Phoebe Ly

Contact: (408)-470-0249 | phoebely7@gmail.com
LinkedIn: www.linkedin.com/in/phoebe-ly
Personal Website: https://phoebely7.github.io/

EDUCATION

University of California, Riverside (UCR)   **June 2017 - June 2021**
Bachelor of Science, Major in Statistics, Minor in Computer Science
Honors Thesis: Contribution of Plant Canopies to Soil Organic Carbon Along an Elevation Gradient

WORK EXPERIENCE

Geospatial Data Analyst, Atlas AI   **April 2022 - present**
- Extracted and analyzed geospatial data, developing a GeoAI platform for various domains (i.e public health, aviation, and agriculture) to enhance client operations while taking on significant responsibilities in a small team environment
- Developed and deployed a range of gradient boosting machine learning models, utilizing MLFlow for comprehensive version tracking from initiation to completion
- Secured an additional $100,000 in project funding by delivering valuable results and proof of concept
- Frequently transformed project concepts into scalable solutions for global customer benefit
- Collaborated with the Engineering team to troubleshoot and resolve GeoAI platform bugs, ensuring smooth transitions to production
- Regularly presented work to clients, internal meetings, and company-wide showcases to foster broader research and business discussions

Research Assistant, Center of Food Security & Environment, Stanford University   **June 2021 - April 2022**
- Sampled and extracted Sentinel-2 data utilizing Google Earth Engine
- Scraped data tables from the Soil Health Card from the Government of India using Python's selenium
- Filtered remote sensing images and calculate various vegetation indices for analysis in R
- Generated countless exploratory maps and graphs, built linear and random forest models with cross-validation
- Authored the scientific abstract and presented a poster at the 2021 American Geophysical Union Conference

Student Researcher, Soil Biogeochemistry Lab, UCR   **September 2019 - July 2021**
- Analyzed X-ray Fluorescence soil data to find trends and characterize soil composition for research
- Automated the generation of user-friendly graphs for audiences with limited statistical expertise
- Collaborated effectively in an interdisciplinary team, gaining insights into an unfamiliar field
- Assisted with multiple projects involving data cleaning and analysis

Data Analyst Intern, Hyperion Development, Cape Town, South Africa   **June 2019 - August 2019**
- Developed a system for automatic weekly generation of dashboards and charts to track employee progress
- Revised and reformatted the Software Engineering and Data Science Bootcamp courses to improve student learning outcomes
- Explored various data visualization tools, including Metabase and Tableau, to identify cost-effective solutions for communicating analysis to employees

TECHNICAL SKILLS

Languages: Proficient in R and Python, Experience with SQL, C++, SAS, MATLAB, HTML/CSS, Javascript, Shell Scripting
Software: RStudio, JupyterLab, QGIS, Tableau, Docker
Platforms: MLFlow, Google Earth Engine, Google Cloud Platform, AWS, GitHub
Technologies: GIS Analysis, Remote Sensing, Spatial Statistics, Data Visualization, Machine Learning, Deep Learning

LEADERSHIP EXPERIENCE

Women in Data Science (WiDS) Ambassador, Stanford University   **January 2020 - May 2022**
- Established the inaugural Riverside WiDS Regional event, attracting over 70 attendees via Zoom

- Organized event logistics and addressed all guest speaker inquiries
- Moderated panels and hosted "Intro to R" workshops for undergraduate students

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

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

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 - June 2021**