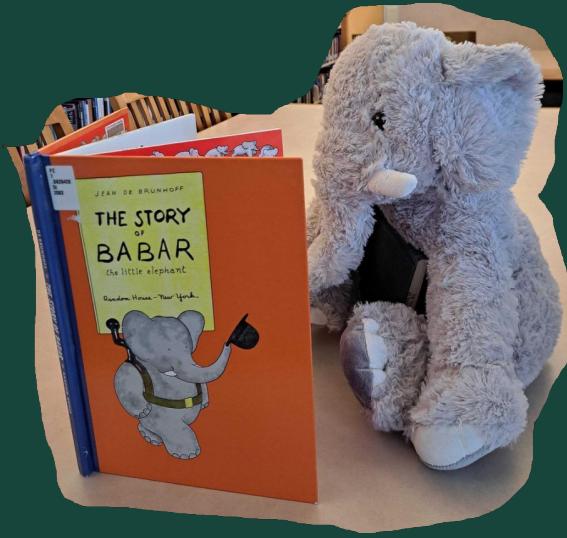


How U.S. Public Libraries Evolve with Community Demographics and Needs?

Jiayi Zhang

Graduate School of Arts and Sciences, Data Analytics

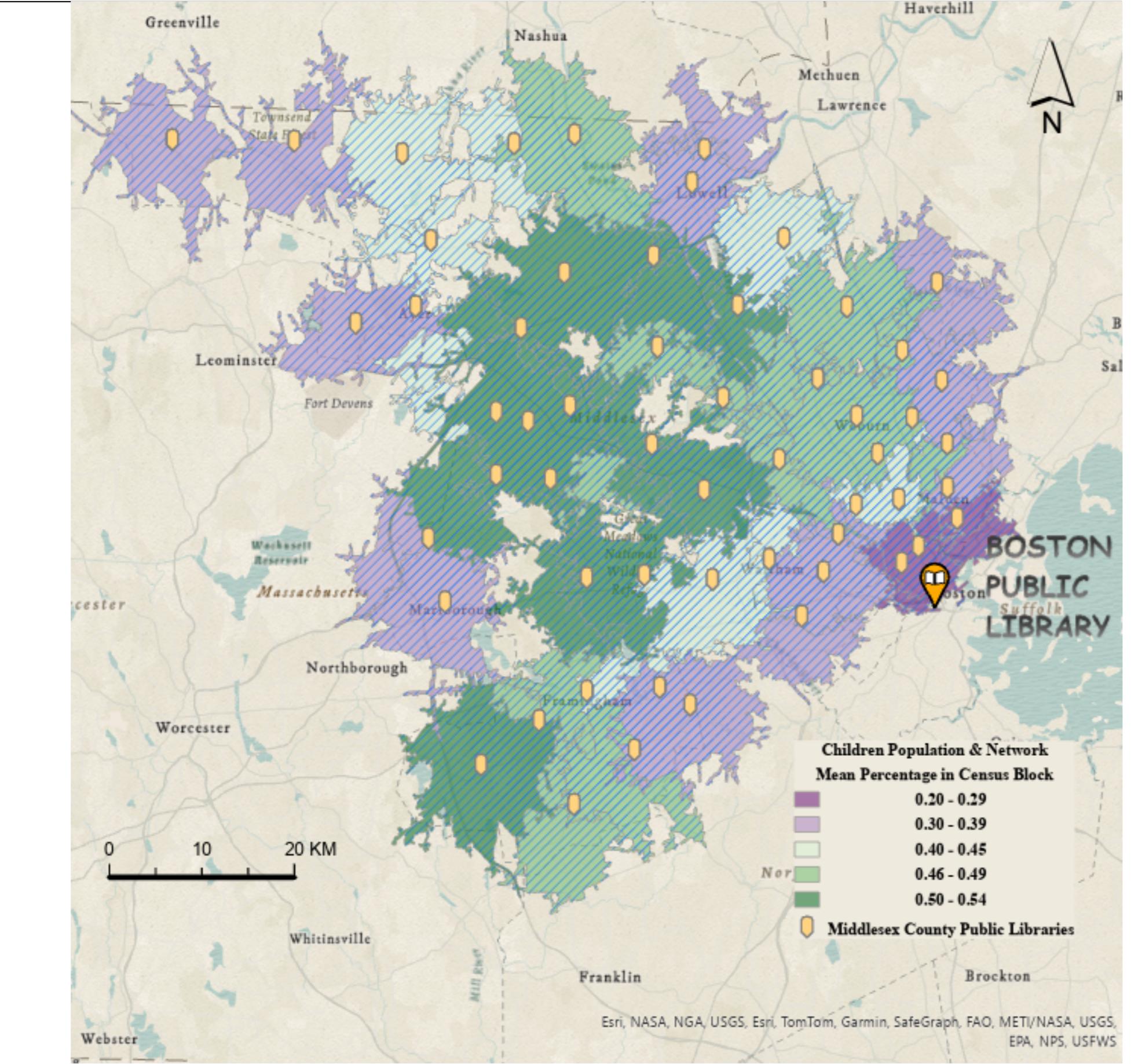


Boston Public Library Network

Random Forest selects the top nine most important features that contribute to future visits. KIDPRO, KIDATTEN, and KIDCIRCL represent the number of children's programs, number of children attendees, and number of children's items in physical circulation, respectively. These are the top three significant features.

Combined with results from forecasting and space cube point aggregation, the Boston Public Library serves as another case study for local demographics and libraries.

The children population percentage per census block in Middlesex County is summarized within the network. The polygon shown in green indicates areas with a high concentration of children, who could be potential attendees at the nearest public libraries.



Network Analysis

Overview: This analysis explores the connectivity and resource distribution among the County and Public Libraries within the established network. It examines the inter-library relationships and the flow of visitors and resources.

Methodology: It studies visitor flow patterns and resource sharing activities, with a particular focus on children's programs, which are pivotal in driving future visits as identified by the Random Forest analysis.

Findings: Libraries within these clusters, especially the Chicago (Boston) Public Library as a central node, demonstrate greater efficiency in program delivery and visitor satisfaction by distributing specialized resources that cater to the region's dynamic demographics. Based on these findings, recommendations include optimizing children's programming distribution and aligning resource allocation with demographic trends, crucial for meeting community needs and ensuring sustainable growth in library visits.

Space and Place

Reconsidering the role of public libraries and human needs in recent years, we recognize a shift in facility functions to accommodate people's behaviors. Instead of merely acting as conduits for information from physical materials, many studies raise the possibility that public libraries can become spaces for joint activities, including intensive meetings with friends, classmates, or colleagues to collaborate on common assignments or leisure activities.

This study focuses only on one of the potential enhancing factors, children's programs, and the target group, children attendees. Programs for children in libraries, most frequently on Saturdays, provide a space where children can behave naturally in ways that might not be allowed in many other learning-oriented environments. Ethnic groups, social functional groups, and behaviors related to affirmative action are all potentially important and need to be further explored in the future.

The idea raised in Aabø's 'Use of Library Space and the Library as Place' suggests that people need three places in their lives: home as the first place, work as the second, and a third place not defined by their social roles. Facilities like public libraries provide a low threshold for people to create or maintain their private space within a shared setting, which is crucial for gathering people in the same community.

Reference:

Svanhild Aabø, Ragnar Audunson. "Use of library space and the library as place." *Library & Information Science Research*, vol. 34, no. 2, 2012, pp. 138-149, ISSN 0740-8188, <https://doi.org/10.1016/j.lisr.2012.03.001>

Introduction

The modern public library began in the 19th century as a facility for storing physical materials and facilitating circulation. Since digital technology and cloud storage had not yet been developed, one of the main functions of public libraries was to satisfy the people's need for information retrieval and ingestion. Planning of Libraries in the social environment in the past were information-oriented. However, in the era of technological exploration, libraries played different roles, corresponding to shifts in people's behaviors and the evolution of society.

The project investigates two perspectives on public libraries, human behavior and contemporary society.

- 1. Longitudinal Study.** The first aspect examines how trends in public library usage have evolved over time to assess fluctuations in library visits across different years and predict the growing demand for specific library services. **Data:** Public Libraries Survey.
- 2. Cross-Sectional Study.** The second aspect involves integrating local demographic information with libraries to leverage the distribution of visitors' characteristics, enhancing library services. **Data:** American Community Survey (ACS).

Space Cube

To investigate the pattern change over time, the study merged five years of PLS data from 2018 to 2022 to perform space-time pattern mining. Aggregating the variables from 2D to 3D by the target cube variables 'total visits' and 'children attendees', the trends in clusters of point densities were further identified through emerging hotspot analysis. The children attendee clusters are included in the total visits clusters, which concluded that they are following the same pattern and the locations shown by deep green are recognized as significant.

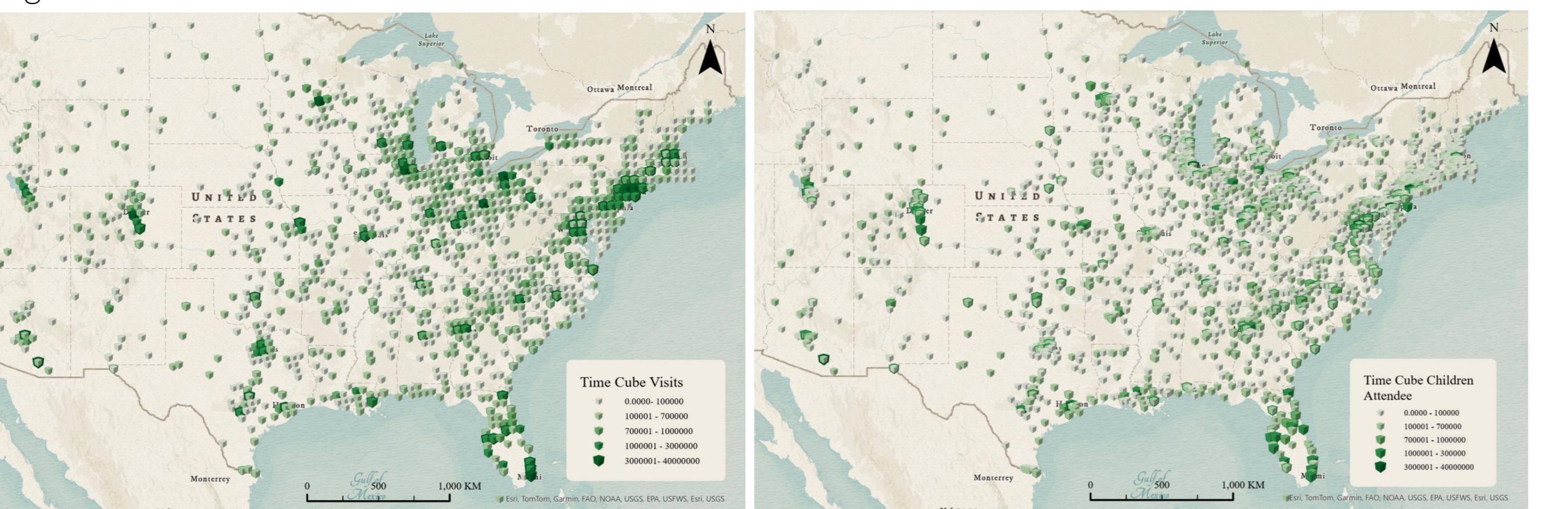


Figure 1. 3D Space Time Cube (Space Time Pattern Mining)

Forecast

Decomposing the cubes at each location into an exponential smoothing forecast model to identify the trends, the trend components will then forecast future time steps, one year ahead in this study, at each location. The light-colored label '5' represents the most significant area for the target, total visits as well as children attendees, scaled by mean RMSE.

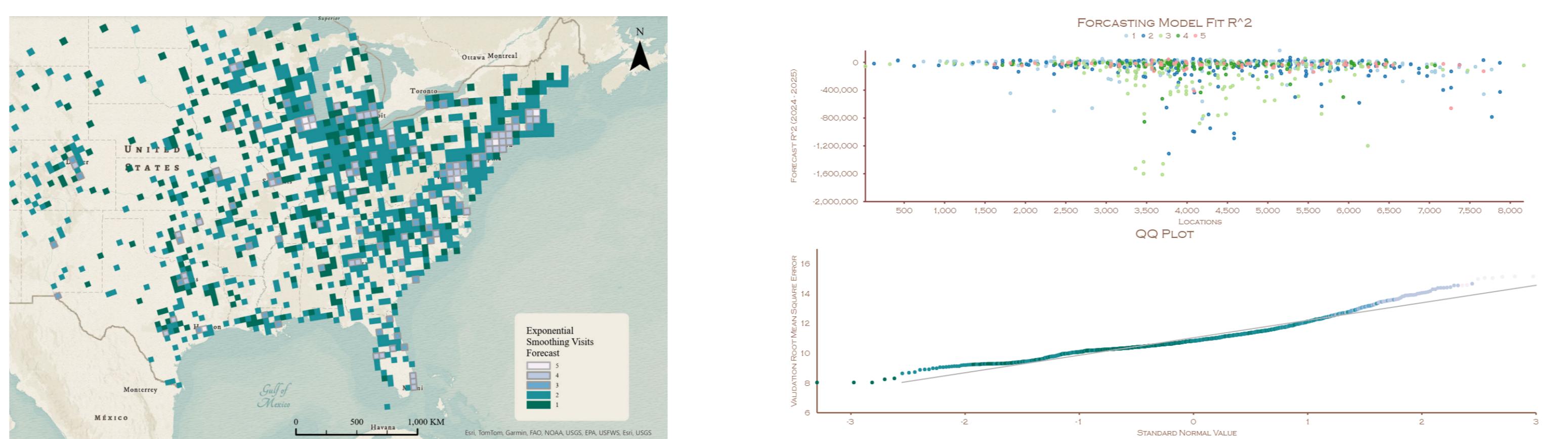


Figure 2. Exponential Smoothing Forecast Result for Public Library Visits, 2024 to 2025

Random Forest

Model Description

Forest-based and Boosted Regression:

$$n_i = \frac{N_t}{N} \left[\text{Imp} - \left(\frac{N_{t(R)}}{N_t} \times R \text{ Imp} \right) - \left(\frac{N_{t(L)}}{N_t} \times L \text{ Imp} \right) \right]$$

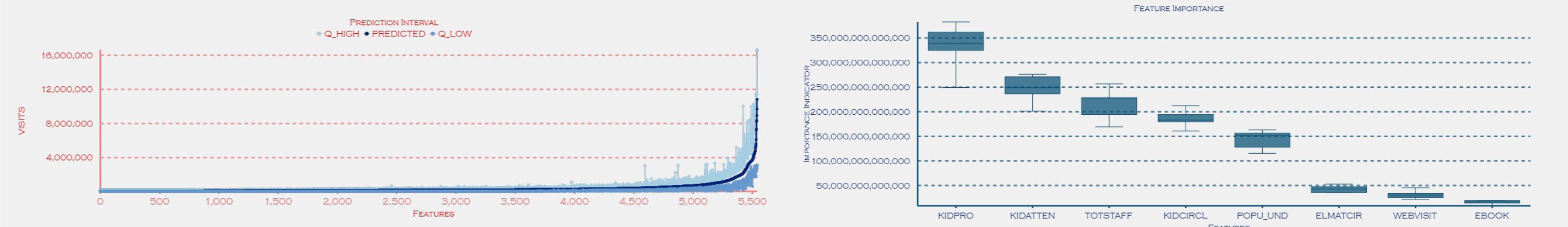
where:

- N_t : number of rows that particular node has,
- N : total number of rows present in the data,
- Imp: Gini Index,
- $N_{t(L|R)}$: number of nodes in the left and right nodes.

Random Forest is processed by creating 'trees'. Each tree represents a different portion of forecasting PLS data and concludes with its own prediction. The overall prediction compiles results from all trees, ranking the importance of n_i , and finally identifies the features that contribute most to the trends of future visits. The selected features from the prediction results will be those that public libraries should consider paying attention to if they aim to increase their visits in 2025.

Random Forest STDEV shown on map, a deeper color means significant variance in existing data pattern.

Feature Importance and Prediction Accuracy



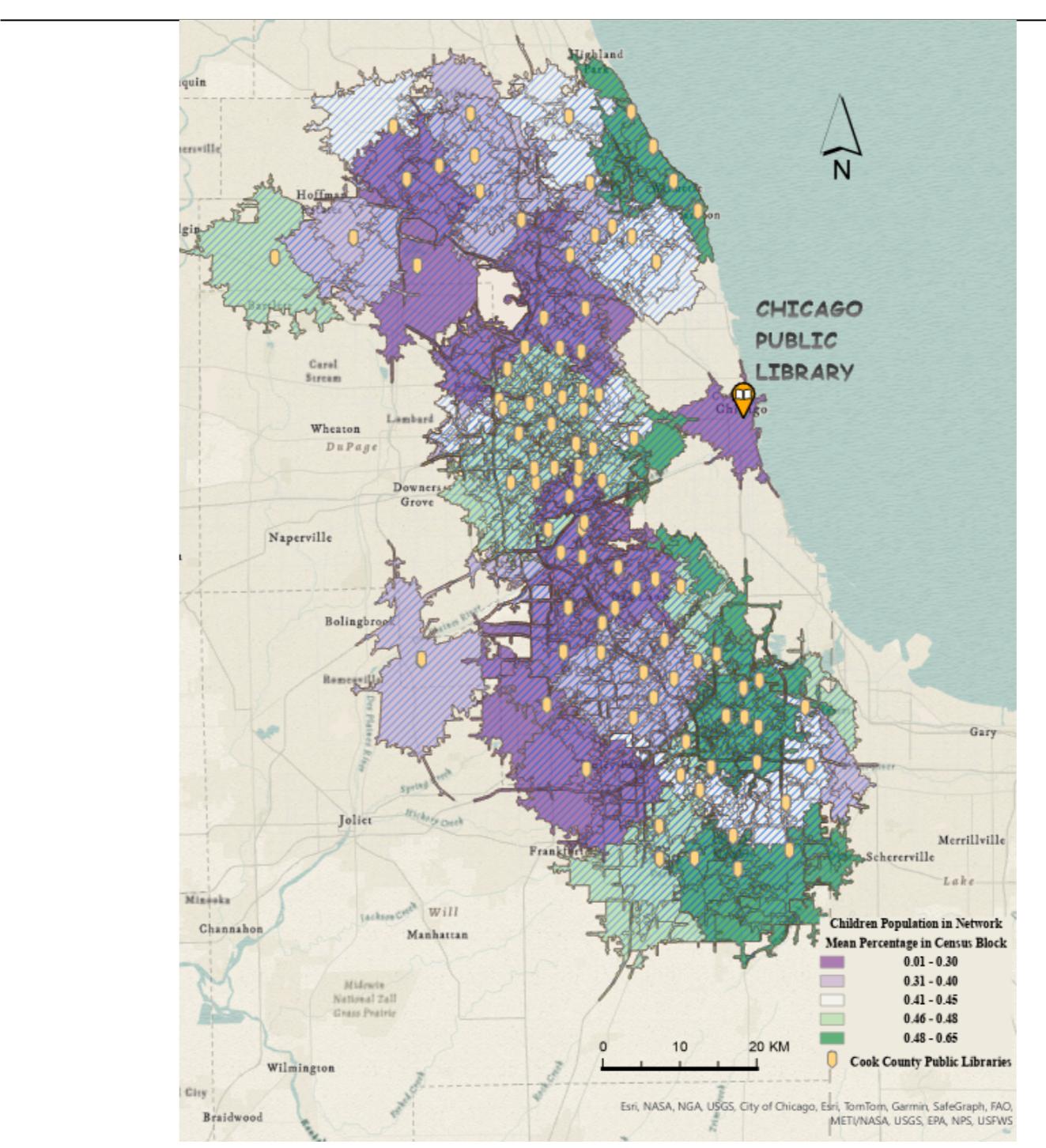
Chicago Public Library Network

Chicago Public Library has been selected for the study as a case to integrate with local demographic information, specifically the children population percentage per census block in Cook County.

All Cook County public libraries act as the nearest facilities, constructing the network with Chicago Public Libraries.

Summarized with the children population percentage, the color shown in green indicates the highest number of children population percentage in the network.

The libraries within the polygon might want to increase the number of children's programs to attract more visits in 2025, and perhaps also in the future.



Acknowledgment

I acknowledge the support of Prof. Sumeeta Srinivasan and TA Laurel Mire from UEP 235-02 Advanced Geospatial Modeling at Tufts University for their assistance in providing resources and support for this study, and the Reading Jumbo photo from Instagram @tischlibrary.