Yuanmo He

+44 7928549255 | y.he54@lse.ac.uk | yuanmohe.com | London

EXPERIENCE

Doctoral Researcher in Computational Social Science   09/2020 - present

The London School of Economics and Political Science (LSE)

Project 1: He, Y and Tsvetkova, M**.** A Method for Estimating Individual Socioeconomic Status of Twitter Users*.*

- **Developed a method** that uses correspondence analysis to estimate Twitter users' socioeconomic status (SES) with the brands they follow, based on classical social theories. The estimated SES achieved significant correlation (0.5-0.7) with conventional proxies for SES: income, education, and occupational social class.
- Used **R, Python, SQL, Azure Cloud Computing, Twitter API, and Google Geocoding API** to collect, process, clean and select **190 million** rows of data and estimated the socioeconomic status of **~3.5 million** Twitter users and **339** brands.
- Conference Presentations: **International Conference on Computational Social Science (IC2S2)**, online, 2021; The **Annual Meeting of the American Sociological Association** (Section on Inequality, Poverty and Mobility: New Approaches to Understanding and Addressing Inequality, online, 2021; General Online Research, online, 2021.
- Publication: forthcoming in ***Sociological Method & Research***. Preprint at https://arxiv.org/abs/2203.11636

Project 2: He, Y and Tsvetkova, M**.** Millions of Mobile-tracked Visits Reveal Socioeconomic Inequality of Daily Consumption. (Manuscript available upon request.)

- **Proposed a new theoretical argument** to incorporated insights from sociology, social psychology, and consumer research, and hypothesized that high SES is associated with more omnivorous consumption practices.
- Used **Python** to process and analyse data of all relevant points of interest (POIs) in the New York State (**24,598 POIs**), linking POI information and mobility pattern from **SafeGraph**, median income for Census Block Groups (CBGs) from **US Census**, and price level for brands from **Yelp**.
- Used standardised regression models to illustrate that the association between income and omnivorousness is significant at the 0.001 level, even after controlling for local availability and geographic mobility.
- Conference Presentations: **IC2S2**, Chicago, US, 2022 (presented by co-author); **European Conference on Social Networks**, London, UK, 2022; ODISSEI Conference for Social Science, Utrecht, NL, 2022

Class Teacher/Graduate Teaching Assistant (multiple posts), LSE   09/2021 - present

- Courses: MY470 Computer Programming, MY474 Applied Machine Learning, ME314 Introduction to Data Science and Machine Learning
- **Communicated** theoretical concepts and practical implementations of computer programming, machine learning, and natural language processing to students with **diverse academic, technical, and cultural backgrounds**.
- **Led** daily 1.5-hour computer lab teaching to two classes of 20+ students on understanding and performing **machine learning and quantitative text analysis with R** at the intensive three-week LSE summer school course ME314. Adjusted the teaching materials to be updated and relevant. Graded mid-term and final projects in R.
- Assisted weekly/biweekly 1.5/2-hour computer lab teaching to classes of 30+ students on **computer programming with Python & R and machine learning & deep learning with R**. Graded weekly/biweekly computer programming & data analysis assignments and final projects in R & Python with GitHub Classroom.

EDUCATION AND HONORS

PhD Social Research Methods, The London School of Economics and Political Science (LSE)   09/2020 – 09/2024
- Designing and executing research that uses advanced quantitative and computational methods to study socioeconomic inequality in daily behaviour and social interactions with large-scale digital trace data.
- Supervisors: Milena Tsvetkova and Kenneth Benoit.
- Awarded the **LSE PhD Studentship** for four years.

MSc Applied Social Data Science (Distinction), LSE   09/2019 - 09/2020
- Relevant modules: Computer Programming, Data for Data Scientist, Applied Machine learning, Quantitative Text Analysis, Multivariate Analysis and Measurement, Fundamentals of Social Science Research Design.
- **Distinction in all modules**.

BSc Social Sciences (First Class Honours), University College London (UCL)   09/2016 - 06/2019
- Relevant modules: Social Network Analysis, Causal Analysis in Data Science, Quantitative Research Methods
- Awarded the UCL Institute of Education and Society **Faculty Medal** (the best final year undergraduate student).
- Achieved **the highest final mark** in the Department of Social Science.

SKILLS

Programming language & statistical software: Python, R, SQL, Git, Stata, SPSS, Stan

Python packages: NumPy, SciPy, pandas, scikit-learn, Matplotlib, NetworkX (non-exhaustive)

R packages: tidyverse, tm, quanteda, glmnet, tree, randomForest, e1071, igraph, statnet (non-exhaustive)

Advanced Data Analysis: machine learning, natural language processing, social network analysis, multivariate analysis, causal analysis, parallel computing, cloud computing

Languages: Chinese (native), English (full professional proficiency)