

Kelly Xie

San Jose, CA 95126

txie@scu.edu, 949-992-6645, <https://www.linkedin.com/in/tingyu-xie-kelly/>

EDUCATION

Santa Clara University, Leavey School of Business
Master of Science, Major in Business Analytics

Santa Clara, CA
December 2025

- Relevant Course Work: Relevant Course Work: Generative AI, Machine learning with Python, Marketing Analytics with R

University of California, Irvine
Bachelor of Arts, Major in Quantitative Economics

Irvine, CA
March 2024

- Relevant Course Work: Econometrics, Probability Theory, Data Programming, Statistics

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, MATLAB

Tools & Platforms: Tableau, Power BI, Excel, AWS, Google Collaboratory, Canva

Analytics: EDA, Data Cleaning, Machine Learning, Predictive Modeling, Clustering, Visualization

EXPERIENCE

Armanino LLP (Practicum)

San Jose, US

Data Analyst Intern

Jan 2025 – Jun 2025

- Developed a Generative AI-powered chatbot to analyze ESG reports, integrating Streamlit, Pinecone, and OpenAI APIs to support vectorized retrieval, semantic search, and regulatory benchmarking.
- Built a structured pipeline that automated PDF parsing, text chunking, and metadata tagging, ensuring high data accuracy.
- Implemented a controlled document upload and classification system, allowing users to categorize and index ESG reports into a database with real-time feedback and audit transparency.
- Automated ESG summary generation and cross-document comparison using prompt-driven logic, producing structured reports with analytical tables.
- Designed and published a user-facing web application to support internal teams with ESG document uploads, analysis, and benchmarking workflows.

First Capital Securities

Hang Zhou, China

Data Analyst Intern

Jun 2023 - Aug 2023

- Collected and standardized key financial and operational data from 50+ publicly listed companies using Excel templates to support internal investment evaluation.
- Analyzed company financials, credit ratings, and operational indicators to identify performance trends, industry risk clusters, and early warning signals.
- Interpreted reporting patterns to highlight high-volatility sectors and potential investment risks, contributing to portfolio monitoring discussions.
- Gained experience with investment screening workflows and data driven assessment methods for fixed-income and equity products.

PROJECTS

Machine Learning: Yelp Restaurant Rating Prediction

- Analyzed over 150,000 restaurant records and 6 million user reviews from Yelp dataset to identify key factors influencing customer ratings.
- Merged and filtered review and business records to ensure high data integrity and minimize missing or duplicate entries.
- Executed model development workflow, including data cleaning, feature engineering, sentiment analysis, and model training using regression-based algorithms with cross validation and regularization to identify top predictors.
- Achieved a high R^2 of 0.9764, with key drivers such as sentiment score, total number of reviews, and geographic effects.
- Delivered business insights for restaurant operators, supported by data-informed strategies.

SQL & Database: Travel Booking System

- Built a MySQL relational database to manage travel bookings, destinations, hotels, flights, and customer data with enforced primary and foreign key constraints.
- Designed and executed 10+ analytical SQL queries to generate insights on top destinations, customer spending, seasonal trends, and destination-level revenue performance.
- Applied window functions, aggregates, and common table expressions to rank destinations, calculate cumulative revenue, and analyze booking patterns over time.
- Built the Flask backend to support query execution and resolved schema issues to ensure accurate data flow.