

SHAMIN SAHEBZADA

ShaminSahebzada@gmail.com | www.linkedin.com/in/shaminsahebzada/ | (516) 554-2704

Epidemiologist with strong skills in epidemiological methods (causal inference, bias, interaction, and study/survey design) and statistical methods (model building, regression, survival analysis, data visualization/storytelling and machine learning). Proficient in statistical programming using SAS, R, and Python (NumPy, Pandas, Scikit, Matplotlib, Seaborn, Plotly). Expertise in healthcare data, systematic literature reviews, public datasets, and written and oral communication.

PROFESSIONAL EXPERIENCE

- Columbia University, Metropolitan New York Registry** (New York, NY) - **Research Coordinator** 2017 - 2020
- Breast and Ovarian Cancer registry of over 8,000 high-risk families actively followed for decades with the goal of improving cancer prevention, detection, and treatment
 - Utilized **SAS** to **clean, operationalize, and verify the integrity of data** from one of the largest family-based cohorts in the US, as well as other **real world data sources (SEER, NHANES)**
 - Built multivariable **regression models** aimed at understanding the **interactions** between a given risk factor and risk-predicting scores on breast cancer risk
 - Co-authored a **systematic literature review** using 100 articles and assisted principal investigator in writing **academic articles, abstracts, and PowerPoint slides** for internal, national, and international conferences
 - Updated and improved follow-up questionnaires/surveys and researched new questionnaire constructs related to breast cancer risk, which were coded into **Qualtrics web surveys** that improve cohort retention
 - Assisted participant **recruiting** efforts and ensured cohort follow-up and return of surveys by conducting phone interviews, emailing, creating newsletters, and **organizing community outreach events**
 - Coordinated with breast cancer diagnosed participants, hospitals, and pathologists to obtain **medical records** and pathology samples for secondary review and input into BCFR database
- Stony Brook University, Sustainability Studies Ecotoxicology Lab** (Stony Brook, NY) - **Researcher** 2013 - 2016
- Ecotoxicology lab focused on investigating the effects of glyphosate herbicide, heavy metals, and fertilizer on earthworm and soil microbial health
 - Designed experiments, **created budgets**, and coordinated a team of 40 undergraduates to guarantee safe and proper execution of eight projects
 - Used **R statistical programming** to run **t-test, ANOVA, and linear regression** analyses on lab datasets
 - Effectively communicated 14 posters at multiple conferences over three years and was awarded Stony Brook's Undergraduate Researcher of the month

TEACHING EXPERIENCE

- Columbia University, Applied Epidemiologic Analysis** (New York, NY) - **Graduate Teaching Assistant** 2018
- Graduate level course on epidemiologic **study design, bias, confounding, interaction, mediation** and the following statistical methods: **linear, logistic, cox proportional hazards, and poisson regression models**
 - Instructed a lab of 30 students on how to implement advanced statistical methods in SAS using real datasets to develop models and critically evaluate results
 - Coordinated with professor to improve lectures, labs, homework assignments, and exams
- Columbia University, Molecular Epidemiology** (New York, NY) - **Graduate Teaching Assistant** 2018
- Graduate level course on the applications of biomarkers within epidemiology. Covered the selection and **validation of biomarkers**, biological sample collection, storage and banking, current laboratory methods for **biomarker analysis**, and **gene-environment interactions**
 - Held weekly office hours, instructed students in one-on-one sessions, graded homework assignments, and met with professor weekly to make decisions on how to improve the course content and layout
- Columbia University, Categorical Data Analysis** (New York, NY) - **Private Tutor** 2017
- Graduate level course on statistical methods for binary and other discrete response data focusing on applications within epidemiologic and clinical studies that covers **probability, hypothesis testing, logistic regression, model building, confounders, mediators, colliders, and interaction analyses**
 - Instructed two MD Post-Doctoral Clinical Research Fellows on how to systematically analyze categorical data based on data structure and hypotheses using **SAS**

EDUCATION

- Columbia University, GSAS** (New York, NY) - **PhD, Epidemiology** 2018 - 2020 (Incomplete)
- Relevant Courses: Randomized Clinical Trials, Causal Inference, Applied Regression, Critical Thinking in Epidemiology
- Columbia University, MSPH** (New York, NY) - **MS, Epidemiology** 2016 - 2018
- Relevant Courses: Public Health Epigenetics, Genetics in Epidemiology, Applied Epidemiologic Analysis, Cancer Epidemiology, Categorical Data Analysis
- Stony Brook University** (Stony Brook, NY) - **BS, Biology, Departmental Honors, cum laude** 2013 - 2016