Keilyn Yuzuki

keilyn.yuzuki@berkeley.edu | (310) 944-1634 | https://kyuzuki.github.io/

Education

University of California, Berkeley — Expected May 2021

B.A. Data Science, B.A. Cognitive Science — GPA: 3.84

Relevant Coursework: Principles of Data Science; Data, Inference, and Decisions; Probability for Data Science; Data Engineering; Data Mining and Analytics; Data Structures; Linear Algebra; Data Ethics; Program Structures; Data Science and the Mind

Awards: Dean's List, Honors to Date, Girl Scout Gold Award, Presidential Volunteer Service Award

Professional Experience

Zymergen, Inc — Emeryville, CA (Remote)

Data Analytics Intern — May 2020 – Dec. 2020

- Developed and automated ETL pipelines with Spark to connect business applications to Snowflake data warehouse
- Collaborated with stakeholders to develop complex data integrations and customized Tableau dashboards
- Implemented ETL Audit strategy to capture process stats and # of rows processed for 95+ batch jobs run daily

UC Berkeley Division of Computing, Data Science, and Society — Berkeley, CA

Curriculum Developer — Feb. 2018 – Aug. 2020

- Developed short programming assignments tailored towards students without prior programming experience
- Presented developed materials to students and explained basic data science concepts and the domain relevance
- Worked with teams and professors across departments to develop specialized data science modules

Research Apprentice — Jan. 2019 – Dec. 2019

- Worked with partner organizations SPUR and Berkeley Unified as part of the Discovery Research Program
- See SPUR experience below and Exploring the Achievement Gap in Berkeley Public Schools project

The Broad Center — Los Angeles, CA

Summer Associate – Data, Research and Evaluation — June 2019 – Aug. 2019

- Designed R scripts for wrangling 100,000+ rows/year of student enrollment and grad rate data for Salesforce upload
- Implemented new Salesforce object to effectively store and display relevant data for 800+ school systems across time
- Optimized the analysis process for The Broad Residency end of program survey data to measure success and impact

SPUR (San Francisco Bay Area Policy and Urban Research Association) — Berkeley, CA

Research Intern — Jan. 2019 – Aug. 2019

- Developed Python scripts to query the Bing Maps API for info on time and length of driving and transit commutes
- Created dataset of 800,000+ Bay Area commutes for comparing time and cost effectiveness of transportation modes

Design for Equity Lab — Berkeley, CA

Researcher — Jan. 2019 – May 2019

- Researched impact of adaptive equity-oriented pedagogies on student achievement and higher education persistence
- Identified effective teaching practices that benefit students with a low-income or underrepresented background
- Evaluated introductory computer science courses and proposed equity-focused pedagogy to improve student outcomes

Projects

Exploring the Achievement Gap in Berkeley Public Schools

- Analyze CAASPP standardized testing data for students across grade levels in Berkeley Unified School District
- Compare schools and explore trends for students meeting standards in different socioeconomic and ethnic subgroups
- Collaborate with district stakeholders and parents on long term district planning based on data from past 5 years

Spam Email Classifier

- Built a classifier using Python and Pandas to identify spam emails through feature engineering with text data
- Parsed text with RegEx and trained model on features including html tags, length of email, and words in emails

Skills & Interests

Programming Languages: Python (Pandas, Numpy, scikit-learn, matplotlib), SQL, R, Scala, Java

Tools/Frameworks: Apache Spark, Snowflake, Tableau, Jupyter Notebooks, RStudio, Postman, Eclipse, IntelliJ

Interests: Ukulele, Audiobooks, Running, Softball, Geocaching, Shiba Inus