Junyan Yao
Email: yaojunyan1202@gmail.com GitHub: https://github.com/junyanyao

EDUCATION

New York University, New York, NY January 2018
Master of Science in Applied Statistics, GPA 3.72/4.0

University of North Carolina, Chapel Hill, NC May 2015
Bachelor of Arts in Economics (Cum Laude), Minor in Mathematics Decision Science, GPA 3.72/4.0

TECHNICAL SKILLS

Natural Language Processing, PostgreSQL/SQL, Stata, SAS, R, Machine Learning, A/B Testing, Python, Scikit Learn

WORK EXPERIENCE

Quantitative Operations Analyst, Elevation Securities 11/2015 - Present

- Lead internal projects pertaining to data gathering, cleaning, large scale report runs from multiple clearing firms' platform; and build pipeline for reconciliation process for trade records with settlement data in Python
- Develop and maintain multiple dashboards integrating PostgreSQL to analyze client trade history and to generate insightful reports using statistical techniques and hypothesis testing to validate findings
- Improve company operational efficiency by supporting back office applications to ensure proper trade settlements and successfully automate trade booking process and end of day trade order completion reporting for equity markets.
- Demonstrated the real-time trade order flow analytics to align with compliance risk control problems and execution

Research Assistant, New York University 12/2017-06/2018

- Took a natural language processing approach to help improve students' collaborative skills by analyzing real data in which students collaborate via online chat and focusing on features of chat content and click-stream data
- Analyzed students' dialogue includes content from assessment materials and affect scores with existing lexicons packages using Python; the sequential dependence between student chat and student response behavior; the extent of temporal synchronization in students' response times
- Explored and interpreted features and resulted in a multidimensional description of collaborative process data
- Applied unsupervised machine learning techniques to infer whether different classes of groups and individuals can be identified using features from processes, team composition, and outcomes

Consultant, Expository Writing Program, New York University 06/2017 - Present

- Perform data cleaning for evaluation data including quantitative and qualitative from Qualtrics using R and Stata.
- Design instructors' evaluation survey in Qualtrics and maintaining department courses databases for each semester.
- Lead and collaborates in developing various statistical reports including tests of significance to evaluate instructors' teaching feedback across the department and provide insightful recommendation to department chairs.

Quantitative Student Consultant, New York University Data Services Lab 02/2017 - 12/2017

- Provided quantitative analysis consultation including data manipulation and statistical data modeling for faculty and students' academic researches, such as A study on Diffusion of Public Health Laws of Tobacco 21.
- Led Data Services instructional tutorials and prepared training documentation for R, Stata, and SAS courses.
- Facilitated research empirical analysis to senior honor students to develop thesis projects including create and test hypotheses based on their research questions in multiple fields and subjects, especially public policy.

ACADEMIC PROJECTS

Cluster and Classification Analysis January 2017

- Researched and explored Single linkage, Centroid, Complete linkage, Ward, and Kmeans as potential clustering methods using the features constructed and examined from various dataset in Python programing lauguage
- Identified the optimized cluster solutions by maximizing the adequacy using Calinki & Harabasz criterion using Ward and Kmean methods along with other qualitative criteria in finding the actionable value
- Validated the result with the known demographics distribution information and principal component plot.