

Eric Li

Columbia, MD | lieric391@gmail.com | 410-369-6751

 Personal Website |  LinkedIn |  Github

OBJECTIVE

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

EDUCATION

Harvard University

M.S. Biostatistics, GPA: 3.94

September 2024 - December 2025

Boston, MA

- Thesis: EB-Coloc: Improving Calibration of Genetic Colocalization Probabilities Using Empirical Bayes
- Coursework: Generalized Linear Models, ML/AI-Ops, Survival, Longitudinal Analysis, Epidemiology

University of Maryland College Park

B.S. Computer Science, B.S. Mathematics, GPA: 3.88

September 2020 - May 2024

College Park, MD

- Awards: University Honors Program, Presidential Scholarship
- Coursework: Data Structures and Algorithms, OOP, Numerical Methods, Bioinformatics, Mathematical Statistics

EXPERIENCE

Regeneron Pharmaceuticals, Inc.

Statistical Programming Graduate Intern

June 2025 - August 2025

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

Graduate Research Assistant

June 2024 - July 2024

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

Risk Analysis Researcher

June 2023 - August 2023

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

Biostatistics Researcher

June 2022 - July 2022

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

AI Note-Taking Application

Tools: Docker, Python Flask, Javascript, Pulumi/Kubernetes, Github CI/CD

2025



- Created an AI note-taking application where users can upload notes/documents and ask questions about it. Answers are enhanced by a RAG model that was trained on introductory college textbooks.

BMI Classification

Tools: Python, sklearn, pandas, matplotlib

2023



- 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

TECHNICAL SKILLS

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