Eric Li

Boston, MA | eli@hsph.harvard.edu | 410-369-6751

Personal Website | LinkedIn | Github

Objective

I am currently a biostatistics master's student at Harvard SPH seeking a data science or machine learning internship, with experience in developing data-driven web applications, creating informative visualizations, and conducting comprehensive statistical analyses. My diverse research experiences have ranged from creating an RShiny application for analyzing genomic data to utlizing NLP methods to investigate the impact of hurricanes on healthcare facilities.

Education

Harvard University *September 2024 - May 2026*
M.S. Biostatistics Boston, MA

- GPA: 4.00

University of Maryland College Park *September 2020 - May 2024*
B.S. Computer Science, B.S. Mathematics College Park, MD

- GPA: 3.88, UH Honors Program, Presidential Scholarship

Experience

University of Maryland School of Public Health *June 2024 - July 2024*
Graduate Research Assistant College Park, MD

- Developed a RShiny web application for analyzing genomic data, enhancing accessibility for research teams
- Applied Bayesian network methods to discover causal relationships between genes and noncoding RNAs

University of Maryland Civil Engineering Department *June 2023 - August 2023*
Risk Analysis Researcher College Park, MD

- Developed and presented an independent research project for the RISE Lab that used approximately 8.7 million Tweets to analyze the impact of hurricanes on healthcare facilities, leading to insights for emergency preparedness
- Analyzed data to gain insights into frequently discussed topics during a storm using pandas and sklearn, which informed strategies for improving communication during emergencies

University of Michigan School of Public Health *June 2022 - July 2022*
Biostatistics Researcher Ann Arbor, MI

- Presented a genomics research project on the association between differential DNA Methylation in population ancestry groups and the prevalence of SNPs with three other students
- Applied quantitative analysis skills using R and Python to topics in biostatistics, data science, and public health

University of Maryland Civil Engineering Department *September 2021 - May 2022*
Undergraduate Researcher College Park, MD

- Led a geospatial modeling project to analyze the operational statuses of U.S. nuclear power plants during hurricanes, providing critical insights to strengthen disaster preparedness and response strategies
- Created dynamic maps using Plotly in Python and R to visualize response patterns in nuclear plant operations

Projects

BMI Classification *2023*
Tools: Python, sklearn, pandas, matplotlib [ ]

- Experimented with three classification models (Logistic Regression, Naive Bayes, SVM) to predict the BMI category of an individual based on health attributes such as physical activity level, sleep quality, and blood pressure

Moneyball Analysis *2022*
Tools: Python, SQLite, pandas, matplotlib [ ]

- Integrated data extraction techniques using SQL to investigate MLB team salaries from 1990-2014 in order to determine if the Oakland Athletics really performed better than other teams while spending less money

Men's Professional Tennis Analysis *2022*
Tools: Python, statsmodels, pandas, matplotlib [ ]

- Applied linear regression and hypothesis testing to determine if certain physical attributes such as height and dominant hand significantly increase a player's chances of winning

Technical Skills

- Data Analysis: EDA, Regression Analysis, Machine Learning, Cluster Computing, Web Development
- Programming: Python, R, SAS, SQL/MySQL, Java, C/C++, PHP, MATLAB
- Libraries/Frameworks: Docker, Spark, MongoDB, RShiny, numpy, pytorch, sklearn, matplotlib