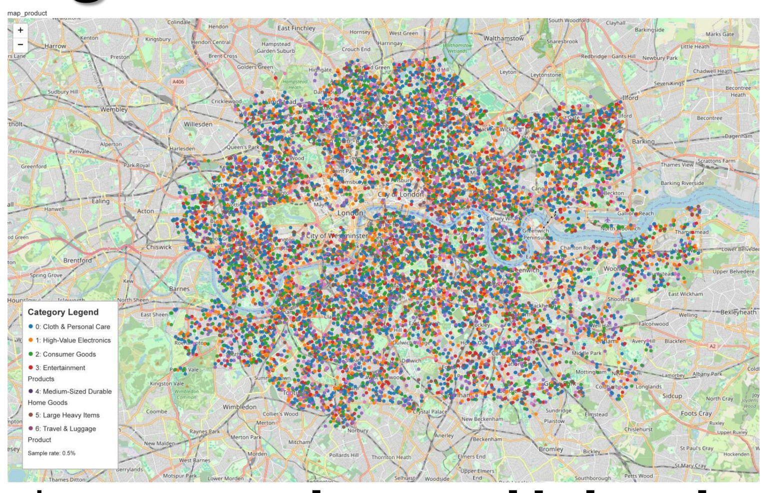
A Data-Efficient Demand Generation Framework for Multi-Agent Last-Mile Delivery via Fuzzy Logic

Yixuan Li
CID: 06013045
Supervisor: Prof Panagiotis Angeloudis

Abstract

With the rapid rise of e-commerce, last-mile logistics face growing demands for efficiency and sustainability. To support delivery optimization, demand-driven simulation is essential, yet existing demand-generation methods often rely on proprietary, opaque, and non-scalable data. This study introduces a population-based framework for synthetic order generation. A synthetic population is constructed from open datasets, while e-commerce purchasing behavior is modeled via fuzzy inference informed by industry reports. Orders are generated using Monte Carlo sampling, and validated through Jensen–Shannon Divergence against real-world data. A case study in Inner London shows that as population scale grows, synthetic orders better align with national purchasing patterns, confirming the model's validity. The framework offers scalability, interpretability, and reproducibility, providing a practical foundation for future last-mile delivery simulation and optimization.

Highlight



This study proposes a **demographic-based**, validated **synthetic order generation** framework and demonstrates its value for last-mile delivery simulation.

Background

Last-Mile Delivery Simulation

- E-commerce growth causes higher pressure on last-mile delivery
- Need for efficient & sustainable logistics
- Demand-dreiven logistics simulation to verify the effectiveness of refined stragteties

Research Gap

Current demand-generation methods are limited:

- **Reproducibility:** Opaque, proprietary data
- **Transferability:** Single-source, non-scalable
- **Heterogeneity:** Missing customer features

Research Objective

Design a Demand Generation Framework

- Utilize open, transparent data
- Easily extend and interpret
- Incorporate customer purchase behavior

Project Framework

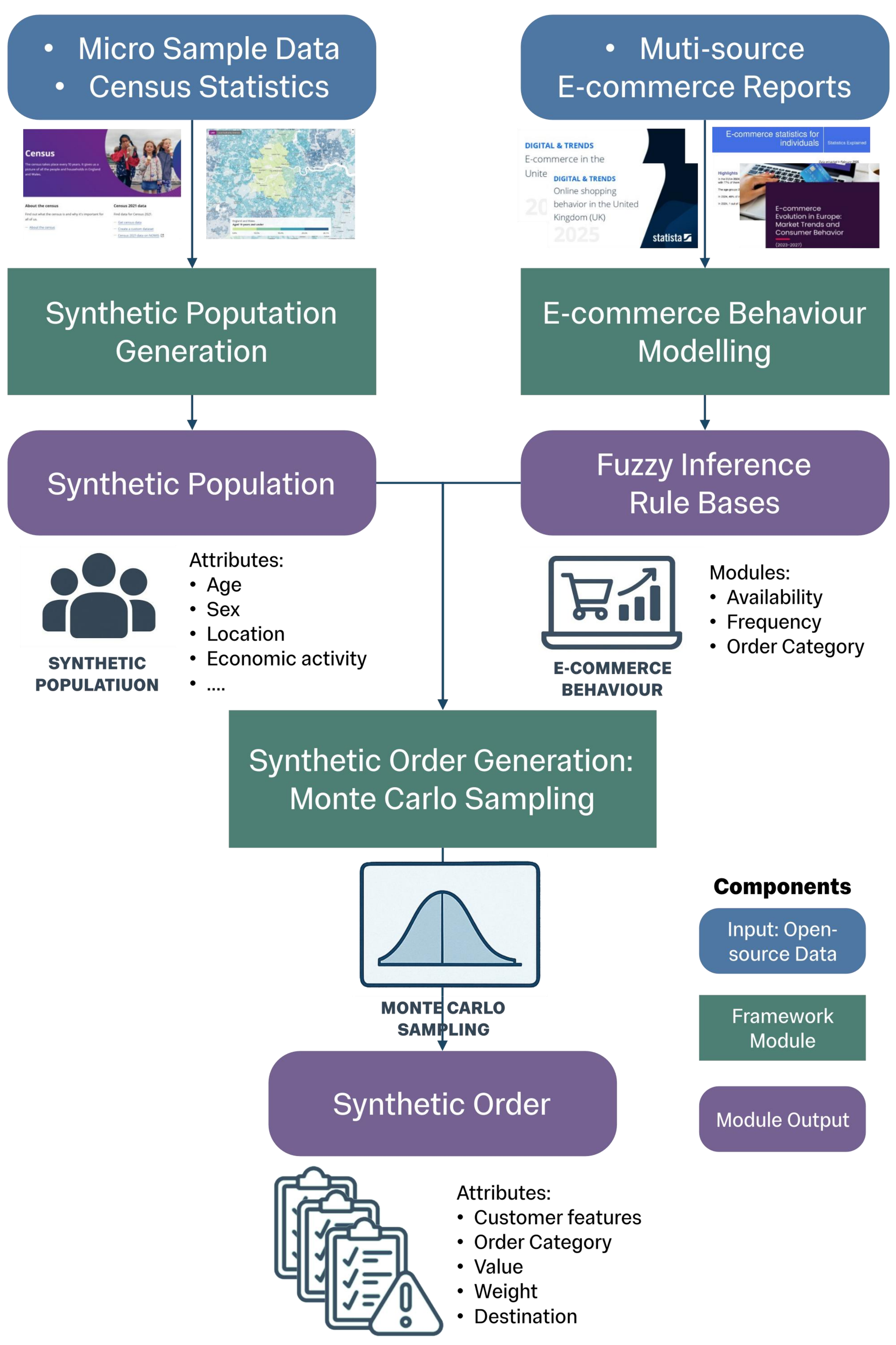


Figure 1 Project Framework

Study Area & Data

Inner London Area

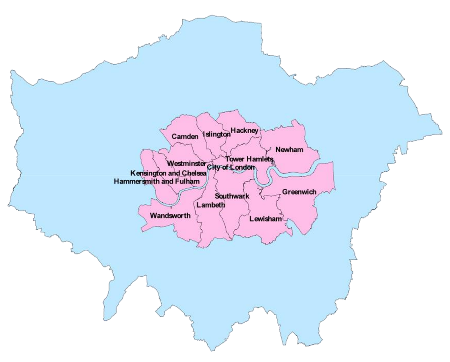


Figure 2 Study Area

Figure 3 Data Granularity: LSOA

Data

All data sources used are open-source data.

Synthetic Population

- UK 2021 Census Statistcs
- UK 2021 Census Micro-sample data

E-commerce Behaviour Modelling

- Research on purchase preference
- Reports of E-commerce Industry

Methodology

Synthetic Population

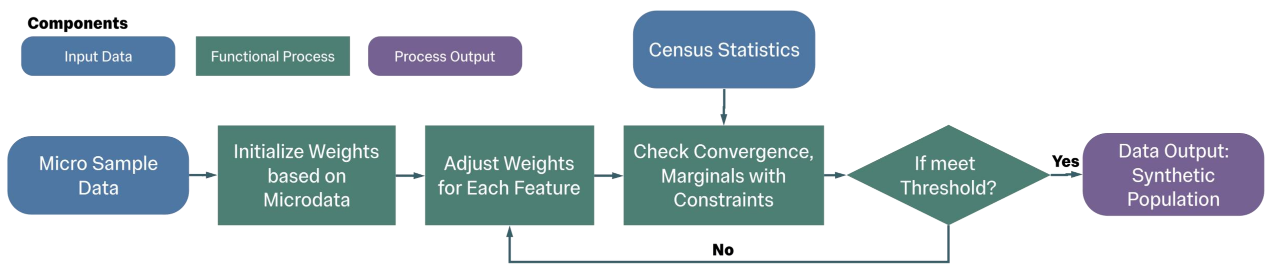


Figure 4 Flowchart of Iterative Proportional Fitting

- Iterative Proportional Fitting

Synthetic Order

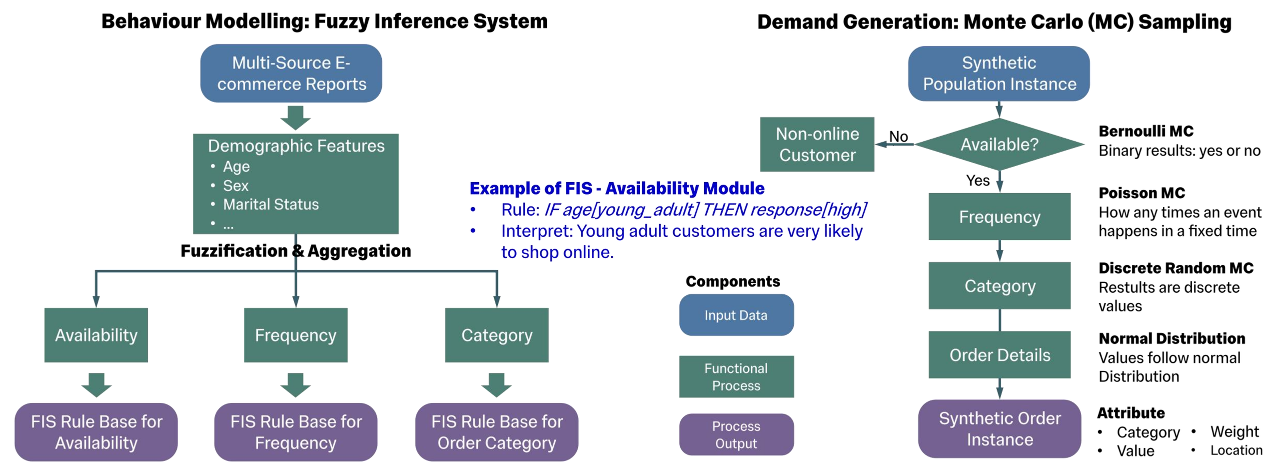


Figure 5 Flowchart of Fuzzy Inference System and Monte Carlo Sampling

Results

Validation Result

- Indicator: Jenson-Shannon Divergence (JSD)
  - Measures similarity between synthetic data and real-world statistics, smaller JSD, higher similarity and fidelity.

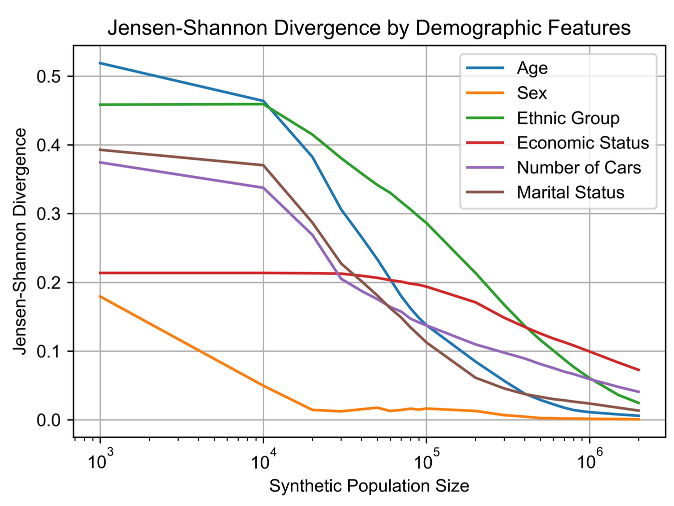


Figure 6 JSD by Demographic Features

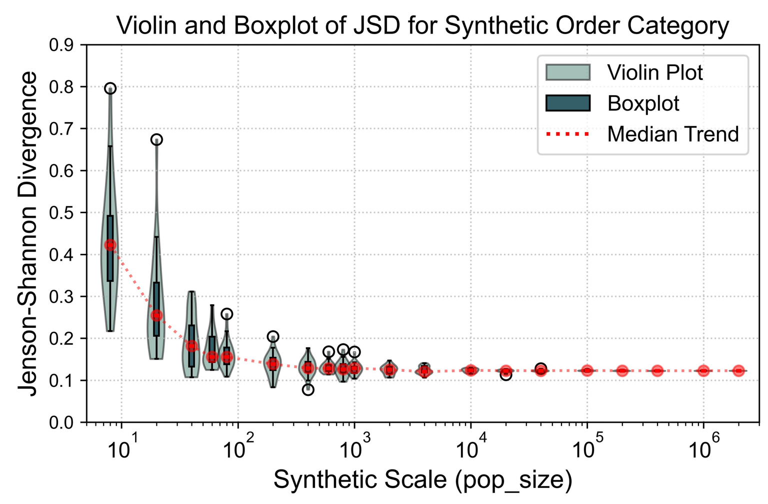


Figure 7 JSD by Order Category

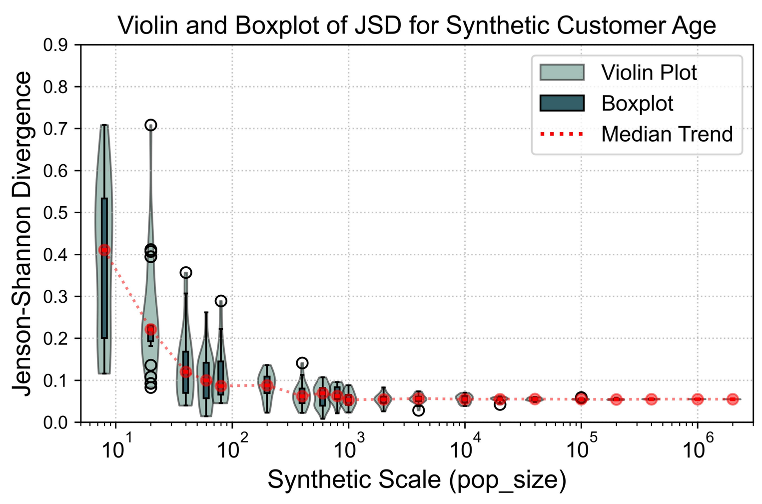


Figure 8 JSD by Customer Age

- As the synthetic population size increases, the JSD steadily decreases and stabilizes, confirming that the proposed framework can generate synthetic populations and orders that closely match real-world distributions.

Visualisation

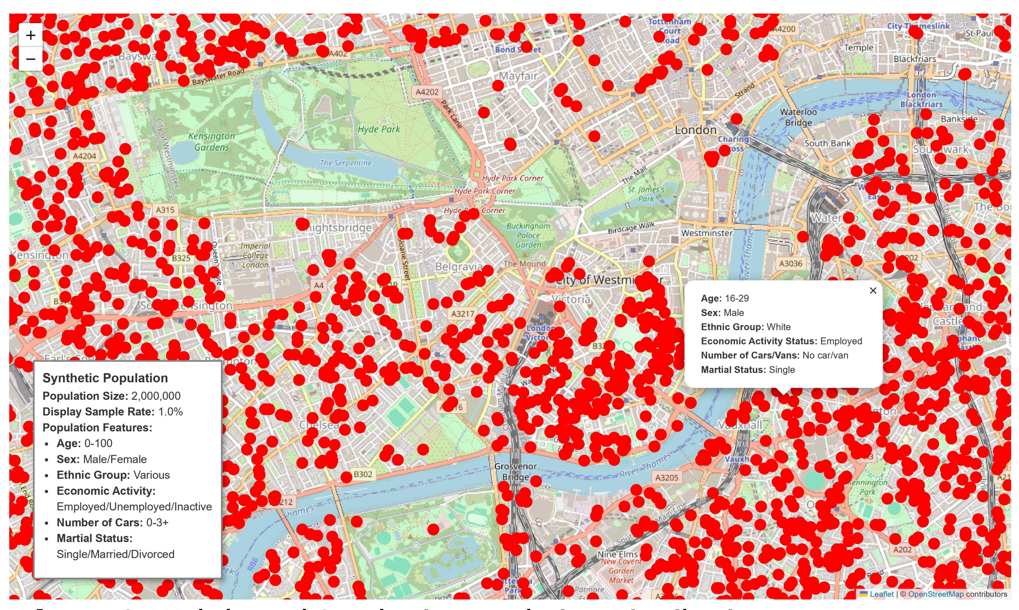


Figure 9 Web-based Synthetic Population Distribution

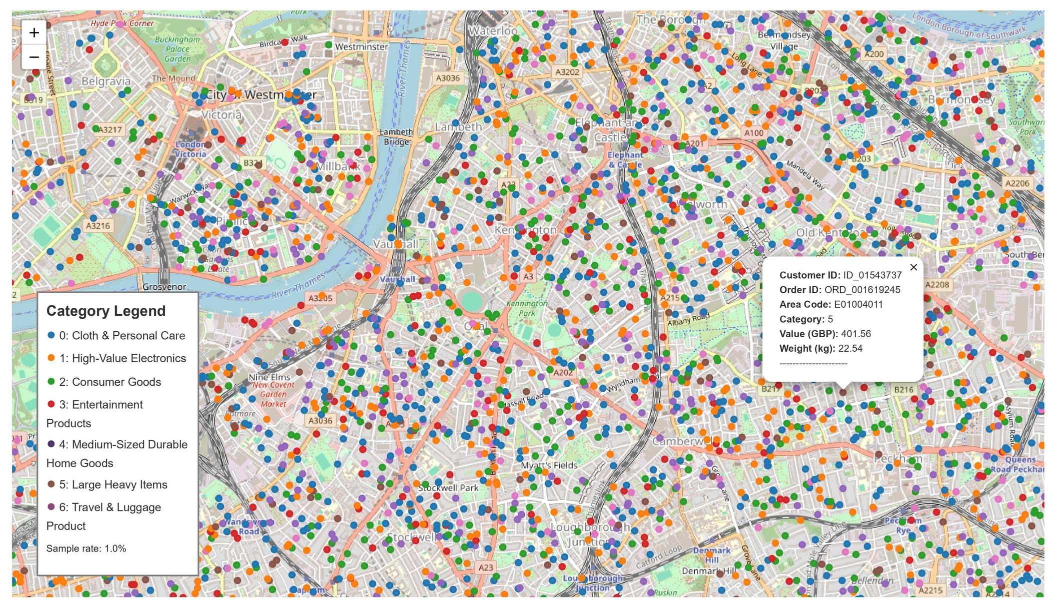


Figure 10 Web-based Synthetic Order Distribution

Conclusion

This study tackles data privacy and data scarcity in last-mile delivery by proposing a population-feature-based framework for generating synthetic orders. Validation results using JSD show that larger-scale synthetic instances (over 100) better capture consumer preferences, ensuring realistic order volumes for logistics simulations.

The method is open-source, reproducible, and extensible, enabling incorporation of additional demographics and delivery preferences (e.g., time windows, modes) to support diverse and fine-grained last-mile logistics simulations.

Project Github