Jialiang Hua

New York, NY, 10032 | (646) 346-0432 | jh4323@columbia.edu | github.com/cuhuajialiang

EDUCATION

Columbia University, Mailman School of Public Health — **New York, U.S.**
Master of Science in Biostatistics Theory and Methods Track — Expected May 2023

Relevant coursework: Epidemiology, Biostatistical Methods I & II, Statistical Inference, Probability, Data Science, Data Science II, Analysis of Longitudinal Data, Survival Analysis, Clinical Trial Methodology, Relational Databases and SQL Programming, etc.

Soochow University — **Suzhou, China**
Bachelor of Medicine (English based, MD-equivalent in the U.S.) — June 2020

SKILLS

Language: English (fluent), Mandarin Chinese (native)
Computer: Proficient in R, SQL, SAS, Microsoft Word, Excel, PowerPoint. Experience with Oracle Health Sciences and Python.
Certifications: Python for Everybody Specialization by University of Michigan on Coursera.

WORK EXPERIENCE

Department of Biostatistics, Columbia University — **New York, U.S.**
Research Assistant — April 2022 – April 2023

- Conducted in-depth preparation and exploration of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) dataset using R, ensuring data quality through meticulous validation, data cleaning, handling missing values, and performing comprehensive exploratory data analysis.
- Used cross-validation to train random forest models to identify important risk factors for future mania in people with major depressive disorder based on NESARC dataset. Applied the weighted random forest algorithm and parallel computing to handle imbalanced data and accelerate the modeling process.
- Examined various techniques for managing imbalanced data, including balanced random forest, weighted random forest, over-sampling, down-sampling, and Synthetic Minority Over-sampling Technique (SMOTE).
- Programmed extensively in R to build logistic regression models, assess and compare model performance, and explore how random forest deals with raw versus preprocessed variables.
- Generated research abstracts and reports in written and graphical form for meetings, seminars and practicum symposium.

Pfizer Inc. — **Shanghai, China**
Clinical Data Manager (Wuxi Clinical Outsourced) — October 2020 – April 2021

- Performed data reviews, data reconciliation and query management with SQL and Excel to support 3 ongoing clinical trials.
- Provided over 30 data management metrics reports to study team and stakeholders.
- Conducted user acceptance tests of eCRFs and ensured filing of all related documentation.

The First Affiliated Hospital of Soochow University — **Suzhou, China**
Intern, Clinical Rotation — June 2019 – June 2020

- Rotated at the Departments of Gastroenterology, Infectious Diseases, Primary Care, Psychiatry, Surgery and more.
- Conducted research on the effect of Kuhuang (TCM) on hepatitis E disease progression. Performed Mann-Whitney U test with SPSS. Proved that Kuhuang has no significant clinical effect on the treatment of hepatitis E.

RESEARCH PROJECTS

Database for Clinical Trial Evaluating Pain Reprocessing Therapy for Chronic Back Pain patients — **New York, U.S.**
Project Lead, Github Link — December 2022

- Designed and built a database in MySQL for a superiority trial analyzing efficacy of pain reprocessing therapy versus standard care for patients suffering from chronic back pain. Designed and added triggers to tables for data validation.
- Tested the database with data entry and queries. Created 10 + complex SQL queries to extract information from the database.

Modeling of Innocent Death Rate Made By Police in the U.S. — **New York, U.S.**
Project Team Member, Github Link — September 2021 – December 2021

- Collected, cleaned, and processed raw data of death records on 50 US states from 2010 to 2020.
- Constructed multiple linear regression model to investigate the association between innocent death rate by police and potential factors. Conducted model diagnostics to validate model assumptions.
- Authored the final report and built the whole website aggregating all the figures, charts and results.