Qinyue(Hailey) Hao

(+1) 646-678-8375 | qh2231@columbia.edu | https://github.com/HaileyHao
155 Claremont Avenue, New York City

Education

Columbia University, Graduate School of Arts and Sciences   **New York**
MA in Quantitative Methods in the Social Sciences, Data Science Focus   Jan 2021 – May 2022
- **GPA:** 4.02/4.0
- **Major courses:** Machine Learning/ Time Series Analysis/ Social Network Analysis/ Data Mining/ Modern Data Structures/ GIS & Spatial Data Analysis/ Practicum in Large-Scale Data Analysis

Communication University of China, School of Journalism   **Beijing**
BA in Communication, Major in Media Market Research and Analysis   Sep 2016 – Jun 2020
- **GPA:** 3.74/4.0   **Ranking:** 6/30
- **Major courses:** Data Visualization (94)/ Sampling Theory and Methods (94)/ Multivariate Statistical Analysis Methods (91)/ Statistics (91)/ SPSS/ Database and Access
- **Awards:** Merit Student of CUC (2018, 2019), Second-class Scholarship of CUC (2019), The Second Prize in the Beijing University Social Survey Paper Competition (1% of 1000 contestants)

Projects

Measuring the Integration and Network Effect of the UN SDGs   Sep 2021 – Dec 2021
- Developed whole network models of the UN Sustainable Development Goals based on the correlation between 100+ variables. Visualized networks, conducted social network analysis, and delivered presentations to UN officials about the trade-offs and secondary benefits of SDG-orientated policy enactment for optimizing investment decisions towards the utmost joint effect.

Machine Learning Competitions on AI Model Share   Sep 2021 – Now
- Utilized feature engineering and tuned machine learning models and neural network models to solve prediction problems including clickbait identification, spam detection, world happiness level prediction, and flower picture classification. Models including penalized linear & logistic models, Random Forest and other Ensemble tree models, SVM, Voting models, MLP, CNN, etc.

FBI Wanted API Python Wrapper   Dec 2021
- Created a python package for retrieving datasets of FBI wanted people by appearance features or administrative regions, as well as wanted posters with photos.

Visualization of Covid-19 Vaccine Side Effects   Apr 2021
- Visualization in R of the side effect reports from a WHO database to understand the side effects of different Covid-19 vaccines. Conducted text analysis and sentiment analysis on relevant Tweets.

Internships

ByteDance   **Beijing**
Data Analyst Intern, Data Strategy Group, Automotive Business Center   Aug 2019 – Nov 2019
- Retrieved data from large internal databases using SQL, collated and cleaned the data with Excel PivotTable and functions, conducted data visualizations and text analysis with Python, and established user models to propose marketing strategies.
- Provided data insights and data strategy reports of client car brands including competitive analysis, user profiles, etc., for the marketing team to develop advertising plans and boost advertising sales.

iResearch Consulting Group   **Beijing**
Assistant Data Analyst Intern, E-Business Group, Industry Research Department   Mar 2019 – Jun 2019
- Assisted analysts to complete industry insight reports: collected public data of the competing companies of the client in the industry, interviewed relative personnel, and compiled data into charts.

Skills

Programming: Python/ R/ SQL/ pandas/ dplyr/ ggplot2/ tidymodels/ scikit-learn/ NLTK/ ipython-sql
Industry Skills: Machine Learning/ Statistical Modeling/ Time Series Analysis/ Data Visualization/ Social Network Analysis/ Spatial Analysis/ APIs/ Git and GitHub/ AWS/ Databricks/ Tableau/ Spark