DAISY LU

Email: daisylu@uchicago.edu
Website: daisylu.github.io
Citizenship: USA

EDUCATION	*Ph.D. in Public Policy*	Expected 2027
	University of Chicago, Harris School of Public Policy	Chicago, IL
	B.A. in Sociology	2016
	University of Chicago	Chicago, IL

SKILLS

Programming: Python, LaTeX, Git, Linux, SQL, Spark, R, Stata
Languages: Cantonese (fluent)

RESEARCH EXPERIENCE

Research Assistant — 2020 - 2022
Professor Steve Cicala, NBER & Tufts University — Cambridge, MA

- Prepared novel data sources for research to estimate the effect of pollution on health using NASA MODIS satellite data to measure particulate matter in the atmosphere and NOAA HYSPLIT to simulate particle dispersion.
- Processed millions of observations of Medicare data using Stata to create county-level mortality rates over different time windows.
- Wrote over 40 Python scripts to scrape decades of hourly energy data for different countries (accounting for about 60% of global electricity consumption) from Tableau dashboards, HTML tables, file directories, and API endpoints.

INDUSTRY EXPERIENCE

Data Science Team Lead — 2018 - 2020
Home Partners of America — Chicago, IL

- Developed a Python algorithm using Multiple Listing Service (MLS) data to automatically find similar homes to a given subject home by minimizing the Euclidean distance between vectors of home feature differences.
- Modeled relationship between resident characteristics, such as FICO and income, and lease outcomes, such as renewal and eviction, by exploring different sampling and synthetic minority oversampling techniques (SMOTE).
- Created an internal Python package with team members for commonly used data tasks such as establishing database connections, API clients, Census data wrangling, and geospatial analysis.

Senior Data Scientist — 2016 - 2018
Nielsen — Chicago, IL

- Completed a two-year rotational program emphasizing technical depth and leadership growth, working on both consumer-packaged goods and digital media measurement methodology.
- Enhanced training data for and evaluated performance of Long Short Term Memory models used to classify millions of item descriptions into under twenty product categories.
- Developed a sales forecasting visualization tool using R Shiny, comparing the accuracy of ETS and ARIMA models.

AWARDS &	The Research in Color Foundation Fellowship	2022
FELLOWSHIPS	Simply Excellent Award, Nielsen	2018
	First Place Team, UChicago Health Tech Hackathon	2018
	General Honors, The College of the University of Chicago	2016
	Dean's List, The College of the University of Chicago	2012 - 2016
	Odyssey Scholarship, The College of the University of Chicago	2012 - 2016