MIAOLEI BAO

Phone: 206-710-8016 ◇ Email: mlbao@uw.edu ◇ GitHub ◇ LinkedIn

SUMMARY OF QUALIFICATIONS

Masters-level Biostatistician passionate about healthcare, offering analytical skills in programming (R, Python, SAS). Strong problem-solving and collaboration abilities; excellent written, verbal, and visual communication skills. Meticulous and thorough, with demonstrated success in handling complex datasets and statistical modeling.

SKILLS

Programming	R, Python, SAS, Github, Latex, R Markdown, Jupyter Notebook
Statistical	Statistical Modeling, Visualization, Data Management, Survey Data, Clinical Trials, Survival Analysis, Nonparametric Regression, Longitudinal Analysis, Spatial Statistics Machine Learning, Regression Analysis, Categorical Data, Statistical Analysis Plan Generalized Linear Models

WORK EXPERIENCE

Fred Hutch Cancer Center, HIV Prevention Trials Network 02/2024-present
Research Assistant for Professor Deborah Donnell, Drug Adherence Analysis *Seattle, Washington*

- Conducted extensive literature reviews, collaborated with biostatistics, medicine, and pharmacy professors to propose and refine metrics for defining adherence categories, and **presented** findings to the principal investigators.
- Extracted and cleaned data from **databases** and **case reports** for about **2,000** participants, identified adherence patterns, analyzed **longitudinal** data, and **visualized** adherence trajectories over time using **R**.

China Guangfa Bank Co. Ltd, Corporate Banking Division 07/2022-09/2022
Financial Analyst Intern *Hangzhou, China*

- Monitored corporate lending activities, conducted both overall and sector-specific analyses, and **presented** insights to a senior management team of three, supporting informed financial decision-making.
- Co-developed **Excel-based** algorithms to rank sub-branches performance, streamlining monthly business reviews.

RESEARCH EXPERIENCE

Valley View Health Center, Capstone Project 09/2024-present
Title: Retrospective Cohort Study of Pharmacist-led Diabetes Interventions *Seattle, Washington*

- Conducted regular meetings with the medical professional sponsor to offer statistical guidance on data collection and analysis methods, and developed a comprehensive **statistical analysis plan**.
- Utilized **longitudinal mixed-effects models** and **survival analysis** to demonstrate the treatment's superiority.

Health Metrics Prediction Using Demographic and Health Surveys GitHub 05/2024-present
Research Assistant, Supervised by Professor Jon Wakefield *Seattle, Washington*

- Automated code extraction and preprocessing utilizing Natural Language Processing (**NLP**) in **Python** for **4000+** metrics, and applied continuous spatial models in **R** with SPDE to generate national and regional estimates.
- Prepared written reports and created various visualizations for group meetings with academic collaborators, government representatives, and United Nations experts, facilitating data-driven decision-making on public health priorities.

Thorax Disease Prediction via Machine Learning-based X-ray Classification GitHub 09/2024-12/2024
Collaborative Biomedical Project *Seattle, Washington*

- Developed **prediction models** using **Convolutional Neural Networks** to classify thorax diseases using chest X-ray data from over 5,600 images, employed **F1 score** to ensure robustness, and achieved an accuracy of **93.3%**.
- Utilized **TensorFlow** and **Keras** for model architecture development, applied dropout techniques to prevent overfitting and improve model generalization, and **optimized** hyperparameters such as batch sizes for high-dimensional image data.

Kernel Density Estimation with Spherical Data GitHub 05/2024-06/2024
Independent Project *Seattle, Washington*

- Developed a kernel density estimator for spherical data in R based on a methodology paper, evaluated performance with simulated and real data, optimized parameters through cross-validation, and created interactive visualizations.

Randomized Clinical Trial Design 04/2024-06/2024
Statistician, Randomized Prospective Trial on Reducing Contact Lens-Related Adverse Events *Seattle, Washington*

· Collaborated with a multi-disciplinary team to formulate the experiment design and led the drafting of the **clinical trial protocol,** employing a factorial design to evaluate the effects of two behavioral interventions on targeted outcomes.
· Specified expected treatment effects, calculated the required trial **sample size**, defined the **randomization** and blinding strategies, outlined methods for handling missing data, and detailed the approaches for primary and secondary analyses.

Genomic Analysis of Tobacco for Identifying Genes Related to Chemical Constituents 10/2022-06/2023
Undergraduate Dissertation *Hangzhou, China*

· Calculated descriptive statistics, performed correlation analysis and conducted variance analysis on the concentrations of commercially important tobacco chemicals.
· Identified significant gene-gene and gene-environment interaction effects using the gene linkage map and predicted 21 candidate genes useful for tobacco breeding using protein databases.

Ecosystem Service Evaluation of Wetlands in the Yangtze River Delta 05/2021-05/2022
Research Project, instructed by Dr. Yuting Xie *Seattle, Washington*

· Obtained data from the past 35 years via ArcGIS using map images, and computed landscape indices.
· Calculated changes in ecosystem service value within a 35-year time span, visualized the results in R, and provided development planning suggestions for the area administration.

EDUCATION

University of Washington, Seattle, WA 09/2023 – 03/2025 (expected)
Master of Science in Biostatistics, Pathway: Data Science

Zhejiang University, Hangzhou, China 09/2019 – 07/2023
Bachelor of Science in Ecology