

Ryan Chui

2564 Walnut Blvd, Walnut Creek, CA 94597 | (925) 683-9874
Email: rchui773@berkeley.edu | GitHub: <https://github.com/ryanchui/Portfolio>

Objective

To pursue graduate studies in Statistics and Data Science, leading to a career in Data Scientist. I am interested in Data Analytics, Natural Language Processing and Machine Learning

Education

University of California, Berkeley, B.S. in *Applied Mathematics*

Fall 2016

Relevant Courses: Data Structures, Machine Structures, Time Series Analysis, Linear Algebra, Data Science, Probability

Certifications

SAS Certified Base Programmer for SAS 9

SAS Certified Advanced Programmer for SAS 9

Technical Skills

Programming: Python, C, C++, Java, Matlab, R, SQL, SAS, Git, HTML, CSS, JavaScript, x86 Assembly

Databases and Office Application: MySQL, MS Access, MS Excel, MS Word, MS Power Point

Packages and Tools: Pandas, NumPy, SciPy, Scikit-learn, matplotlib, ggplot2, MS Visual Studio, Eclipse IDE

Professional Experience / Leadership

Arbor Tutors, *Mathematics & Statistics Tutor, San Francisco, CA*

Jan'17 – Present

- Emphasized interactive learning by asking students questions instead of routinely presenting the solutions when solving problems, ranging from AP, SAT, GRE exams and courses up to Linear Algebra

UC Berkeley Undergraduate Research Apprentice Program, *Data Research Analyst, Berkeley, CA*

May'16 – Sep'16

- Applied data mining techniques to unstructured datasets such as time series, sequence, and spatial datasets
- Led a data team of 3 students to analyze precipitation data, identified trends and determined the implications for practices in core areas such as rainfall extremes, climate change, and global warming
- Applied Bayesian approaches to univariate threshold selection with justification to determine suitable thresholds for extreme value models using Extreme Value Theory and Generalized Pareto Distribution
- Performed logistic regression, clustering and optimization methods to estimate rainfall in Colorado using R
- Used Shiny package to display 57 weather stations in Colorado from MySQL database on a Google Map using the Google Maps JavaScript API

Amazon Inc., *Online Warehouse Manager, Berkeley, CA*

Sep'13 – Dec'15

- Managed 1500 pair of Eyeglasses, addressed shipment of product and issued refund of payment to customers
- Ensured customers receive a delivery confirmation tracking number and resolved customer concerns promptly to maintain satisfaction

Key Academic Projects

Forecasting with Google Trends

Nov'16 – Dec'16

Course Project in STAT 153 with Prof. Aditya Guntuboyina

UC Berkeley, CA

- Achieved stationary time series by differencing a random walk model and interpreted the statistical significance of autocorrelation values using Ljung-Box test
- Performed ARIMA model selection, model fitting and time series forecasting of a particular Google query in R and tuned model parameters based on criteria (AIC, BIC, CV, MSE) reaching 88% accuracy

Classification from the Million Song Dataset

Apr'16 – May'16

Course Project in CS 8 with Prof. John Denero

UC Berkeley, CA

- Built classification models (k-NN, Feature Selection) in Python to classify 5000 words appear in a song's lyrics into 2 categories: hip-hop, country, and determined how frequently certain words appear in a particular song
- Produced the predictions and classification errors by 5-fold cross validation to tune model parameters so that resulting 85% overall prediction accuracy on the random forest decision tree

Predicting Location via a Statistical Indoor Positioning System

Apr'15 – May'15

Course Project in Stat 133 with Prof. Deborah Nolan

UC Berkeley, CA

- Built a model for the location of a device as a function of the strength of the signals between the device and each access point using training data
- Examined 150000 measurements of signal strength recorded at 6 stationary WiFi access points (routers) within a building at the University of Mannheim and developed a statistical IPS
- Accomplished to compute Euclidean distances between vectors of 6 signal strengths and collapse 110 signals for each position orientation combination into one record by using Cross-Validation