Rhiann Zhang

rhiannzhang@gmail.com | (657) 217-7557 | Fullerton, CA | github.com/rhiannz

ANALYTIC SKILLS

- **Languages and Software**: Python, R, SQL, Git, HTML, bash scripting, C++, Markdown, Power BI, LaTeX
- **Libraries**: Pandas, NumPy, Matplotlib, seaborn, Plotly, scikit-learn, SciPy, PySAL, tidyverse (ggplot2, dplyr)

EDUCATION

University of California, Berkeley Aug 2021 - May 2022

Master of Arts in Statistics GPA: 3.81

- Relevant Coursework: Applied Machine Learning, Statistical Computing, Advanced Statistics, Advanced Probability, Linear Models, Urban Informatics and Visualization

University of California, Riverside Sept 2017 - Mar 2021

Bachelor of Science in Mathematics GPA: **3.97**

- **Chancellor's Honor List** & **Dean's Academic Distinction Award**
- Relevant Coursework: Numerical Analysis, Intermediate Analysis, Applied Linear Algebra, Multivariable Calculus, Discrete Structures, Number Theory, Mathematics Education

SELECTED PROJECTS

Geospatial Explanatory Modeling of U.S. Drug Epidemic Jan 2022 - May 2022

- Procured and cleaned over 500 datasets and performed backward stepwise feature selection in tandem with OLS regressions to determine county features that best explain drug overdose rates
- Conducted geospatial-significance analysis, achieved normalized RMSE of 0.078, and developed recursive estimation procedure to fill in map previously missing overdose rates
- Effectively collaborated with a team of four undergraduate students and two graduate students to conceptualize, organize, develop, and present all findings of research

Analysis of AP Exam Participation Feb 2022 - May 2022

- Implemented linear regressions, random forests, and entity-demeaned fixed effect modeling to explore school characteristics most related to AP exam participation using datasets with over 100,000 observations
- Effectively communicated findings by creating interactive Plotly maps across time and GitHub Pages website

genetic-selectR Package Development Nov 2021 - Dec 2021

- Developed an R package implementing a genetic algorithm that efficiently selects the most optimized combination of covariates from multidimensional datasets for generalized linear modeling
- Constructed and performed rigorous testing on several dynamic functions carrying out operations such as fitness calculation, genetic mutation, allele cross over, and parent selection

RESEARCH EXPERIENCE

Research Assistant in Persistent Homology Feb 2020 - June 2020

University of California, Riverside (UCR), Riverside, CA

- Studied and applied topological data analysis to evaluate the algebraic facets of the persistence of holes in a given space and display high dimensional data trends
- Developed Python code that utilized persistent homology to successfully identify possible academic misconduct in student typing patterns collected by online proctoring sites