

# Olivia Nguyen

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

## EDUCATION

<b>University of California, Los Angeles (UCLA)</b> B.S. Statistics and Data Science, Minor in Data Science Engineering Coursework: Data Science, Data Analysis & Regression, Mathematical Statistics, Optimization, Probability, Linear Algebra	<b>Sept 2023 - June 2026</b> GPA: 3.65
--	---

## EXPERIENCE

<b>Joshi Lab - David Geffen School of Medicine</b> <i>Data Research Assistant</i>	<b>Oct 2024 - Present</b> <i>Los Angeles, CA</i>
<ul style="list-style-type: none"><li>Analyzing IBS dietary, microbiome, and clinical data in R (tidyverse, glmnet, ggplot2) using linear modeling, multivariate analysis, and exploratory data visualization.</li><li>Contributed to research on estrogen-metabolizing gut bacteria across menopausal stages and menstrual cycle phases in IBS patients and healthy controls; poster presentation through the American Gastroenterological Association.</li></ul>	

<b>EQ.app</b> <i>Data Engineer - Contract</i>	<b>Oct 2025 - Dec 2025</b> <i>Los Angeles, CA</i>
<ul style="list-style-type: none"><li>Designed a cloud-based ETL pipeline using Airflow to process data from AI chatbots, phone bots, and email bots.</li><li>Built a scalable data infrastructure to enable real-time analytics and performance insights for recruiting agents.</li></ul>	

<b>AGENTS HQ (AGENTS.inc)</b> <i>AI Engineer Intern</i>	<b>June 2025 - Aug 2025</b> <i>Berlin, Germany</i>
<ul style="list-style-type: none"><li>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.</li><li>Reduced runtime from 15 minutes to &lt;1 minute using asynchronous processing.</li></ul>	

<b>Mobility Lab - Henry Samueli School of Engineering</b> <i>Machine Learning Undergraduate Researcher</i>	<b>Apr 2025 - June 2025</b> <i>Los Angeles, CA</i>
<ul style="list-style-type: none"><li>Detected ~15K anomalies in 3M+ GPS staypoints in Japan using Python, pandas, and scikit-learn.</li><li>Built interactive geospatial visualizations with Folium and Plotly to explore patterns to forecast where, when, and how people travel, enabling urban transportation strategies.</li></ul>	

<b>Zeitios</b> <i>Software Developer - Contract</i>	<b>Apr 2025 - June 2025</b> <i>Los Angeles, CA</i>
<ul style="list-style-type: none"><li>Collaborated on a full-stack AI tutoring platform using React and Node.js to generate interactive courses and videos.</li><li>Integrated OpenAI, ElevenLabs, and MagicSlides APIs to create AI-generated content with voiceovers and slides.</li></ul>	

## EXTRACURRICULARS

<b>Data Resolutions at UCLA</b>   <i>Data Learn Co-Head</i>	<b>Apr 2025 - Present</b>
<ul style="list-style-type: none"><li>Creating and delivering in-person lectures, creating practice materials, and covering data science topics beyond UCLA's standard curriculum including Python, SQL, and Tableau for over 200 students each quarter.</li></ul>	

<b>UCLA Statistics Club</b>   <i>STAND/DEI Chair ('25-'26), Career Development Chair ('24-'25), DataFest Events Committee ('23-'24)</i>	<b>Oct 2023 - Present</b>
<ul style="list-style-type: none"><li>Leading inclusivity-focused events, community service, and diversity initiatives with the Statistics Department.</li><li>Presented 2-3 career development workshops each quarter on career development, such as panels and resume-building.</li><li>Coordinated workshops through collaboration with UCLA faculty to support and guide participants at DataFest.</li></ul>	

- Built a user-friendly web platform facilitating roommate search using JavaScript, HTML/CSS, React, Figma, and Google Firebase
- Developed skills in software engineering and product management through weekly curriculums.

## PROJECTS

---

**Skin-tuition: Unmasking Skincare Secrets | Data Resolutions - Data Blog Project**

- Analyzed 1,400+ skincare products to evaluate ingredient functionality, customer feedback, and profitability.
- Built a Scrapy web scraper to collect 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 | Data Resolutions - Data Blog 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.

**California Food Deserts | Digital Humanities Final Project**

- Analyzed 2.5K California census tracts to identify gaps in food access and created interactive Tableau visualizations and a WordPress site to communicate actionable insights to stakeholders.

**IMDb Movie Review Analysis | National Student Data Corporation at UCLA**

- Analyzed 50,000 IMDb movie reviews using NLP techniques such as sentiment analysis, named entity recognition, tokenization, lemmatization, and TF-IDF in Python.
- Predicted positive/negative reviews, sentiment towards directors, actors, actresses, and movie genres with ~60% accuracy.

## ABSTRACTS & PRESENTATIONS

---

Mahurkar-Joshi, S., Labus, J. S., Dong, T., Jacobs, J. P., **Nguyen, O.**, et al. Differential abundance of estrogen-metabolizing gut bacteria across menopausal stages and menstrual cycle phases in irritable bowel syndrome patients compared to healthy controls. Poster presentation, *American Gastroenterological Association (AGA), Microbiome in Gastrointestinal and Liver Diseases Track*, 2025, Los Angeles, CA.

## CERTIFICATIONS

---

**IBM Data Science Professional Certificate**

- A 12-course Coursera certificate in Python, data analysis, machine learning, SQL, data visualization, and other data science skills with hands-on projects and model deployment.
- Gained proficiency in Python libraries such as Pandas, NumPy, and Matplotlib for data manipulation and visualization.

## SKILLS

---

**Programming languages:** Python, R, SQL, C++, Java, HTML/CSS, JavaScript

**Tools:** Tableau, PowerBI, Apache Hadoop, Airflow, Jupyter, Git, Excel

**Hobbies:** Baking (cinnamon muffins), cooking (stir fry), hiking, exploring