

# RAY YI

chirayyi.github.io · linkedin.com/in/ray-yi · chirayy@uci.edu · 510-3246621

## EDUCATION

### University of California, Irvine

Irvine, CA (August 2016 - June 2020)

- Major: B.S., Data Science, Minor: Management, GPA: 3.5, Dean's Honours
- Relevant Coursework: Programming in Python I-III, Data Structures, Analysis of Algorithms, Machine Learning, Database Management, Statistical Methods for Data Analysis, Bayesian Statistics, Information Visualization

## SKILLS

**Programming** Python, R, SQL, C++, HTML/CSS/JS

**Frameworks/Applications** Scikit-Learn, Tensorflow, Tableau, RShiny, Alteryx, MongoDB, Git, Unix, Latex

**Languages** Fluent in English and Mandarin; Conversational Proficiency in Spanish

## EXPERIENCE

### Pacific Life

Data Science Intern

Aliso Viejo, CA (June 2019 - August 2019)

- Built RShiny dashboard that visualizes and ranks policy abnormality from >100,000 policynumbers using PCA, K-Medoids Clustering, and Manhattan Distance, data wrangling done in SSMS, Alteryx, and R
- Led weekly meetings to discuss feature engineering and app design/functionality for insurance disbursement fraud

### Wing AI Technologies

Data Science Intern

Irvine, CA (October 2018 - May 2019)

- Identified KPIs such as user browsing habits, wrote respective mySQL queries for data visualization with Tableau, taught Tableau Data Storytelling, increasing query output by 60%
- Scraped web pages of competing prices with BeautifulSoup and merged data with Alteryx for pricing strategy

### UCI Center in Writing & Communication

Peer Tutor

Irvine, CA (December 2016 - May 2018)

- Suggested ways to improve students' textual analysis logic and received annual evaluation of 9.3/10
- Prioritized teaching the principles behind grammatical and logical errors so students can apply in future papers

## SELECTED PROJECTS

### Presidential Candidate 2020 Sentiment Analysis

(March 2020)

- Jupyter Notebook that reveals Twitter perception of Presidential Candidates using Tweepy and PostgreSQL
- Designed database schema for Tweets, streamed Tweets, and conducted Sentiment Analysis with NLTK and sklearn

### Police Killings Data Analysis

(May 2018)

- Jupyter Notebook that visualizes and interprets impact of variables involved with police killings in the U.S.
- Utilized Pandas, Matplotlib, and Numpy to generate graphs and data frames to reveal patterns in shootings

### KinectBeats - 1st Place HackTech 19 - Python Motion-to-Music Platform

(March 2019)

- Music platform implemented with Kinect API to map hand gestures to sounds coded with Sonic Pi in Python
- Improved gesture recognition precision and accuracy by tweaking existing gesture images for a more diverse dataset

### Spell Checker

(October 2018)

- C++ spell check and suggestion program, uses hash map for text input and hashed index value assignment

## ACTIVITIES

### Undergraduate Society of Statistics and Data Science

Resource Acquisition

Irvine, CA (February 2017 - Present)

- Invite employees from tech companies to educate the club of data science application across different domains, lead data wrangling, ML, and linear regression model selection workshops, increasing active members by over 40%
- Gathered workshop material and chose representative live coding examples suitable for audience of all levels

### Boy Scouts of America

Eagle Scout, Senior Patrol Leader, Eagle Project Leader

(September 2014 - September 2016)

- Led a group of 10 scouts to construct a compost bin for the local church, church of Our Lady of Guadalupe
- Planned and led weekly meetings for over 100 active scouts as Senior Patrol Leader, increasing attendance by 20%

### UCI Men's Varsity Table Tennis Team

Coach, Player

(September 2016 - Present)

- Competed in 2019 NCTTA iSET Nationals finishing 4th, led team to 2 western regional conferences finishing 3rd