Christopher Lum

lum@ucsd.edu • (425) 295-8028 • San Diego, CA

ch-lum.github.io • github.com/ch-lum • linkedin.com/in/ch-lum

PROFESSIONAL EXPERIENCE

UC San Diego   **La Jolla, CA**
Research Assistant – Visualizing Pandas Tables   *Apr 2024 – Present*
- Develop truncating code for Pandas Tutor backend to raise memory limit from 1 MB to 100 MB, significantly enhancing its capability to handle larger datasets
- Research how people reason about data tables and how they get displayed in notebooks to understand data comprehension and presentation

San Diego County Taxpayers Association   **San Diego, CA**
Data Science Intern   *May 2023 – Sep 2023*
- Developed a ProPublica scraper for 990 forms with BeautifulSoup, scraped documents with 99% accuracy with error flagging for inaccuracies and increased the size of our dataset by 20% over the prior method of manually scraping the data
- Created data visualizations with Altair based on SANDAG public data for the organization's social media, focused primarily on city budgets/spending
- Digitized San Diego County's city budgets from PDF to XML and prepared groundwork for future financial and meta-analysis

EDUCATION

University of California, San Diego   **La Jolla, CA**
Bachelor of Science in Data Science, Minor in Business   *Grad: June 2025*
9500 Gilman Drive, La Jolla, CA, 92093

- **Cumulative GPA**: 3.86/4.00
- Relevant Coursework: Probabilistic Modeling and Machine Learning, Data Visualization, Web Mining and Recommender Systems

PROJECTS

Rocket League Possession Model   *https://ch-lum.github.io/projects/rl_possession.html*
- Scraped over 10 GB of match data for video game "Rocket League" to build an XGBoost model to predict the team that scores next as a proxy for possession
- Training dataset came to over 300,000 rows

Soccer Domestic League Possession Analysis   *https://github.com/ch-lum/soccer_possession*
- Scraped 10 years of match data from Football Reference for 24 different domestic soccer leagues to compare leagues to each other with statistical methods such as KS and ANOVA tests

Facebook Comment Sentiment Analysis   *https://ch-lum.github.io/comment-vis/*
- Implemented a Hugging Face sentiment analyzer on a comment section from a Biden Facebook post
- Visualized data with Altair for a static image and with D3.js and Svelte for an interactive website
- Highlighted as best project in class of 150 students by instructor

San Diego Parking Visualization   *https://ch-lum.github.io/sd-parking/*
- Developed a webpage with D3.js and Svelte to explore the use of parking meters in San Diego
- Highlighted as top 5% of projects in class of 150 students