Beyond Books: Exploring the Societal Impact of Public Libraries in the U.S. Through Data Visualization

CPSC 490 Project Proposal (Fall 2022)
Pinyu (Pearl) Hwang
Adviser: Prof. Holly Rushmeier

Research Goals

This project aims to examine and present the societal impact of public library systems in the United States through data visualization. In addition to visualizing statistics specific to public libraries and public library systems such as circulation volume, population served, and size of holdings, I am also interested in exploring the connections, if any, between statistics on public libraries and demographic data that are accessible through the U.S. Census Bureau.

Background & Motivations

There aren't many truly public places left in America. Most of our shared spaces require money or a certain social status to access. Malls exist to sell people things. Museums discourage loiterers. Coffee shops expect patrons to purchase a drink or snack if they want to enjoy the premises.

—Jennifer Howard, "The Complicated Role of the Modern Public Library"

Public libraries form the backdrop of many fond memories for me, and they're one of my favorite places to be. Whenever I have to stay somewhere for longer than a few days, I find myself either within or looking for its public libraries: New York, NY; Baltimore County, MD; New Haven, CT; Seattle, WA; King County, WA…

And it's not just about the books (and audiobooks, DVDs, even video games), although they are great. The services of public libraries extend far beyond the multitude of materials they manage and lend out; public libraries provide its patrons with services such as internet and computer access, study spaces, programs, classes, and clubs for people of all age groups and a variety of interests, homework help and online tutoring, college prep services, job-hunting assistance and assistance with other social services, cultural events, access to online resources like databases and streaming services, information about community engagement, etc. The list can go on.

In an 2018 op-ed in the *New York Times*, sociologist Eric Klinenberg describes public libraries as *social infrastructure*, "physical places and organizations that shape the way people interact."

Moreover, in a time when the explosion of information online can often be confusing and overwhelming, libraries and librarians are still among the top sources of information trusted by Americans, according to a 2016 study by the Pew Research Center.

As such, I believe that the existence of public libraries is important to society, contributing greatly to the wellbeing and quality of life of the communities they serve. This is also a hypothesis that I hope to explore throughout the course of this project.

Potential Statistics of Interest

I plan to explore relationships between the following statistics:

1. Circulation of materials
2. Size of holdings
3. Population served
4. Library visits
5. Funding
6. Census data such as in regards to:
    - Poverty
    - Homelessness
    - Literacy/education
7. Internet access
8. Statistics and data specific to specific public library systems (tentative, depending on time and availability of data)

Potential Sources of Data

- Institute of Museum and Library Services: Public Libraries Survey
  https://www.imls.gov/research-evaluation/data-collection/public-libraries-survey
- Institute of Museum and Library Services: State Library Administrative Agency Survey
  https://www.imls.gov/research-evaluation/data-collection/state-library-administrative-agency-survey
- Census.gov (e.g. SAIPE datasets, level of education)
  https://www.census.gov/
- National Telecommunications and Information Administration (data on internet access)
  https://www.ntia.doc.gov/data/explorer#sel=libCommInternetUser&disp=map

Deliverables

- Processed data files
- 3-5 visualizations exploring a subset of relationships between statistics mentioned above
- Final report

Timeline

Dates	Objectives
Mid-Sep. – Oct. 5	**Background research + preliminary exploration of datasets** - Familiarize myself with available datasets (i.e. what statistics are available, what various statistics are measuring, etc.)
Oct. 6 – Oct. 12	**Data selection + data processing/cleaning** - Determine which statistics/fields are interesting to visualize - Extraction of relevant fields from datasets
Oct. 13 – Oct. 26	**Reviewing d3.js + prototyping possible visualizations** - Explore various combinations of data and visualization idioms
Oct. 27 – Nov. 2	**Finalizing visualizations + set up webpage(s)** - Finalize combinations of data and idioms that will be used - Begin setting up a basic webpage on which to host the final visualizations
Nov. 3 – Nov. 16	**Building final visualizations**
Nov. 17 – Nov. 23	**Refining + documentation**
Nov. 24 – Dec. 7	**Final report + submission**