Wenqing Qian

EDUCATION

Master of Science in Survey and Data Science *Exp. April 2024*
Institute for Social Research, University of Michigan, Ann Arbor, United States
Bachelor of Science in Statistics *June 2022*
School of Management, Fudan University, Shanghai, China

RESEARCH EXPERIENCE

Data quality assessment of the Life History Mail Survey (LHMS)
University of Michigan *January 2023-Present*

- Item missing rates of LHMS data and their correlates
  - Developed rules to identify missing items in LHMS data based on questionnaire design and constructed indicators to measure the extent of missingness.
  - Conducted literature review on factors influencing item nonresponses.
  - Described and visualized missing patterns and their variation across survey waves, different sections of the questionnaire, and respondent groups.
  - Performed Kruskal-Wallis tests and Dunn's tests to investigate the correlation between item missingness and respondent-specific covariates.
  - Designed and fitted a beta regression model that integrated stability selection techniques to filter essential predictors and model item missing rates.
  - Evaluated the impacts of selected covariates on the data quality using marginal effects and formulated follow-up models based on previous results to assess potential interventions among inter-correlated predictors.
- Consistency check on LHMS data
  - Examined the consistency across related questions in LHMS and compared the congruity between LHMS data and other data products of the Health and Retirement Study.
  - Analyzed the variation in data consistency across respondent groups defined by characteristics of interest using Chi-squared tests.

Sequence analysis on residential movement based on data from LHMS
University of Michigan *September 2022-January 2023*

- Cleaned and coded grid data about residential movements in LHMS into sequences using R.
- Carried out sequence analysis on residential movement data and performed clustering algorithm to investigate movement patterns.
- Conducted literature review on missing data imputation methods applicable to sequence data.

Demographics of Weibo users in China
Max Planck Institute for Demographic Research *March 2021-Present*

- Carried out preparatory research on social media usage in China and algorithms applied by the Weibo Advertising Platform.
- Programmed a web crawler in Python to collect and organize user coverage estimates from the Weibo Advertising Platform. Collected and cleaned 2020 census data and the latest available socioeconomic data.
- Performed exploratory analysis and visualization on Weibo's penetration and user composition.
- Analyzed outliers and the correlation pattern between the demographics and socioeconomic statuses of Weibo users. Adopted compositional data analytic techniques to summarize the demographic characteristics of Weibo users.
- Conducted spatial analysis on the geographical distribution of penetration rates and sex ratios of Weibo users.

PUBLICATION

Qian, W., Hexel, O., Zagheni, E., Kashyap, R., & Weber, I. (2023). Demographic inequalities in digital spaces in China: the case of Weibo. *Workshop Proceedings of the Seventeenth International AAAI Conference on Web and Social Media (ICWSM-23)*. AAAI Press. http://doi.org/10.36190/2023.01.

POSTER PRESENTATION

The Demographic Profile of Weibo Users — *April 2023*
PAA Annual Meeting — New Orleans, LA, US
The Demographics of Weibo Users in China — *June 2022*
European Population Conference — Groningen, NL

WORK IN PROGRESS

Respondent Correlates of Item Missing Rates in LHMS Data. *(Supported by the Health and Retirement Study)*

COURSE PROJECT

Correlates of internet usage in US seniors

- Framed research questions and conducted a literature search on internet usage among the elderly population in the US.
- Developed unweighted and weighted logistic regression models to explore the effects of respondents' features on their internet adoption.

Bayesian statistical analysis on influencing factors of songs' popularity

- Designed and fitted a beta regression model and a hierarchical beta regression model using JAGS to look for indicators associated with the popularity of songs on Spotify.

Estimating depression via Google Trends

- Proposed time series models to predict the development of depressive issues based on past Google Trends data on selected keywords.
- Built and launched an R Shiny app to visualize the temporal and geographical distribution of each type of depressive symptom and prediction results.

SKILL & LANGUAGE

Technology skills: R, Python, SQL, Java, MATLAB, SAS, SPSS
Languages: Mandarin (Native), English (Proficient), Japanese (Elementary)