Eric Li

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

Ⓗ Personal Website | in LinkedIn | ⓞ Github

OBJECTIVE

I am currently a biostatistics master's student at Harvard SPH seeking a data science, biostatistics, or statistical programming fulltime role, with experience in developing data-driven web applications, creating informative visualizations, and conducting comprehensive statistical analyses. My diverse experiences have ranged from creating an RShiny application for analyzing genomic data to applying CDISC standards for clinical trial data.

EDUCATION

Harvard University   *September 2024 - December 2025*
M.S. Biostatistics   Boston, MA

- Thesis: Markov Models in Population Genetics

- Coursework: Generalized Linear Models, Stochastic Processes, Survival, Longitudinal Analysis, Epidemiology

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

- Coursework: Data Structures and Algorithms, OOP, Numerical Methods, Bioinformatics, Mathematical Statistics

EXPERIENCE

Regeneron Pharmaceuticals, Inc.   *June 2025 - August 2025*
Statistical Programming Graduate Intern   Basking Ridge, NJ

- Created SDTM, ADaM, and TFL datasets for a clinical trial study using SAS

- Developed a RShiny web application for analyzing protocol deviation processes

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 a Bayesian network analysis 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

- Presented an independent research project that used approximately 8.7 million Tweets to analyze the impact of hurricanes on healthcare facilities, leading to insights for emergency preparedness

- Gained insights into frequently discussed topics during a storm using NLP methods, which informed strategies for improving communication practices 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

PUBLICATIONS

- Ke, H., Ye, Z., Feng, L., Xu, Z., Pan, R., **Li, E.**, Qi, J., Chen, S., Liang, M., & Ma, T., (2025, April 16). **Inferring non-coding RNA regulatory network from transcriptomic data and curated databases** [Preprint]. bioRxiv. https://doi.org/10.1101/2025.04.10.648218

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

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: SAS, R, Python, SQL/MySQL, Java, C/C++, PHP, MATLAB
- Libraries/Frameworks: sklearn, pytorch, MongoDB, Spark, Docker, RShiny, numpy, matplotlib