



Teaching Spatial Data Analysis; A Case Study with Recommendations

Journal:	<i>Nonprofit Policy Forum</i>
Manuscript ID	Draft
Manuscript Type:	Case Study
Classifications:	
Keywords:	Nonprofit education, spatial, Open data

SCHOLARONE™
Manuscripts

Teaching Spatial Data Analysis; A Case Study with Recommendations

Duncan J. Mayer^a
Robert L. Fischer^a

^aThe Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University. 11235 Bellflower Rd, Cleveland, Ohio, 44106, USA.

Duncan J. Mayer is a PhD candidate in social welfare at the Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University. His research centers on nonprofit management and community-organization dynamics.

Robert L. Fischer is associate professor in the Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, and codirector of the Center on Urban Poverty and Community Development. His research focuses on nonprofit accountability and the application of evaluation research methodologies in applied and nonprofit settings.

Declarations: Funding: No funding was received to assist with the preparation of this manuscript.

Conflicts of interest: The authors have no relevant financial or non-financial interests to disclose.

Keywords: Spatial, Programming, Open data, Nonprofit education

The data and code for the case study can be found https://github.com/duncan-mayer/spatial_teaching_case_study

Teaching Spatial Data Analysis

Abstract

Learning from data is a valuable skill for nonprofit professionals and researchers. Often, data have a spatial component, and data relevant to the nonprofit sector are no exception.

Understanding spatial aspects of the nonprofit sector may provide immense value to social entrepreneurs, funders and policy makers, guiding programmatic decisions, facilitating resource allocation, and possibly development policy. As a result, spatial thinking has become an essential component of critical thinking and decision making among nonprofit professionals. The goal of this case study is to support and encourage instruction of spatial data analysis and spatial thinking in nonprofit studies. The case study presents a local nonprofit data set, along with open data and code, to assist the instructors teaching spatial aspects of the nonprofit sector.

Pedagogical approaches are discussed.

Teaching Spatial Data Analysis

1 Introduction

Data are an increasingly important component of nonprofit operations, as managers and other organizational members regularly employ a range of data in an effort to evaluate or improve programs, communicate with stakeholders and donors, and satisfy accountability concerns (author). Educational programs focused on nonprofit management have taken note of this emphasis; a search of existing course offerings in the Seton Hall University database revealed 10 courses at 10 different universities focused on working with data (Mirabella, 2022). These included titles such as “data analytics for public and nonprofit managers,” “data analysis for social impact,” “data analytics/metric in the nonprofit sector,” and “nonprofit data-based decision making.” Importantly, data increasingly have a spatial component which can provide crucial context and assist in decision making (Huang & Wang, 2020). Yet, using the same database, a search for “spatial,” “space,” and “geography” returned no results. This is inconsistent with the repeated calls of nonprofit scholars for increased attention to spatial aspects of the nonprofit sector (MacIndoe & Oakley, 2022; Never, 2011, 2016) as well as in organizational studies more broadly (van Wissen, 2004). To illustrate the teaching of spatial data analysis in nonprofit studies, this brief article presents a case study of Cuyahoga County Ohio, including data and code, intended to assist with instruction of spatial data analysis and the importance of spatial reasoning in nonprofit studies.

2 The Importance of a Spatial Perspective

Nonprofit organizations often provide benefits to those in proximity to their service locations. These benefits may be part of services associated with the mission of the organization, or auxiliary benefits, such as employment, improved communication and goal alignment, or good will (Haslam et al., 2019; Marwell, 2004; McQuarrie & Marwell, 2009). The local benefits

Teaching Spatial Data Analysis

nonprofit organizations provide has resulted in a pragmatic emphasis on their spatial arrangement (Joassart-Marcelli & Wolch, 2003; Yan et al., 2014). This has led to consideration of metrics such as the concentration of nonprofits in counts and as a per capita measure. However, the spatial distribution of nonprofits is of theoretical interest as well, providing opportunities to test a range of theories (Carroll & Hannan, 2000). Yet, MacIndoe & Oakley (2022) suggest spatial dynamics of the nonprofit sector remain understudied and provide a range of questions that require spatial thinking and spatial data analysis.

Outside of policy makers and researchers, incorporating data into decision making and planning processes is increasingly important in management activities in nonprofit organizations (author). Never (2011) argued that maps are an essential tool for understanding the nonprofit sector and can help with identifying service gaps. Nonprofit professionals, including foundations, have taken note of these needs and over the past decade several initiatives have responded to the need for spatial integrating spatial information (see Roudebush et al., 2013). For example, in 2010 the Urban Institute's National Center for Charitable Statistics procured a grant to create a "Community Data Platform," a data tool with the expressed purpose of facilitating the use of local information with a spatial dimension (e.g., through geographic information systems, GIS) by nonprofit organizations. Brudney et al. (2016) interviewed key nonprofit stakeholders that used this GIS platform and found the organizations procured the local information to understand their community, seek collaboration, support programming, and obtain funding. Although highly valued by nonprofits, these initiatives often fail to persist as the implementation may lack community engagement or the nonprofits may lack the technical ability or financing to fully embrace the sustained use of the technology (Brudney et al., 2016; author).

Teaching Spatial Data Analysis

Spatial information has the potential to provide substantial value to nonprofit professionals, including managers, and they may also alter the conclusions for policy makers, philanthropists, and managers (MacIndoe & Oakley, 2022; author; Never, 2016). Yet our search revealed few case studies and materials available for instructors in nonprofit studies to illustrate the importance of spatial thinking. In the next section, this paper presents a case study, with open data and code, intended to be used to illustrate the advantages of entering a spatial dimension into nonprofit studies.

3 Case Study

The case study in this article is computational in nature, with the expressed purpose of facilitating the teaching of spatial thinking to nonprofit professionals. Accordingly, the accompanying code can be found in the authors’ GitHub, the link to which is in the appendix. GitHub is a reliable and widely used platform for the storage and dissemination of open-source software. It is recommended that readers download the associated zip file from the repository, which can be used to reproduce the analyses and figures in this paper and used in their classes to facilitate or enhance instruction. Although the data found in the repository is in a general format (e.g., .csv), the associated code is written in R (R Core team, 2020), an open-source language that is compatible with a range of existing analysis platforms.

The case study focuses on the spatial arrangement of the nonprofit sector, by census tract, in Cuyahoga County Ohio (USA) in 2016. The census tract is an apt choice of geography, as it is a widely used proxy for neighborhood and corresponds to theories related to the local benefits of nonprofit organizations. Further, while spatial dynamics may be at play in a range of situations, accounting for them is often most important with smaller, clustered units (Dale, 2014). Cuyahoga County is just over 1,200 square miles with a population over 1.2 million. Cuyahoga County is

Teaching Spatial Data Analysis

also an interesting location for a study of the nonprofit sector as it maintains a rich philanthropic history, home to some of the oldest community foundations and federated organizations, and has previously received scholarly attention in this journal (Roudebush & Brudney, 2012).

This case study focuses on two measures of the size of the nonprofit sector, density and mass. Consistent with the extant literature, density is defined as the number of active nonprofits in the census tract (Carroll & Hannan, 2000; Wo, 2018). Yet, density presents a single dimension of the nonprofit sector; it is not a direct measure of activity and neglects the variable size of organizations (Amburgey, 1996; Carroll & Hannan, 2000). Responding to this, several scholars have considered alternative measures that draw on nonprofits' financial information (Joassart-Marcelli & Wolch, 2003; Never, 2016). Accordingly, this case study considers mass in addition to density, where the latter is defined as the total revenue received by nonprofits in the census tract. All nonprofit information is drawn from the 2016 Business Master File (BMF) provided by the National Center for Charitable Statistics. The addresses of all nonprofits that have submitted tax documents in the previous two years, after extensive cleaning, were geocoded. Consistent with prior research, in the event that a nonprofit lists a post-office as their location, the post-office address is used (author; Yan et al., 2014). The successful geocoding rate was over 99 percent. The BMF is limited insofar as it may undercount the smallest organizations and those focused on religious services, however, is the best available data source for scholars working on the broader population of nonprofits in a region (Wo, 2018; Yan et al., 2014).

Given the local benefits of nonprofit organizations, a crucial question for the development of a nonprofit sector is the location of nonprofits in relation to need (Joassart-Marcelli & Wolch, 2003; Never, 2016; Yan et al., 2014), and scholars have taken different approaches to quantifying underlying community need. This case study uses the Neighborhood

Teaching Spatial Data Analysis

Deprivation Index (NDI) developed by Messer et al. (2006), which includes a series of variables from the US Census in the domains of education, employment, housing, occupation, poverty, residential stability, racial composition, marital status, inequality, as well as other variables (Buller, 2022). Higher levels of the NDI indicate a higher level of deprivation and need in the geographic area. Three census tracts have been removed from the sample for having a resident population of zero in 2016, which makes several variables used in the NDI undefined (e.g., percent of the population under the federal poverty line). The final sample contains 433 tracts within Cuyahoga County.

Table 1 shows the descriptive statistics for the variables in the case study. The table shows the average tract has a population of just under 3,000, with 8.5 nonprofits and just under 40 million in revenue. In total, Cuyahoga County shows a population of XXX nonprofits in 2016 based on tax filing data. The table also includes the distribution of nonprofits by primary purpose, showing the most common nonprofit types, on average, including human service nonprofits, or those in “other.” The NDI appears in Table 1 as well, which is unitless and ranges from just below 0 to .21.

[Table 1]

Several scholars have emphasized the benefits of exploratory data analysis when spatial components are present in data (MacIndoe & Oakley, 2022; Never, 2016). Good (1983) suggests the goal of exploratory data analysis is to present the data in a way that matches our ability to process information, identify non-random patterns, develop and refine hypotheses, and maximize expected utility by estimating the cost of computation and thinking. Although table 1 gives a variety of information about the distribution of each variable, it does not include the spatial element, and may obscure any spatial clustering or high-density regions.

Teaching Spatial Data Analysis

The presence of relevant spatial information suggests a map may be a better way to understand these data. Figure 1 shows four maps, where panel A displays the spatial distribution of nonprofit density and panel B shows the spatial distribution of nonprofit revenue log transformed, centered, and scaled to unit variance. Panel A shows the relative presence of nonprofit organizations across the county, with higher density in the city center (Cleveland) and in several suburban areas. Panel B shows the relative financial capacity of the nonprofits, with greater revenue in the city center and in the eastern suburbs. These maps are straightforward, however, additional descriptive statistics can be helpful to understand the spatial dimension of the county's nonprofit sector. There are number of well-developed measures of spatial autocorrelation, which call for different assumptions and provide different interpretations (see, Bivand & Wong, 2018 for a more comprehensive review of these measures and their implementation). Of course, calculating spatial autocorrelation requires information about the proximity of a unit to surrounding units. This is done with the information provided in the shapefile (see the case study documents), a simple format used to store geographic information. After reading the shapefile, we have several functions available to process the information for data analysis in the *spdep* package (see, namely, *spdep::poly2nb* in the associated code, Bivand et al., 2013).

[Figure 1]

MacIndoe & Oakley (2022) encourage the use of Moran's I, a measure of global spatial autocorrelation, which is typically bounded between -1 and 1. A positive or negative Moran's I suggests that values in a region are positively or negatively correlated with their neighbors. In this case study, the global Moran's I for nonprofit density and mass are 0.07 ($p < .001$) and 0.12 ($p < .001$), respectively. However, the global measure of spatial autocorrelation may belie

Teaching Spatial Data Analysis

important patterns across subgeographies that average out to a given global measure. To further explore spatial patterns, Panels C and D show the local Moran’s I, relaxing the homogeneity assumption of the global measure by providing a measure of spatial autocorrelation in each spatial unit (Anselin, 1995). Clearly, the figures identify spatial patterns that are likely to be non-random. Panel C shows several regions where nonprofit density tends to be positively clustered, including pockets of tracts in the north-central and north-east of the County. The local Moran’s I for revenue shown in Panel D show very different patterns over space. The pattern of revenue is generally smoother, (i.e., closer to homogenous) across the county, with several negative areas, suggesting revenue may be concentrated in one tract in those regions.

3.1 Bivariate spatial exploratory data analysis

The preceding discussion illustrates the importance of the spatial dimension of a county’s nonprofit sector. It shows that nonprofit density and mass are not uniformly distributed over space and may be highly concentrated in specific areas. However, Never (2011) argues that for funders and managers, effective maps must contain measures of organizational presence as well as “measures of public problem intensity” to allow donors and foundations to ensure “their funds are reaching organizations that are providing the services to a specific population at a time of need” (p. 177). This suggests that looking exclusively at measures of density and mass are inadequate for an exploratory analysis of the nonprofit sector, and scholars have long considered the co-occurrence of density and need (Joassart-Marcelli & Wolch, 2003; McDonnell et al., 2020; Never, 2016; Yan et al., 2014).

In 2016, nonprofit density and the NDI have a correlation of $-.16$ ($p < .001$). While this suggests that regions with higher deprivation may have slightly fewer nonprofits on average, it does not help guide investment for managers, planners, or foundations. A map that meets the

Teaching Spatial Data Analysis

requirements proposed by Never (2011) will help redress these limitations, and may involve a bivariate scale, which can be accomplished in either *ggplot2* or the *biscale* package, used for simplicity in this case study (Prener, 2022; Wickham, 2016). By splitting the data into evenly spaced quantiles, figure 2 illustrates that many tracts with higher nonprofit density have lower deprivation, particularly near the outer fringes of the county. The north-central area contains “downtown” Cleveland, which contains several tracts that appear to have relatively high need and nonprofit density. In contrast, the north-west region, containing the city of East Cleveland, has several areas with high deprivation and low nonprofit density. Though this figure presents the broad nonprofit presence by tract, depending on the issue area of focus, it could be recreated focusing only on specific types of nonprofits (e.g., educational, human services). This would better inform interpretation and action by more narrow categories of funders.

[Figure 2]

4 Discussion and Suggestions for Instruction

Data are increasingly important to the operations of nonprofit organizations, including tasks typically undertaken by managers such as community need assessment, program design, outreach and engagement. Accordingly, several nonprofit education programs have undertaken efforts to increase data literacy among nonprofit professionals. Yet, the spatial element, which benefits nonprofit professionals by providing essential context for philanthropic efforts, has received less attention in this area. In this environment, we have presented a case study, with associated code and data, which can be used to illustrate the benefits of incorporating a spatial perspective in nonprofit work. By encouraging a hands-on approach, this case study is consistent with theories of adult learning which suggest adults learn best when learning is applied and experiential. Engaging students in a experiential learning creates opportunities for a feedback

Teaching Spatial Data Analysis

loop, where experience itself brings further learning opportunities (Merriam & Bierema, 2013). For example, instructors may follow this case study and use students’ past experience to discuss information that may benefit the managers, foundations, or policy makers. These conversations can prompt discussion regarding communication with executives and policy makers in this area, for example, regarding the best ways to communicate the information to relevant stakeholders. Additionally, the associated data and existing code provide students a foundation to begin exploring finer-grain level considerations or alternative representations, sparking self-directed learning (Brookfield, 1991). Short of a hands-on experiential approach to this case study, instructors may incorporate the maps used in this case study to illustrate the importance of the spatial dimension, using these data to raise questions related to equity in access to nonprofit organizations, and the benefits they provide.

Schools of nonprofit and philanthropic studies, including those focused on management, have taken note of the need to incorporate data use into their courses. Despite ongoing calls emphasizing spatial elements of these data (MacIndoe & Oakley, 2022; Never, 2011, 2016), there is little presence of a spatial perspective in nonprofit course work. This case study has provided open data with code to assist in the facilitation of conversations and skill development related to spatial data in nonprofit studies.

Teaching Spatial Data Analysis

6 References

- Amburgey, T. L. (1996). Organizational ecology: Past, present, and future directions. *Academy of Management Journal*, 39(5), 23.
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R* (2nd ed. 2013 edition). Springer.
- Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3), 716–748. <https://doi.org/10.1007/s11749-018-0599-x>
- Brookfield, S. D. (1991). *Understanding and Facilitating Adult Learning: A Comprehensive Analysis of Principles and Effective Practices* (Reprint edition). Jossey-Bass.
- Buller, I., D. (2022). *ndi: Neighborhood Deprivation Indices*. The Comprehensive R Archive Network. <https://cran.r-project.org/package=ndi>
- Carroll, G., & Hannan, M. T. (2000). *The demography of corporations and industries*. Princeton University Press.
- Dale, M. R. T. (2014). *Spatial Analysis: A Guide For Ecologists* (2nd edition). Cambridge University Press.
- Good, I. J. (1983). The Philosophy of Exploratory Data Analysis. *Philosophy of Science*, 50(2), 283–295. <https://doi.org/10.1086/289110>
- Haslam, A., Nesbit, R., & Christensen, R. K. (2019). The Dynamic Impact of Nonprofit Organizations: Are Health-Related Nonprofit Organizations Associated with

Teaching Spatial Data Analysis

Improvements in Obesity at the Community Level? *Nonprofit Policy Forum*, 10(3).
<https://doi.org/10.1515/npf-2018-0040>

Huang, B., & Wang, J. (2020). Big spatial data for urban and environmental sustainability. *Geo-Spatial Information Science*, 23(2), 125–140.
<https://doi.org/10.1080/10095020.2020.1754138>

Joassart-Marcelli, P., & Wolch, J. R. (2003). The Intrametropolitan Geography of Poverty and the Nonprofit Sector in Southern California. *Nonprofit and Voluntary Sector Quarterly*, 32(1), 70–96. <https://doi.org/10.1177/0899764002250007>

MacIndoe, H., & Oakley, D. (2022). Encouraging a Spatial Perspective in Third Sector Studies: Exploratory Spatial Data Analysis and Spatial Regression Analysis. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*.
<https://doi.org/10.1007/s11266-022-00459-6>

Marwell, N. P. (2004). Privatizing the Welfare State: Nonprofit Community-Based Organizations as Political Actors. *American Sociological Review*, 69(2), 265–291.

McDonnell, D., Mohan, J., & Norman, P. (2020). Charity Density and Social Need: A Longitudinal Perspective. *Nonprofit and Voluntary Sector Quarterly*, 49(5), 1082–1104.
<https://doi.org/10.1177/0899764020911199>

McQuarrie, M., & Marwell, N. P. (2009). The Missing Organizational Dimension in Urban Sociology. *City & Community*, 8(3), 247–268. <https://doi.org/10.1111/j.1540-6040.2009.01288.x>

Merriam, S. B., & Bierema, L. L. (2013). *Adult Learning: Linking Theory and Practice* (1st edition). Jossey-Bass.

Teaching Spatial Data Analysis

Messer, L. C., Laraia, B. A., Kaufman, J. S., Eyster, J., Holzman, C., Culhane, J., Elo, I., Burke, J. G., & O'Campo, P. (2006). The Development of a Standardized Neighborhood

Deprivation Index. *Journal of Urban Health*, 83(6), 1041–1062.

<https://doi.org/10.1007/s11524-006-9094-x>

Mirabella, R. (2022). *Nonprofit Management Education; Current Offerings in University-Based Programs*. Seton Hall University. <https://academic.shu.edu/npo/>

Never, B. (2011). The Case for Better Maps of Social Service Provision: Using the Holy Cross

Dispute to Illustrate More Effective Mapping. *VOLUNTAS: International Journal of*

Voluntary and Nonprofit Organizations, 22(1), 174–188. [https://doi.org/10.1007/s11266-](https://doi.org/10.1007/s11266-010-9123-y)

[010-9123-y](https://doi.org/10.1007/s11266-010-9123-y)

Never, B. (2016). Place matters: The spatial effects of human service expenditures. *Nonprofit*

Policy Forum, 7(3), 20. <https://doi.org/DOI 10.1515/npf-2015-0025>

Prenner, C. (2022). *biscale: Tools and Palettes for Bivariate Thematic Mapping*.

R Core team. (2020). *R: A language and environment for statistical computing* (4.0.2). R

Foundation for Statistical Computing. <https://www.R-project.org/>

Roudebush, M. M., & Brudney, J. L. (2012). Making Policy Without Parameters: Obtaining Data

on the Nonprofit Sector in a Local Community. *Nonprofit Policy Forum*, 3(1).

<https://doi.org/10.1515/2154-3348.1044>

van Wissen, L. (2004). A Spatial Interpretation of the Density Dependence Model in Industrial

Demography. *Small Business Economics*, 22(3/4), 253–264.

<https://doi.org/10.1023/B:SBEJ.0000022232.12761.a9>

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.

<https://ggplot2.tidyverse.org>

Teaching Spatial Data Analysis

Wo, J. C. (2018). Understanding the density of nonprofit organizations across Los Angeles neighborhoods: Does concentrated disadvantage and violent crime matter? *Social Science Research*, 65(7), 56–71. <https://doi.org/10.1016/j.ssresearch.2018.01.001>

Yan, J., Guo, C., & Paarlberg, L. E. (2014). Are Nonprofit Antipoverty Organizations Located Where They Are Needed? A Spatial Analysis of the Greater Hartford Region. *The American Statistician*, 68(4), 243–252. <https://doi.org/10.1080/00031305.2014.955211>

For Review Only

Teaching Spatial Data Analysis

Table 1.

Descriptive statistics for Case Study Variables (Cuyahoga County 2016, N = 443)

Variable	Mean	SD	Min	Max
Resident population	2,841.33	1,348.65	145	10687
NDI	0.01	0.053	-0.07	0.21
Total Revenue ¹	3,903.89	41,003.48	0	754,438.30
Nonprofit Density	8.58	24.66	0	455.00
Human services	2.40	3.84	0	37.00
Education	1.26	3.83	0	69.00
Public	1.27	3.24	0	58.00
Religious	0.54	3.25	0	63.00
Arts	1.71	18.21	0	383.00
Health	0.56	1.36	0	11.00
Other	2.11	4.48	0	75.00
Note: All nonprofit information is from the 2016 BMF provided by the NCCS. NDI = Neighborhood deprivation index				
¹ In units of 10,000 2016 dollars.				

Teaching Spatial Data Analysis

Figure 1. Distribution of Nonprofit Density (A) Nonprofit (log) Revenue (B) and the Local Moran's I of each tract (C-D)

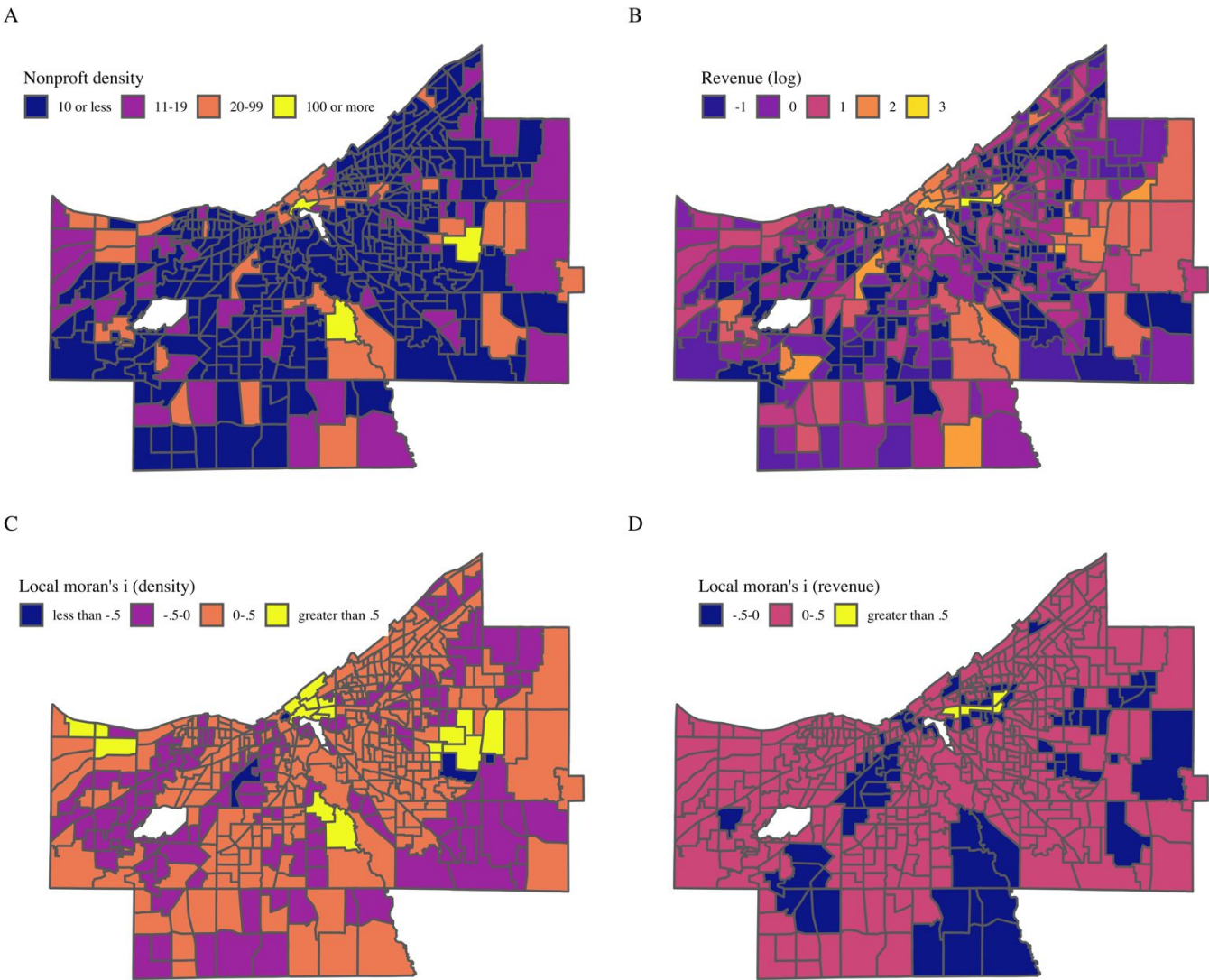


Fig 1. Panel A shows nonprofit density (discretized) by tract, panel B shows log revenue, after mean centering and scaling to unit variance. Panel C shows the local moran's I for nonprofit density, while D shows the local moran's I for log revenue.

Teaching Spatial Data Analysis

Figure 2. Bivariate Map of Neighborhood Deprivation and Nonprofit Density in Cuyahoga County in 2016 ($N = 433$)

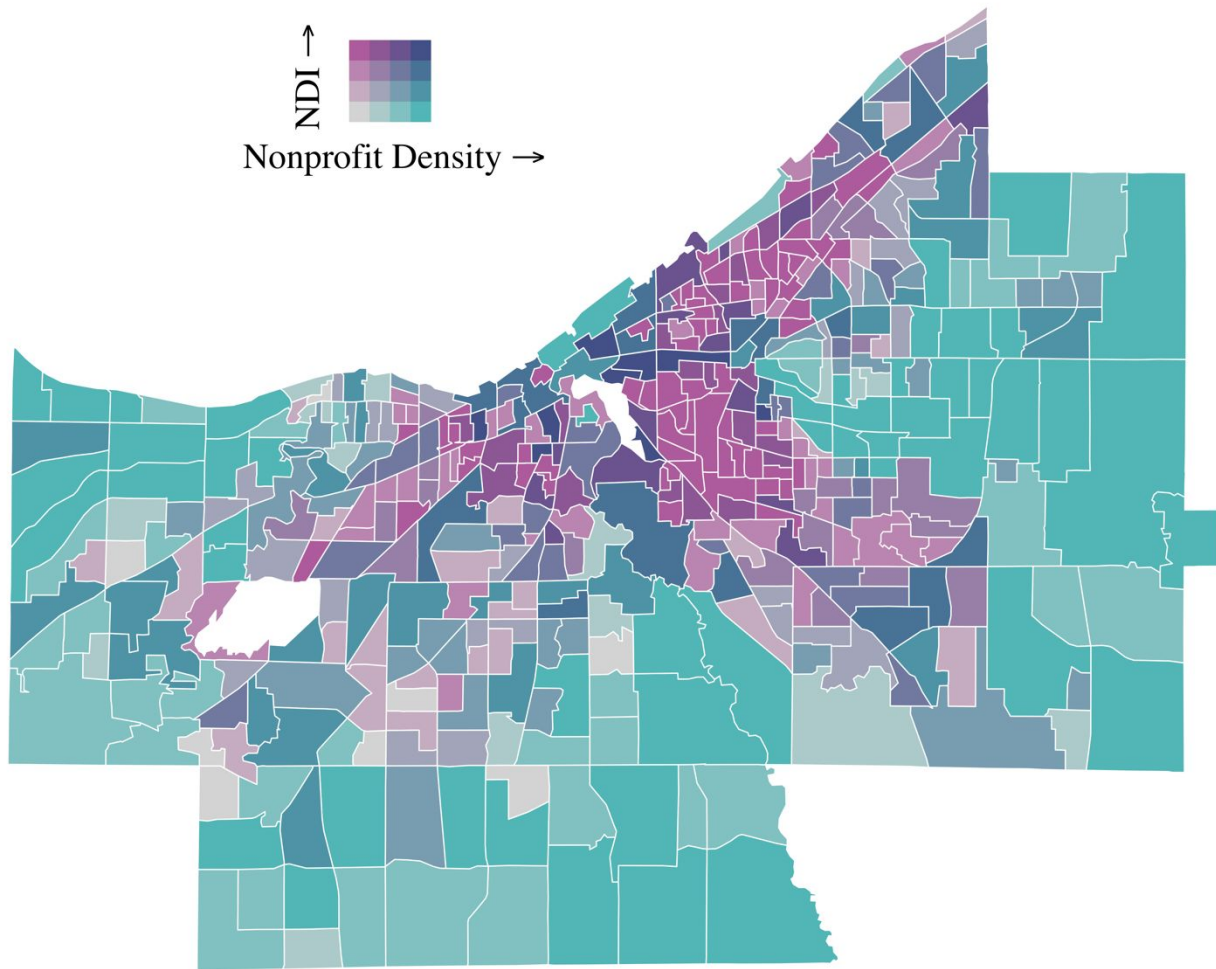


Fig 2. Bivariate map showing the neighborhood deprivation index developed by Messer et al., (2006), and nonprofit density.