SHIRUI (Oli) ZHOU

(571) 444-9537 | sz614@georgetown.edu | Washington, D.C.

EDUCATION

Georgetown University, McCourt School of Public Policy, Washington, DC Expected July 2024
Master of Science in Data Science for Public Policy GPA: 3.9/4.0
Advanced Modelling Techniques, Advanced Statistical Learning, Data Visualization, Massive Data Fundamentals, Network Analysis, Data Communication and Ethics, Congressional Power and Asia Policy, Comparative Politics

University of Nottingham, Ningbo, China July 2018 - July 2022
Bachelor of Science in Economics GPA: 4.0/4.0
Econometrics, Macroeconomics and Microeconomics Theory, Intermediate Mathematical Economic, Political Economy, Labor Economy, Experimental and Behavioral Economics, International Trade, Firm Strategy and Internationalization

SKILLS SUMMARY & CERTIFICATE

University of Zurich|Deep Dive into Blockchain Certificate July 2023 - July 2023
University of Chicago|Data & Policy Summer Scholar Certificate July 2019 - July 2019

- **Programming & Tools:** Python (Pandas, Sklearn, Matplotlib, Seaborn, NetworkX, Nature Language Toolkit), SQL (Advanced), PySpark, Tableau, Power BI, R (RCT, DID, RDD, IV), AWS, Google Analytics, MS Office, WordPress
- Advanced Knowledge and work experience in **Complex System Design, Decentralized Social Graph, Digital Governance**
- Effective communication skills in **English and Mandarin**, analytic and critical thinking, and time management skills

PROFESSIONAL EXPERIENCE

Ashoka (NGO) - ***Changemaker Index Product Team Data Intern*** | Python, Power BI May 2023 – August 2023

- Developed a data ingestion pipeline using **Python,** processing 20K user records from Drupal database through methods such as standardization and weighted scoring.
- Defined and computed over 20 KPIs using **Power BI** with integration to **Google Analytics**, leveraging custom SQL queries, to offer in-depth insights into the performance of survey platform.
- Curated diverse visualizations by **PyeCharts** and **PyVis** to construct region-level network maps, user feature matrix and distribution analysis, empowering partner organization to understand cohorts' strengths.

Matters (Web3 Media Company) - ***Research and Operation Intern*** | Python, MySQL, Network Analysis. Jun 2022 – Sep 2022

- Developed a detailed dashboard capturing NFT transaction trends using **MySQL**, covering metrics such as daily/monthly/weekly volumes, transaction prices, and distribution of holders.
- Employed **NetworkX** and **Plotly** to visualize the network of token holders improved community detection.

Intellisia (Intelligent Institution) - ***Data Analyst and Policy Intern*** | R, Tableau, International Policy Analysis Jun 2021 – Sep 2021

- Executed **data wrangling** on the CFPS database and created a streamlined multi-year database using **R**
- Analyzed potential **time-series correlations** by economic models, focusing on the relationship between the mental health of left-behind children and the migration status of their parents, culminating in a comprehensive report.
- Visualized vaccine importation trends in **Tableau** on 20+ countries to derive insights on geopolitical strategies.

University of Nottingham, Ningbo, China (On-campus)- ***Behavior Economics Research Assistant*** **|** STATA Jun 2021 – Sep 2021

- Collected primary data via **field research**, telephone interviews, and email outreach, specifically exploring discrimination based on titles that indicate social class in the UK, producing records including over 300 responses for further analysis.
- Leveraging **STATA** for data management and variable construction, labelling raw data and conduct exploratory analysis.

PROJECT EXPERIENCE

Network Analysis of Twitter to Identify Opinion Leader, Emotion Cascade and Community Structure

- Deployed **K-core decomposition** to examine the community structure, applied **NLP** including Name Entity Recognition, Topic Modelling and Sentimental Analysis on tweets to investigate the emotion cascade across 4 different time periods.

Predicting Attitudes toward UBI in EU using Machine Learning Techniques

- Built and trained Logistic Regression, Decision Tree, SVM, Random Forest, XGBoost and GBDT to identify 5 primary indicators and 35 secondary indicators to determine most important factor influencing EU citizens' attitudes towards UBI.

How the consensus built and involved through a voting mechanism? Using Taiwan's voting regarding Uber compliance issue on Polis as an example

- Using **PCA** and **UMAP** to visualize the participants' stance on a 2-demensional map, use **K-means** to cluster and classify group A and B, and use centroid coordinates calculation to get the distance between two opinion groups.