ALAINA LIU

Orange County, CA | https://github.com/alainaliu
(562) 298-3605 | alainaaliuu@gmail.com

EDUCATION

BACHELOR OF SCIENCE STATISTICS AND DATA SCIENCE

University of California, Santa Barbara

EXPECTED GRADUATION: JUNE 2024 / GPA: 3.92

- Relevant Coursework: Statistical Machine Learning, Data Structures and Algorithms, Time Series, Regression Analysis, Design of Experiments, Data Science Concepts and Analysis, Statistical Data Science, Probability and Statistics
- Relevant Employment: Website chair of Data Science Collaborative: Built and managed website through Quarto using HTML, JavaScript, and CSS

WORK EXPERIENCE

DATA ENTRY CLERK

Rubber-Cal, Inc. | Fountain Valley, CA

JULY 2023 – SEPTEMBER 2023

- Worked with Microsoft Excel, NetSuite, and various back-ends of online storefronts to manage and organize large datasets.
- Found and analyzed sales trends and decided order acceptances based on freight cost-benefit analysis.
- Coordinated with three warehouses daily to order inventory and handle paperwork.

ASSISTANT STUDENT COOK

UCSB Dining Commons | Santa Barbara, CA

2021 - PRESENT

- Worked front and back of house, helping chefs make, prepare, and serve food.
- Effectively communicated with staff and students and finished tasks in a timely manner.

SKILLS & PROJECTS

R (ggplot2, tidyverse, dplyr, shiny, tidyr), Python (pandas, numpy, matplotlib, Altair, sklearn, seaborn, scipy, cvxpy), C++, SQL, Microsoft Excel, SAS, Quarto, HTML, Javascript, CSS, LaTeX, shell

SANTA BARBARA WEATHER (2023)

Time series project looking at past data of Santa Barbara weather from 2015-2022 with over 2500 observations. Looked at interactions and correlation between a variety of climate factors like daily average temperature, humidity, and wind speed. Validated stationarity of the data and performed transformations. Fitted two methods, SARIMA and ARMAX, to find the best model to forecast future predictions. Compared forecasted values to real data collected during the same period and evaluated models.

NETFLIX RATINGS PREDICTION (2023)

Used machine learning tools to build a model to predict Netflix genres for movies and shows given predictors from a dataset of over 600 observations. Models included k-nearest neighbors, elastic net regression, pruned decision trees, random forest, and boosted decision trees to find the best fit model. Assessed the performance of each model with ROC metrics and analyzed results and tradeoffs.

BEST BANG FOR YOUR BUCK (2023)

Gathered observations on nutrition information data of 17 popular fast food restaurant chains and generated analytical reports with graphs and visualizations, exploring best price worth for calories and serving size, and best predictors of popularity. Built a shiny app for user interaction, letting users directly visualize the best restaurants for their needs and concerns. Found out the best restaurants for different metrics: best value, best protein, and most overall health.