicle combinations using Python.
- Visualized performance trade-offs using statistical analysis and charts, determining the best setup to win.

SKILLS

Programming languages: Python, R, SQL, C++, Java, HTML/CSS, JavaScript
Tools: Tableau, PowerBI, Apache Hadoop, Airflow, Jupyter, Git, Excel