Eric Li

Cell: 410-369-6751 E-mail: xiaodalee@gmail.com

Personal Website LinkedIn GitHub

Education

University of Maryland - College Park (Expected to graduate in May 2024)
Bachelor of Science: Statistics and Computer Science Dual Degree
Unweighted Cumulative GPA: 3.90 (out of 4.0)

Awards/Honors

UMD Presidential Scholarship
UMD University Honors Living Learning Program
Dean's List Fall 2020, Spring 2021, Fall 2021, Spring 2022, Fall 2022, Spring 2023

Technical Experience

Software: Python (pandas, numpy, matplotlib), R (dplyr, ggplot), SAS, MySQL, MATLAB, Java, C, Ruby, Ocaml, Rust
Skills: Data extraction, cleaning, visualization, web-scraping, linear regression, hypothesis testing, machine learning
Relevant Coursework: Data Science, Probability Theory, Statistical Inference, Real Analysis, Multivariable Calculus, Differential Equations, Linear Algebra, Algorithms, Object-Oriented Programming, Biology and Chemistry w/ Labs
Languages: English, Chinese, Spanish

Work/Research Experience

Undergraduate Summer Researcher, College Park, MD — June - August 2023

- Conducted and presented an independent research project for the RISE Lab at UMD that utilized Twitter data to investigate the impact of hurricanes on healthcare facilities. Experimented with NLP methods such as sentiment analysis and wordclouds, gained experience with handling large datasets and creating informative visualizations.

Big Data Summer Institute, Ann Arbor, MI — June - July 2022

- Conducted and presented a genomics research project on differential DNA Methylation between ancestry groups with three other students. Developed quantitative analysis skills through an intensive, interdisciplinary training program in biostatistics, data science, and human health. Heavy use of R programming language and packages.

Projects (found on website and github)

Men's Professional Tennis Data Analysis

- Created a data science tutorial using data from the ATP Tour. Used linear regression and hypothesis testing to investigate whether certain physical attributes such as height and dominant hand are significant advantages.

MoneyBall Analysis

- Investigated MLB team salaries from 1990-2014 to determine if the Oakland Athletics really performed better than other teams while spending less money. Gained experience extracting data using SQL queries. Link

NASA Solar Flare Data Analysis

- Analyzed NASA solar flare data in order to gain experience working with and cleaning datasets. Web-scraped the data directly from the NASA website using BeautifulSoup. Used Pandas to clean and organize the data. Link

Gapminder Life Expectancy Prediction

- Using data from gapminder.org, fitted a generalized linear model factored by continent in order to predict future life expectancies. Learned about regression techniques and residual analysis. Link