Simeng Li

+1 804-937-8961 | sl3245@cornell.edu | Ithaca, NY

EDUCATION BACKGROUND

Cornell University — Ithaca, NY
Master of Applied Statistics, Data Science Track — *Expected Dec 2025*
University of Richmond — Richmond, VA
Bachelor of Science, Mathematical Economics (Major), Computer Science (Minor) — *Aug 2019 - May 2023*
Major GPA: 3.925 / 4.00
Cumulative GPA: 3.87 / 4.00 (Summa Cum Laude)

- Honors and Awards: Dean's Lists, University of Richmond 2021-2022 Herman P. Scholarship, Departmental Honor Graduate
- Relevant Coursework: Algorithms, Data Structures, Database Systems, Advanced Econometrics, Pandas & SQL Basics, Mathematical Economics, Machine Learning, Mathematical Statistics, Probability, Statistical Learning
- Skills: MySQL, Python (NumPy, scikit-learn, pandas, matplotlib), SAS, R, Java, Tableau, MongoDB
- Certifications: SAS Certified Specialist: Base Programming, Tableau Desktop Specialist

EXPERIENCE

Regeneron Pharmaceuticals — Basking Ridge, NJ
Statistical Programming Intern — *Jun 2024 – present*

Medical Practice Evaluation Center, Massachusetts General Hospital — Boston, MA
Clinical Research Coordinator, Simulation Modeling and Decision Science Pillar — *Jul 2023 – Apr 2024*

- Conduct analyses for research projects using "Cost-Effectiveness of Preventing AIDS Complications (CEPAC)", a computer-based mathematical microsimulation model.
- Perform model runs, assist in writing manuscripts for academic journals and present study results to project teams and external stakeholders.

Clinical Impact and Cost-effectiveness of Long-acting Antiretroviral Therapy (LA-ART) for Breastfeeding People and Their Infants in Zimbabwe (Planned Submission Fall 2024)

- Assess relevant data for Zimbabwe postpartum women studies and perform statistical calculations to derive appropriate model inputs.
- Design simulation model runs and debug the results to identify abnormal trends.
- Perform sensitivity and cost-effectiveness analyses comparing outcomes of high-adherence, low-adherence, and virologic-failure women who are on oral ART regimens vs LA-ART, informing the cost-effectiveness of LA-ART injectable for postpartum women in Zimbabwe.

Incorporating HIV Transmissions into Adolescent-focused Cost-effectiveness Analyses: A Scoping Review (Planned Submission Summer 2024)

- Conduct blind screening for 2400 studies that use mathematical models to simulate averted adolescent HIV transmissions and resolve screening conflicts with research scientists.
- Coordinate weekly meetings, document project progress, and distribute concise meeting notes to ensure team alignment.

University of Richmond — Richmond, VA
Student Researcher

Length Bias Estimation of Small Businesses Lifetime
Mathematical Economics Honor Thesis — *Sept 2022 – Apr 2023*

- Conducted in-depth analysis of lifetime data for 24 restaurants located on Carytown Commercial Street in Richmond, employing advanced statistical techniques to estimate the average lifespan in the area while minimizing the impacts of length bias and right-censoring.
- Implemented Kaplan-Meier estimators to account for right-censored observations in R. Performed rigorous goodness-of-fit tests, such as the Anderson-Darling (AD) test, to assess the efficacy of the developed estimator against the exponential distribution. The results unveiled precise lifetime estimations devoid of length bias.
- Showcased the findings as an honors thesis research project within the mathematics department, elucidating the concept of length bias using relatable examples from cancer detection to non-technical audiences.

Analyzing the Statistical Effectiveness of Diagnostic Tests for COVID-19 Using the ROC Curve — *May 2022 – Aug 2022*
Student Researcher and Group Leader

- Worked with university librarians and located recent diagnostic data sets on COVID-19 hospitalization statistics and death rates due to COVID-19, including 566,600 observations.
- Developed a preliminary ROC curve with Python using patients' age as the threshold that achieved an AUC-ROC of 0.81 and plotted the error trade-offs of different decision rules using R.
- Constructed univariate and multivariate logistic regression models and utilized the predicted values to construct an updated ROC curve with an improved AUC of 0.86, indicating a decent performance of the logistic regression classifier. The ROC curve informs possible decision rules to allocate limited medical resources during the COVID-19 pandemic.

ACADEMIC PROJECTS

Customer Subscription Prediction Using Machine Learning Models for a Freemium Music Business *Mar 2022 – Apr 2022*
- Imported data from the university MongoDB server and created a pandas data frame for analyzing and building models.
- Designed and developed machine learning models for predicting whether a freemium music subscriber will become a paid member in Python, including KNN, logistic regression, decision tree, and random forest algorithm.
- Presented 4 most important features for increasing subscribers based on the decision tree model with an accuracy score of 85.266% after applying SMOTE over-sampler.

Machine Learning on Restaurant Reviews in Las Vegas *Apr 2022*
- Utilized k-means clustering on 7,000 customer reviews to discern underlying restaurant types, revealing the potential restaurant classification strategies in Las Vegas.
- Constructed log-likelihood ratios for context-driven keywords and made compelling distinctions within each cluster of restaurants. The patterns uncovered include unique food types and restaurant-related activities.

CAMPUS INVOLVEMENT

Department of Mathematics and Computer Science, University of Richmond Richmond, VA
Tutor and Grader *Jan 2022 – May 2023*
- Tutored one-on-one for Introduction to Programming and Machine Learning for Business Analysts and gave students comments.
- Graded homework and labs for 25 students on Data Structures and gave comments.

New Spider Orientation, University of Richmond Richmond, VA
International Orientation Advisor *Jul 2021- Jan 2022*
- Organized orientation materials and performed administrative work and tasks throughout the orientation.
- Led three groups of new students from more than ten countries to various campus locations for tours and events.