

Jialin Jiang

(408) 816-4693 | jialinn09@berkeley.edu | [Linkedin](#)

Berkeley, California, 94720

Education

University of California, Berkeley

May 2026

B.A. Molecular Cell Biology and B.A. Data Science (Business & Industrial Analytics Track), GPA: 3.81

Relevant Coursework: CS 61B (Data Structures), Data 100 (Principles and Techniques of Data Science), UGBA 141 (Production and Operations Management), Econ 2 (Introduction to Economics), ENVECON C118 (Introductory Applied Econometrics Using R)

Upcoming Coursework: Data 101 (Data Engineering), CPH C100 (Foundations for Computational Precision Health), MCB 120 (Therapeutics Discovery and Development), Stat 156 (Causal Inference)

Experience

Data Science Varsity - Berkeley, CA

September 2024 – December 2024

Women in Big Data Student Team Lead

- Scoped project timeline, aligned team on key milestones, and ensured seamless access to technical resources for integrating **Knowledge Graphs** and **Retrievers** within the **RAG chatbot architecture**.
- Led exploratory data analysis (EDA) to optimize for product goals and scalability, while researching industry best practices for chatbot development to manage 100,000 Wikipedia articles.
- Developed a [reproductive health chatbot](#) providing reliable, tailored advice on birth control, STIs, abortion, and consent, using a retrieval-augmented generation framework. **Secured 2nd place** out of 8 teams, although the project was not completed due to time and resource constraints.

Develop For Good - San Francisco, CA

May 2024 - August 2024

Building Beats Student PM

- Scoped and maintained project milestones by drafting a comprehensive **Product Requirements Document (PRD)**, conducting **UX audits**, revising **information architecture**, and troubleshooting design challenges.
- Led a team of **8 designers** in the **redesign and implementation** of the **Building Beats** website on **WordPress**, enhancing content accessibility for underserved youth and creative arts instructors, with the potential to impact **10,000+ users**.
- Documented the **end-to-end design process** in a detailed [project case study](#) to serve as a reference for future initiatives and stakeholders.

CITRIS Health - Berkeley, CA

February 2024 - present

Student Project Specialist

- Assisted in streamlining **data cleaning and management** for a database of **500+ age-tech companies**, narrowing the focus to **293 high-potential candidates** for potential outreach and partnerships supporting CITRIS's collaboration with Germany on entrepreneurship initiatives.
- Compiled information to support future analysis of **trends in pain points, technologies**, and **user groups** within the **age-tech landscape**, with plans to culminate in a **white paper** as a foundational resource for the age-tech community.
- Supported **stakeholder engagement** by maintaining **funding/grant trackers**, updating **CITRIS Health websites**, conducting **literature reviews** on health tech topics, and curating content for **event promotions**.

Projects

Golden Bear Cafe Operational Improvement Analysis

December 2024

- Awarded "Most Creative and Rigorous Project"** out of 15 teams for delivering actionable insights on queue time optimization, demand forecasting, quality control using p-charts, and value creation strategies

 UGBA 141 Final Project Report.pdf

- Directed **team brainstorming sessions** and aligned focus on **key operational metrics**, ensuring cohesive and effective data analysis despite limited access to internal data sources.

Housing Price Prediction – Data 100 Project

October 2024

Cook County Housing Property Valuation Data

- Conducted exploratory data analysis (EDA) using **Pandas** and **Matplotlib** to identify correlations between property features and housing prices, applying **Tukey's bulging diagram** for axis transformations.
- Built a data pipeline in **Python** utilizing **NumPy** and **Pandas** for feature extraction, outlier removal, and one-hot encoding, optimizing data for machine learning models.
- Developed a regression model using Scikit-learn, achieving an **RMSE of 180K** for both training and test sets, significantly improving on the baseline model, which had an RMSE over 280K using only the bedroom feature, while analyzing socio-economic implications of over-valuations in lower-income neighborhoods.

Skills & Interests

Languages: Mandarin (native), Cantonese (native), English (native), German (conversational);

Skills: Java, Python, SQL, Excel, Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn, Figma, Tableau, ArcGIS

Interests: hip-hop dance, traveling, blogging, neuroscience, health tech, HEOR, UX/UI, product operations/strategies.

DataCamp Certificates In Progress: Data Scientist Certificate in Python, Data Analyst Certificate in Tableau