Jiaqi Li

100 Haven Avenue, Tower 3 17F, New York, NY 10032
(201) 887-9945, jl5025@columbia.edu
https://ljq312.github.io/

EDUCATION

Columbia University, Mailman School of Public Health, New York, NY   *May 2019*
Master of Science, Biostatistics, GPA: 3.6
- Relevant courses: Biostatistical Methods, Data Science, Statistical Learning & Data Mining

University of Washington, Seattle, WA   *December 2013*
Bachelor of Arts, Psychology; Bachelor of Arts, Economics, GPA: 3.5
- Honors: Dean's List, Cum Laude

RELEVANT PROJECTS

Data Science: Flight delay patterns (website is available here)
- Explored factors associated with flight delay patterns from 2012 through 2016 using R.
- Manipulated 2.8GB of data and analyzed 29M+ flights using a virtual server and Linux operating system.
- Created a Github website to display project findings, including a report, video, and interactive data visuals made by Plotly and ggplot.

Statistical Learning: Using statistical learning to analyze Airbnb listings
- Built a predictive model of rental price; model was selected from various statistical learning algorithms of linear and nonlinear regression based on performance (using methods include ridge, lasso, PLS, splines, GAM).
- Classified high/low review score by selecting model from logistic regression, LDA, QDA, tree methods, and SVM.

Biostatistical Methods (report is available here)
- Identified variables associated with hospital length of stay (LoS) patterns to construct a predictive model using Multiple Linear Regression with R and SAS.
- Ran rigorous model diagnostics and model validation (cross-validation, bootstrap).

Using SAS to compare effects of LAGB and RYGB, and examine associations in weight loss (post is available here)
- Built a model to identify measured variables associated with weight loss using multiple linear regression.
- Compared obesity treatments effect overtime using longitudinal data analysis.
- Created descriptive statistics for relevant variables.

EXPERIENCE

Synyi Artificial Intelligence   *Shanghai, China*
Internship – RWE Biostatistician   *May 2018-August 2018*
- Designed regional EHR analysis data set query specifications.
- Built models using Propensity Score Matching, and built attractive data visuals to display analysis results and demographic information.
- Communicated with pharmaceutical company sponsors and participated in protocol review and edit.
- Consulted with clinical investigators on sponsors' needs with respect to statistical design of studies.
- Developed SQL builder prototype using Sketch.

Columbia University Medical Campus   *New York, NY*
Research Assistant   *Oct 2018 - present*
Structural Determinants & Social Transitions among Adolescents and Young Adults in Rakai
- Clustered reported characteristics of subjects' sexual partners into different classes using Latent Class with Random Effect Analysis in R.
- Compared the resulting partner classes from Latent Class with Random Effect Analysis with Latent Class Analysis qualitatively and quantitatively.
- Created codebook to clean historical data and standardize coding rules.

SKILLS

- Statistical languages: R (& Shiny), Python, SAS, SQL
- MS Office: Word, Excel, PowerPoint, Access
- Bilingual, Chinese and English