

The Census in Library Applications

Census data is invaluable for practitioner-based library work as well as LIS research. Census data can be used to help public libraries understand their communities, so they can provide relevant services and meaningful collections. Coupled with GIS, census data can also be used to measure library access and availability and for the purpose of siting new libraries. In this chapter we will briefly review the literature and methods for using census data and GIS to study library issues.

Public libraries have a mission to serve communities within a defined legal area, such as a county, city, or town. Studying census data for the county or town can help the library understand the community it serves and can help inform decision-making. How many children are there relative to senior citizens? Is there a large population of residents who speak a language other than English at home? Is there a large percentage of people who don't have computers at home? What is the unemployment rate? The answers to all of these questions can inform library decisions in terms of what materials to buy and what programming to conduct in terms of workshops, lectures, and other events.

Beyond the legal service area, the LIS literature has recognized that there is a functional service area where the library attracts more people who live and work nearby as opposed to those who are farther away. In county and city library systems with multiple branches, studying the community that's in the neighborhood will be more meaningful than studying the entire city or county. Neighborhoods are informal areas that are not strictly defined, and the Census Bureau does not delineate or publish data for anything resembling a neighborhood. To study neighborhoods, one needs to use small census geography such as block groups or census tracts and aggregate the data for these geographies to form neighborhood-like areas based on formal and informal definitions used in that local area. Alternatively, one could select

all census geographies at a given distance of a library and aggregate those areas to measure the population of the functional service area.

GIS is usually coupled with census data for doing these types of analyses, given its ability to collate data from multiple sources in the same geographic space, aggregate smaller geographies and data into larger units, and select or measure features based on geographical relationships or distance. The datasets produced by the Institute of Museum and Library Services (IMLS) as part of the annual Public Library Survey provide GIS and attribute data for library systems and outlets (individual library facilities) in the United States.

There are different methods for measuring library service areas, and the principal ones were outlined in an early study by Jue and colleagues (1999): container, fixed-radius or buffer, and a gravity-based or distance approach. For the container method, the census geography where the library falls serves as the de facto service area. This method is the least desirable, as the shape of the area may not meaningfully represent the true service area. The buffer approach is more widely used, where a circle is drawn around the library at a given distance and census areas within that circle are selected as part of the library's service area. An alternative to the buffer method is a Thiessen or Voronoi polygon approach, where a boundary is drawn between each library and its nearest neighbors to delineate nonoverlapping service areas for every library. With a distance-based approach, the distance from each census geography (an origin) is measured to the nearest library (a destination), and geographies are assigned to the library based on closest distance.

A number of factors must be considered in each of these analyses. For the buffer and Thiessen approaches, a method must be chosen for selecting the census areas that fall within the buffer or polygon, as the census areas may cross the boundaries of these features. One could assign all census areas that

partially fall inside a buffer to that library, or assign areas that have a certain percentage within the buffer, or assign areas whose geographic center is contained within the buffer (Schlossberg 2003). The population data of the areas could be apportioned based on the percentage of area that falls within the buffer. This approach was widely used prior to the introduction of the ACS but is more problematic today given the low precision of ACS estimates at the block group and tract levels and the difficulty with apportioning these estimates. The size of the census geographies is a key factor: while it is desirable to use the smallest areas possible to limit overlap between census features and the library buffer, ACS data becomes more unreliable the smaller the unit is, and block-level data is limited to the decennial census.

Another consideration for the buffer approach is how large the buffer should be. Most studies use a range of one to three miles based on past convention. This is based on how far people would be willing to walk or generally travel, but more simply it represents an area that would meaningfully constitute a neighborhood or community, which would typically be small. An alternative would be to use variable buffers, based on the size of the library (and its ability to attract people from larger distances) or whether the library is in an urban or rural community (which affects willingness to travel lesser or greater distances).

The distance-based approach eliminates this problem, as every census area would be associated with the nearest library based on distance. With distance-based approaches, the primary issue is how distance should be measured. Simple straight-line or Euclidean distance is the easiest to compute, but network distance that follows actual roadways and paths is more realistic. Again, the size of the census geography plays an important role, as the geography serves as an aggregated summary of every person's individual home. Distances would be measured from the central point or centroid of each geography to the closest library. A better alternative to geographic centroids are population centroids, which are summary points representing how population is distributed within an area. The Census Bureau publishes population centroids for each DEC for block groups, tracts, counties, and states, and this data can be plotted in GIS and associated with tabular census data.

Figure 7.1 illustrates a one-mile buffer approach with census tracts and an urban public library. Census tracts 214.01, 218, and 219.02 have been assigned to the central library's service area; as none of the tracts are perfectly contained within the one-mile buffer, they have been assigned based on whether the center of the tract falls within the buffer. The population data for these three tracts would then be summarized to describe this library's service area. The

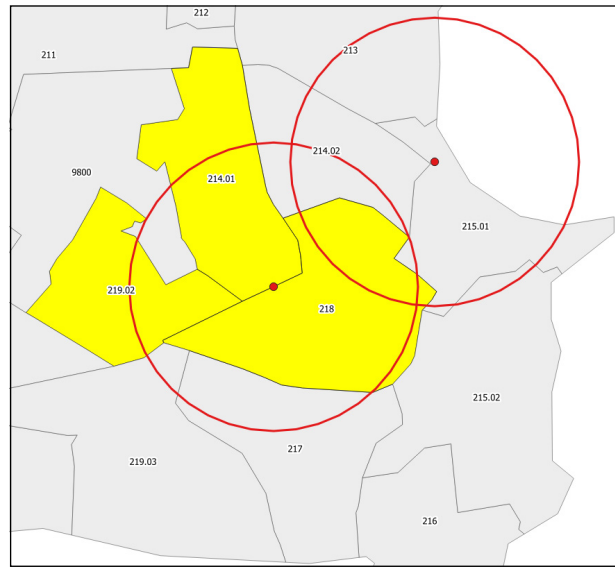


Figure 7.1
Using GIS to assign census tracts to libraries using one-mile buffers

second library to the northeast would be assigned the two tracts whose centers fall within its buffer. Smaller block groups could be used instead of tracts to minimize overlap, but if ACS data was being used, the precision of the estimates would be much lower. Alternatively, every tract could be assigned to a library based on distance to the closest branch.

Works by Palmer (1981) and Hayes and Palmer (1983) are among the earliest studies that reviewed the literature and measured library accessibility and service areas using a buffer-style approach in the pre-GIS era. Recent literature reviews by Bishop and Mandel (2010) and Mandel, Bishop, and Orehek (2020) summarize the applications of GIS to LIS research, which includes the use of GIS and census data for studying public library accessibility and distribution. Most of these studies focus on small areas such as counties or cities as case studies, while a few are national in scope. The earliest studies were practitioner-based and highlighted the utility of using GIS in conjunction with census data and library-generated data on circulation and patron registration to study library user communities, site new libraries, and understand library use and accessibility (Park 2012a; Bishop 2008; Hertel and Sprague 2007; Kinikin 2004; Ottensmann 1997). Other studies emphasized library policy in conjunction with social issues, such as the presence or absence of libraries in areas of high poverty (Jue et al. 1999), minority and low-income communities (Koontz, Jue, and Lance 2005), and the impact of library closures on different communities (Koontz, Jue, and Bishop 2009). Statistical analysis has been used to study the relationship between the locations of libraries and

the characteristics of users from both local (Japzon and Gong 2005) and national (Sin 2011) contexts. The application of buffer-based (Donnelly 2014) and distance-based (Donnelly 2015) approaches in conjunction with census data has been used to characterize the distribution of public libraries throughout the United States.

More recent work has focused on refining various aspects of the accessibility measurement process. The census counts people based on where they live, but many library users visit libraries based on where they work or don't make single trips to and from home and the library but combine them in a chain of trips (Park 2012b). Library users also use a variety of means to get to the library, including driving, public transit, biking, and walking, and this affects accessibility and distance that users are willing to travel (Allen 2019). It's important to consider these factors when studying library service areas and measuring accessibility and to remember that the approaches described here are all attempts to model reality. While a circular distance of one to three miles around a library may not perfectly measure the community or accessibility for all library users, it still serves as a general characterization of the area that likely constitutes the library's core population and generally measures the presence or absence of a library and whether a community will accrue its benefits or not.

The most recent studies demonstrate increasing internationalization, as researchers use census data in their countries and various GIS methods to study these issues in Wales in the UK (Page, Langford, and Higgs 2020; Higgs et al. 2018), Canada (Allen 2019), Hong Kong (Guo, Chan, and Yip 2017), South Korea (Lim and Park 2015), and Slovenia (Zakrajsek and Vodeb 2020), as well as in different cities in the United States (Cheng et al. 2021). These studies are also characterized by increasing complexity, where the focus has shifted to the refinement of models and methods and away from the practitioner and policy-driven applications of the earlier studies. There are exceptions, such as the studies in Wales, where the analyses are explicitly tied to the impact on society of cuts to library services and the need for informing public library policies. Recent practitioner-based works include studies on mapping rural library employment opportunities in Kansas (Lund 2017) and adult literacy outreach programs in Nashville (Rosichan 2020).

US census data is an invaluable public resource that supports policy and research about all facets of American society. For librarians, it's a resource that we rely on for meeting the diverse needs of our user communities, by connecting them to information they need to do their research, and for aiding our understanding of who our communities are so that we can better serve them.

References

- Allen, Jeff. 2019. "Mapping Differences in Access to Public Libraries by Travel Mode and Time of Day." *Library and Information Science Research* 41: 11–18.
- Bishop, Bradley Wade. 2008. "Use of Geographic Information Systems in Marketing and Facility Site Location." *Public Libraries* 47, no. 5: 65–69.
- Bishop, Bradley Wade, and Lauren Mandel. 2010. "Utilizing Geographic Information Systems (GIS) in Library Research." *Library Hi Tech* 28, no. 4: 536–47.
- Cheng, Wenting, Jiahui Wu, William Moen, and Lingzi Hong. 2021. "Assessing the Spatial Accessibility and Spatial Equity of Public Libraries' Physical Locations." *Library and Information Science Research* 43, no. 2: 101089. <https://10.1016/j.lisr.2021.101089>.
- Donnelly, Francis P. 2014. "The Geographic Distribution of United States Public Libraries: An Analysis of Locations and Service Areas." *Journal of Librarianship and Information Science* 46, no. 2: 110–29.
- . 2015. "Regional Variations in Average Distance to Public Libraries in the United States." *Library and Information Science Research* 37: 280–89.
- Guo, Yingqi, Chee Hon Chan, and Paul S. F. Yip. 2017. "Spatial Variation in Accessibility of Libraries in Hong Kong." *Library and Information Science Research* 39: 319–29.
- Hayes, Robert M., and E. Susan Palmer. 1983. "The Effects of Distance upon Use of Libraries: Case Studies Based on a Survey of Users of the Los Angeles Public Library—Central Library and Branches." *Library and Information Science Research* 5: 67–100.
- Hertel, Karen, and Nancy Sprague. 2007. "GIS and Census Data: Tools for Library Planning." *Library Hi Tech* 25, no. 2: 246–59.
- Higgs, Gary, Samuel Jones, Mitchel Langford, and Jesse Heley. 2018. "Assessing the Impacts of Changing Public Service Provision on Geographical Accessibility: An Examination of Public Library Provision in Pembrokeshire, South Wales." *Environment and Planning C: Politics and Space* 36, no. 6: 548–68.
- Japzon, Andrea C., and Hongmian Gong. 2005. "A Neighborhood Analysis of Public Library Use in New York City." *Library Quarterly* 75, no. 4: 446–63.
- Jue, Dean K., Christie M. Koontz, J. Andrew Magpanatay, Keith Curry Lance, and Ann M. Seidl. 1999. "Using Public Libraries to Provide Technology Access for Individuals in Poverty: A Nationwide Analysis of Library Market Areas Using a Geographic Information System." *Library and Information Science Research* 21, no. 3: 299–325.
- Kinikin, JaNae. 2004. "Applying Geographic Information Systems to the Weber County Library System." *Information Technology and Libraries* 23, no. 3: 102–7.
- Koontz, Christie M., Dean K. Jue, and Bradley Wade Bishop. 2009. "Public Library Facility Closure: An

- Investigation of Reasons for Closure and Effects on Geographic Market Areas.” *Library and Information Science Research* 31: 84–91.
- Koontz, Christie M., Dean K. Jue, and Keith Curry Lance. 2005. “Neighborhood-Based In-Library Use Performance Measures for Public Libraries: A Nationwide Study of Majority-Minority and Majority White/Low Income Markets Using Personal Digital Data Collectors.” *Library and Information Science Research* 27: 28–50.
- Lim, Ho-kyun, and Sung Jae Park. 2015. “Designing a GIS-Based Planning and Support System for a Public Library Building Project.” *Journal of Librarianship and Information Science* 47, no. 3: 254–64.
- Lund, Brady. 2017. “Rural Library Opportunity Zones: Mapping Rural Library Employment Opportunities Using Quantum GIS.” *Endnotes: The Journal of the New Members Round Table* 8, no. 1: 1–12.
- Mandel, Lauren Heather, Bradley Wade Bishop, and Ashley Marie Orehek. 2020. “A New Decade of Uses for Geographic Information Systems (GIS) as a Tool to Research, Measure, and Analyze Library Services.” *Library Hi Tech*, in press.
- Ottensmann, John R. 1997. “Using Geographic Information Systems to Analyze Library Utilization.” *Library Quarterly* 67, no. 1: 24–49.
- Page, Nicholas, Mitchel Langford, and Gary Higgs. 2020. “Exploring Spatiotemporal Variations in Public Library Provision Following a Prolonged Period of Economic Austerity: A GIS Approach.” *Area* 52, no. 2: 342–53. <https://doi.org/10.1111/area.12575>.
- Palmer, E. Susan. 1981. “The Effect of Distance on Public Library Use: A Literature Survey.” *Library Research* 3: 315–54.
- Park, Sung Jae. 2012a. “Measuring Public Library Accessibility: A Case Study Using GIS.” *Library and Information Science Research* 34: 13–21.
- . 2012b. “Measuring Travel Time and Distance in Library Use.” *Library Hi Tech* 30, no. 1: 151–69.
- Rosichan, Sharra. 2020. “Application of GIS in Mapping of Adult Literacy Outreach Programs and Potential Areas of Impact and Need.” *Public Library Quarterly* 39, no. 4: 346–62.
- Schlossberg, Marc. 2003. “GIS, the US Census and Neighbourhood Scale Analysis.” *Planning, Practice, and Research* 18, no. 2–3: 213–17.
- Sin, Se-Ching Joanna. 2011. “Neighborhood Disparities in Access to Information Resources: Measuring and Mapping U.S. Public Libraries’ Funding and Service Landscapes.” *Library and Information Science Research* 33: 41–53.
- Zakrajsek, Franc J., and Vlasta Vodeb. 2020. “Agent-Based Geographical Modeling of Public Library Locations.” *Library and Information Science Research* 42: 101013. <https://doi.org/10.1016/j.lisr.2020.101013>.