

Phoebe Ly

4523 Sandano St. Riverside, CA 92507 | (408)-470-0249 | phoebely7@gmail.com

EDUCATION

University of California, Riverside (UCR)

June 2017 - June 2021

Bachelor of Science, Major in Statistics, Minor in Computer Science

GPA: 3.60

Honors Thesis: Contribution of Plant Canopies to Soil Organic Carbon Along an Elevation Gradient

TECHNICAL SKILLS

Languages: Proficient in R and C++, Experience with Python, SAS, MATLAB, HTML, CSS, SQL

Software: RStudio, Google Earth Engine, Tableau, Minitab, Metabase, Zoom

Relevant coursework: Data Structures & Algorithms, Discrete Math, Nonparametric Techniques, Linear Algebra, Vector Calculus, Optimization, Probability and Statistical Theory, Time Series Analysis, Regression Analysis, Software Construction, Database Management Systems

Average typing speed: 95 wpm

WORK EXPERIENCE

Research Assistant, Center of Food Security & Environment, Stanford University

June 2021 - present

- Researching under Dr. David Lobell in Earth System Science
- Funded by the Stanford Summer Undergraduate Research in Geoscience and Engineering (SURGE) program — a competitive nationwide program (17 students selected out of ~250)
- Used Google Earth Engine to randomly sample and extract Sentinel-2 data
- Used Python to scrape data tables from the Soil Health Card from the Government of India
- Used R to filter out unwanted remote sensing images and calculate various vegetation indices for analysis
- Generated countless exploratory maps and graphs, built linear and random forest models with cross-validation
- Wrote the scientific abstract and presented a poster at the 2021 American Geophysical Union Conference

Student Researcher, Soil Biogeochemistry Lab, UCR

September 2019 - July 2021

- Researched under Dr. Samantha Ying in Environmental Science
- Performed appropriate statistical tests in R with X-ray Fluorescence data to analyze trends and characterize soil composition
- Automated the production graphs that are easy to understand for those who have little statistical expertise
- Managed to work in an interdisciplinary team and learn about an unfamiliar field
- Assisted in a number of other projects that required data cleaning and analysis

Data Analyst Intern, Hyperion Development, Cape Town, South Africa

June 2019 - August 2019

- Developed a system that automatically generates dashboards and charts every week to track employee progress
- Revised and reformatted the company's Software Engineering Bootcamp and Data Science Bootcamp courses to enhance student education
- Explored different data visualization tools such as Metabase and Tableau to find the most efficient and low-cost way to communicate analysis to employees

LEADERSHIP EXPERIENCE

Women in Data Science (WiDS) Ambassador, Stanford University

January 2020 - present

- Established the first Riverside WiDS Regional event that had over 70 people in attendance over Zoom
- Connected with UCR professors to participate in the event planning
- Organized event logistics and ensured that all guests speakers had their questions answered
- Introduced the WiDS Regional event, as well as moderate the Research Talk, Graduate Student Panel, Data Scientist Panel, and R Workshop sessions

President, Highlander Statistics Society, UCR

June 2019 - June 2021

- Integrated, organized, and led coding workshops for students to learn R at club meetings
- Produced a presentation regarding undergraduate research to help other students find opportunities
- Partnered with faculty members to organize club field trips to companies such as Google and Uber

- Collaborated with UCR alumni to speak to the club about career topics and academia
- Coordinated street-clean up events to fundraise for annual Datafest competition
- Re-elected to serve as President for another year

Representative, Diversity, Equity, Inclusion Committee (DEI), UCR

April 2021- June 2021

- Brainstormed ideas and initiatives to build DEI efforts within the community
- Compiled a list of organized campus research programs for students in the Statistics Department
- Organized a Department coffee hour to bridge the gap between students (undergrad/grad) and faculty
- Proposed to officialize mission statement and action plans for the year to demonstrate commitment towards DEI

UNDERGRADUATE PROJECTS

UCR Healthy Campus Initiative 3-Year Longitudinal Survey Analysis

December 2020 - February 2021

- Led a group of 15 undergraduate students to perform data analysis for a on-campus initiative
- Cleaned messy and inconsistent big data across three years for analysis
- Mass produced side-by-side bar charts to compare faculty, staff, and students' attitude and perception on health programs, as well as individual health
- Held coding sessions on independent time to guide peers
- Met all time deadline while working on a tight schedule

Assessing Effects of COVID-19 on Undergraduate Students, UCR Dept. of Statistics

November 2020

- Led a team of 8 students to carry out the analysis
- Surveyed students studying Statistics and Data Science to help the Statistics Department better understand how to help them during the COVID-19 pandemic
- Generated descriptive statistics regarding students' feelings towards online schooling, housing situations, and mental health
- Tested for and found significant results that led to productive conversations
- Promoted diversity by encouraging students from different majors to learn how to perform statistical analysis

American Statistical Association - Datafest, Chapman University & UCLA

April 2018, 2019, 2021

- Participated in a 48 hour data-hackathon to analyze a large and rich dataset and present findings to a panel of judges and a room of a hundred competitors in a visually appealing way
- Chapman (2018) — data donor: Indeed
 - Investigated *Indeed's* construction employment postings and found a spike in interest after a natural disaster; suggested that marketing efforts should be allocated based on forecasted start dates of incidents
- UCLA (2019) — data donor: Canadian Women's Rugby Team
 - Analyzed womens' survey answers regarding health measures and confirmed the hypothesis that the total duration of workouts before a tournament resulted to more games won on average, despite increase in fatigue levels; suggested to not decrease total duration based off fatigue, but not complete high intensity workouts at the start of the season to reduce burnout
- UCLA (2021) — data donor: Rocky Mountain Poison and Drug Safety Center
 - Explored how misuse of drugs is related to one's external surroundings, namely the political culture of where they reside; observed a noticeable trend between Democrats and Republicans and suggested that the organization include a survey question regarding political affiliation

VOLUNTEER WORK

Statistics Tutor, STAT100A & STAT048, UCR

September 2019 - November 2019

- Recommended by a Professor to tutor a student in their class
- Guided student through homework, lectures, and practice tests in Introduction to Statistics courses
- Succeeded in helping a student elevate their exam score by 2 letter grades
- Reserved libraries rooms 4x a week to hold study areas for more students to learn

HONORS & ASSOCIATIONS

Outstanding Student in Statistics Award, ASA OCLB Chapter

January 2022

Honors Excellence In Research Scholarship Award, UCR

January 2021

University Honors, UCR

August 2018 - present

Women in Data Science, Stanford University

January 2020 - present

