Tianran ZHANG

Homepage | GitHub | LinkedIn | +1 (646) 269-6106 | im.trzhang@gmail.com | LA, CA

EDUCATION

Cornell University, New York, NY **Sep 2019 – Aug 2020**

M.S. in Biostatistics & Data Science

GPA: 4.0/4.0 | Certificate of Academic Excellence

Fudan University, Shanghai, China **Aug 2015 – Jul 2019**

B.S. in Mathematics (Data Science)

Undergraduate Scholarship, STEM Scholarship, China Mathematical Modeling Contest

SKILLS

Technical Skills	Python (NumPy, Pandas) • SQL • R• Tableau • MATLAB • Jupyter • Stata • Git • C • SAS • Bash • LaTeX • AWS
Statistical Modeling	Linear/Logistic/Ridge/Lasso Regression • Tree Based Models • Classification • k means • k-NN • Clustering • ANOVA Bayesian Network • Non-linearity • PCA • SVM •GAM • Causal Inference • Survival Analysis • Longitudinal Analysis

RELEVANT EXPERIENCE

GE AVIATION | Virtue Digital Technology Data Analytics Intern **Aug 2020 – Sep 2020 | New York, NY**

- Used GE Dataiku platform to combine flights and airport data to determine the distance traveled for each airplane.
- Built and published insights on a dashboard on GE Dataiku to visualize the KPI tables and other useful information.

WEILL CORNELL MEDICINE | Research Assistant **May 2020 – Aug 2020 | New York, NY**

- Developed a more general analysis template for analyzing Human Microbiome Project (HMP) 16S microbiome data.
- Reprocessed the raw sequence data of human microbiome composition with a higher speed and newer algorithm.

ZHONG OU AMC (Quantitative Investment Department) | Data Scientist Intern **Jul 2018 – Oct 2018 | Shanghai, China**

- Conducted data wrangling on over 5 million funds data from 1480 companies with Python; Applied OLS and FGLS regressions in STATA.
- Compared the affection of different factors on market share of companies under different strategies nested within China and U.S.

FUDAN UNIVERSITY | Data Mining Teaching Assistant **Sep 2018 – Jan 2019 | Shanghai, China**

- Summarized the data analysis methods in the financial area and prepared the lecture notes for students.
- Gave instructions and helped students understand deeply in building financial and time series models.

FUDAN UNIVERSITY | Data Visualization Research Assistant **Jan 2018 - April 2018 | Shanghai, China**

- Constructed and performed free-form deformations for any real-world object based on its shape and position.
- Innovation of matrix calculation was carried out to speed up the deformation procedure by 5 times with Python.

INSIDESHERPA | Virtue Data Analyst Intern **April 2020 - May 2020 | New York, NY**

- Calculated the transaction volume and spending over the course for each day.
- Conducted predictive analytics to identify the annual salary for each customer.

PROJECTS

ASSOCIATION OF NUMERACY (OBJECTIVELY AND SUBJECTIVELY ASSESSED) WITH SELF-RATED HEALTH | Poster, GitHub

Submitted to Journal – Patient Education and Counseling (Short Communication)

- Conducted logistic regressions with post-stratification weights and design weights for generalization.
- The final results outperform than PIAAC analysis with a 92% power and 15.2% prediction error rate.

ENVIRONMENTAL IMPACT CALCULATOR | Website, GitHub

- Built UI and Server to deploy an interactive calculator shiny app for Happy Valley Meat Company with R.
- Used bar plots and a word cloud plot to visualize the environmental impact differences among cuts.

"SCR" R PACKAGE | Website, GitHub

- This package serves as an associate tool to help students get a deep understanding about the book "Statistical Computing with R".
- Designed and implemented nearly 100 R functions in the "SCR" R package; allocated work and did the follow-up maintenance.

BREAST CANCER PREDICTION | Report, GitHub

- Conducted k-means to cluster the patients according to their top 2 out of 30 nuclei's features correlated with breast cancer by PCA.
- Applied two-sample t-test to compare the features' difference between patients with recurred breast cancer and those who did not.