Investigating the modifiable areal unit problem in Baltimore, Maryland

Proposal, GES Independent Study

Harrison DeFord

2025-03-06

### Introduction

Spatial data are and often used in decision-making, providing utility to disciplines including public health, urban resource management, crime mitigation, or disaster management. Across all domains, the choice of spatial units in analysis can change the patterns that emerge, influencing both the conclusions drawn from data and the decisions based on them. This issue, known as the Modifiable Areal Unit Problem (MAUP), pervades spatial analysis and must be carefully managed. The MAUP, first described by Gehlke and Biehl (1934), is manifested in two ways: the scale issue, where patterns are emphasized or suppressed based on the spatial scale of areal units, and the boundary issue, where data are separated by arbitrary borders which often have no relation to the spatial processes being investigated.

GIS software enables users to create and manipulate spatial data with ease. However, there are countless ways to aggregate data, depending on the pattern being analyzed and the context of its use. My analysis will explore different methods for defining spatial units and their impact on observed patterns, using 311 calls for service in Baltimore, Maryland, as a case study.

### Data sources

I will use Baltimore City historic 3-1-1 customer service requests from Open Baltimoreas my observational data, specifically l focusing on the “311 Customer Service Requests” for 2024since it is the most recent complete calendar year at the time of publishing. U.S. Census Bureau administrative units, including Census tracts (U.S. Census Bureau 2024), will be included the comparison of spatial units, as well as providing a framework from which to build novel areal units. In addition, Baltimore City maintains Neighborhood Statistical areas (Baltimore City Department of Planning 2024), which will be used as a standalone spatial unit as well as a starting point for the creation of novel units.

### Methods

While there is much research available regarding the effect of the MAUP across various existing administrative units, less attention has been given to spatial units developed to analyze more specific spatial processes. R Statistical Software (R Core Team 2024) and ArcGIS Pro (ESRI 2022) will be used to generate novel aggregation units, which units will be used to summarize 3-1-1 calls for service. These units will include the traditional administrative boundaries, which will then be manipulated via geoprocessing tools such as centroids, buffers, and isochrones. In addition, tessellations will be generated across the region to facilitate cell-based analysis. Descriptive statistics and geographic distributions created by each of the spatial units will be compared, allowing for a complete examination of the effectiveness of each type of aggregation as it relates to the spatial patterns of calls for service.

### Deliverable

The goal for this independent study is, as mentioned previously, a research paper suitable for publishing in an academic journal. The paper will contain written comparison of numeric distributions, as well as thematic maps depicting the geographic distribution of 3-1-1 calls when aggregated into various spatial units. This work is scheduled to be carried out over the summer, with an estimated completion in late August-early September, 2025. Maintaining this timeline will allow the authors to collaborate and edit the written portion of the document before submission to a peer-reviewed journal.

### References

Baltimore City. 2024. “311 Customer Service Requests 2024.” *Open Baltimore*. Baltimore City. <https://data.baltimorecity.gov/datasets/68a1136acff444bba6c93e845dfc00e1_0/explore>.

Baltimore City Department of Planning. 2024. “Neighborhood Statistical Area (NSA) Boundaries.” *Open Baltimore*. Baltimore City. <https://data.baltimorecity.gov/datasets/68a1136acff444bba6c93e845dfc00e1_0/explore>.

ESRI. 2022. “ArcGIS Pro.” Esri Inc. <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>.

Gehlke, C. E., and Katherine Biehl. 1934. “Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material.” *Journal of the American Statistical Association* 29 (185): 169. <https://doi.org/10.2307/2277827>.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

U.S. Census Bureau. 2024. “tl\_2024\_24\_tract.” *TIGER/Line Shapefiles*. <https://www2.census.gov/geo/tiger/TIGER2024/TRACT/>.