

Fuman (Annie) Xie

412-579-3228
fumanx@andrew.cmu.edu
www.linkedin.com/in/fuman-xie
fuman-annie-xie.github.io

Education

Carnegie Mellon University, Heinz College 5/2024
M.S. in Public Policy & Management – Data Analytics; GPA: 4.10/4.33
A highly selective program that includes robust coursework in machine learning, statistics and computer programming
Shanghai Jiao Tong University, School of International and Public Affairs 3/2021
Master in Public Administration; GPA: 3.89/4.0; Honor: National Scholarship
University of Michigan – Shanghai Jiao Tong University Joint Institute 8/2018
B.S. in Electrical and Computer Engineering; GPA: 3.51/4.0; Honor: Dean's List

Skills & Coursework

Programming: Python, SQL(PostgreSQL, MySQL), R(ggplot2, tidyr, tidycensus), C++, Stata, git
Analytics: sklearn, Pandas, NumPy, GeoPandas, R, EXCEL, SPSS, Seaborn, matplotlib, econometrics
Courses: Data Focused Python; Database Management; Machine Learning Foundations with Python; Data Structure and Algorithm Applied Econometrics; Decision Analytics for Business and Policy; Statistics; Probabilistic Methods;

Intern Experience

[Python | R] Intern at Social and Decision Analytics Division, Biocomplexity Institute, UVA 05/2023 – 08/2023

- Improved data quality of large household-level datasets by geocoding, web scraping and data wrangling in Python
- Used cross validation to tune the KNN and kernel smoothing model to impute internet speed in R
- Co-designed and ran the difference-in-difference model used to evaluate a rural development program; matched treatment and control group in based on housing features; created a workflow that can be reproduced to assess other projects
- Led an analysis using Panel Study of Income Dynamics; ran a panel regression to study the effects of economic recessions on household-level charitable givings over a 20-year period

Project Experiences

[Data Dashboarding] Capstone Project with Bipartisan Policy Center 8/2023 - present

- Automating the process of downloading and integrating immigration data from Customs and Border Patrol
- Creating a data dashboard that can provide insights for the general public and be updated automatically

[Classification | Regression] Analyzing Trip Data of Citi Bike in New York City 03/2023 – 05/2023

- Consolidated and cleaned over 35 GB of weather data, demographic data and bike trip data
- Classified user type and bike type with different models including logistic regression, naïve bayes classifiers and kNN; predicted daily usage with regularized linear, regularized and decision tree regression

[Python | PostgreSQL] Web Scraping and Database Management 8/2022 – 12/2022

- Collected information of over 20000 wedding venues across the U.S. with BeautifulSoup, NumPy and Pandas in Python
- Designed a PostgreSQL database to manage all the datasets in this group project

[Python | MySQL | Web Development] Production Document Management System 12/2016 – 3/2017

- Cooperatively launched an online production document management system for suppliers of Siemens in China using the framework Flask in Python, MySQL, JavaScript, HTML and Git
- Implemented table-related functions such as creating/deleting tables, filling in the tables, adding/deleting rows, uploading pictures, autosaving, exporting to Excel files and controlling different versions

[SPSS | Regression] College Students' Volunteering and Donating Behaviors in China 9/2019 – 3/2021

- Analyzed data from 8000+ questionnaires (over 1 million entries) in SPSS and used binomial logistic regression and linear regression to investigate the predictors of volunteering frequency and donating behavior
- Published a paper as first author in an A&HCI journal and presented the project at the 2020 ARNOVA conference

Work Experiences

Manager of a Family Business in Manufacturing in Guangdong, China 5/2021 – 7/2022
Volunteer Teacher in Guangxi, China 8/2018 – 8/2019