Mapping Multidimensionality

Using Census Data to Model Neighborhood Communities

Nathan N. Alexander, PhD
Assistant Professor
Department of Curriculum and Instruction, School of Education
Program in Applied Data Science and Analytics, CADSA
Director, Quantitative Histories Workshop

Meet the talented individuals who make up our lab!

Nathan Alexander, PhD

PI, Howard University

Dr. Alexander is the founder of the Workshop and he has led the lab since its inception.

Helen Jang, PhD

PI, Morehouse College

Dr. Jang directs our sister lab, the HJLab, in the Atlanta University Center (AUC).

Lyrric Jackson

Spelman College

Prof. Jackson is our coordinator, in-house technical expert, and lab manager.

Jalil Coooper

Stanford University

Jalil is a senior design major at Stanford and he has been with the lab since its inception.

Myles Ndiritu

Morehouse College

Myles is a junior political science major at Morehouse College and a lab TA.

Kade Davis

Morehouse College

Kade is a senior sociology major at Morehouse College and a lab TA.

Kyshan Nichols-Smith

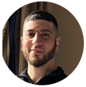
Morehouse College

Kyshan is a senior political science major at Morehouse College and a lab TA.

Qyana Stewart, MS

Howard University

Qyana is a PhD candidate in Higher Education Leadership and Policy Studies.



Basil Ghali

Howard University

Basil is a doctoral student in psychology and he has been with the lab since its inception.

Zoe Williams

Howard University

Zoe is a junior political science major at Howard University.

Jibek Gupta

Howard University

Jibek is a junior computer science major at Howard University.

Bayowa Onabajo

Howard University

Bayowa is a MS student in Applied Data Science and Analytics.

Amari Gray

Morehouse College

Jahmere Jackson

Howard University

Nicholas Angelo

Howard University

Lab Alum

Julian Amaya, Morehouse College

Jeremiah Lowther, Morehouse College

Maya Phillips, Spelman College

Sayid Achilov, Brown University

Kenthia Roberts, Columbia University

Gabriella Walker, Spelman College

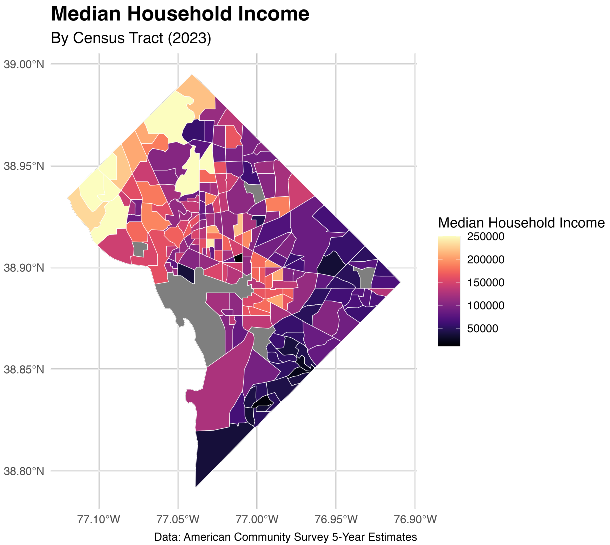
A History of Dividing Walls

There are dynamic and complex historical structures that have isolated and segregated communities based on a host of factors.

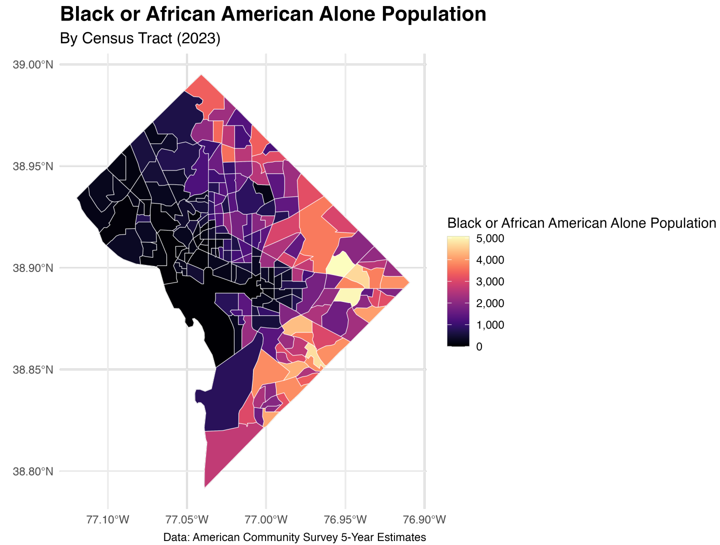
- Legacies of Jim Crow
- Redlining
- Gerrymandering
- Gentrification

Multiple year study on how modeling historical injustices better frames educational, economic, and political outcomes.

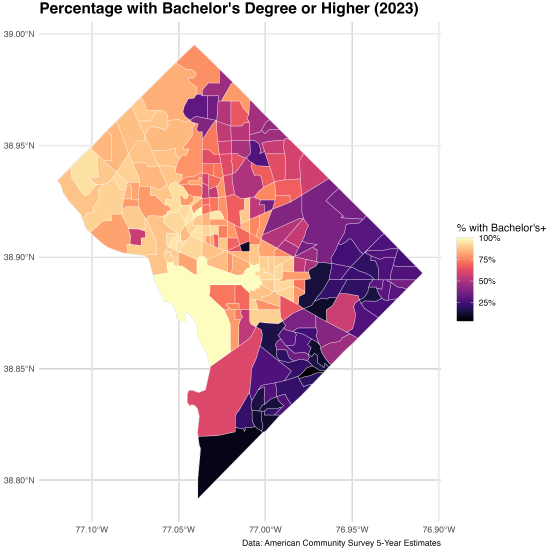
Single Dimension - Income



Single Dimension - Race



Single Dimension - Education



Considering Multiple Dimensions

How might we analyze the multidimensional patchwork of a city's fabric while retaining critical insights from theorists that challenge monolithic narratives of any single group?

Theoretical framework: A Patchwork Nation

"*If you pay attention to the complexity of the USA, its diversity and differences you soon realize that the ways we try to understand it – red and blue, Northeast and Midwest – are too simplistic. They are inadequate and misleading.*" -Patchwork Nation Project

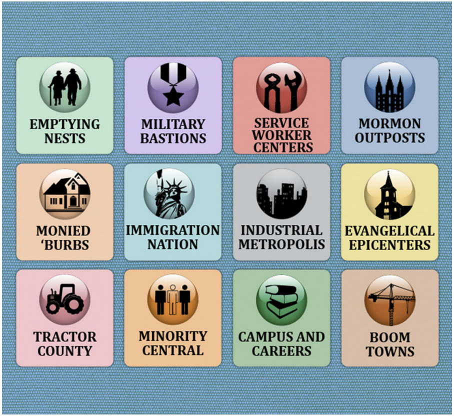


Figure 1: Community types in the Patchwork Nation

Analytic Framework: Theory of Dividing Walls (city-level)

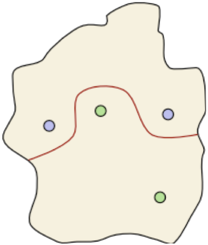
Consider a neighborhood as an "island" containing towns — the green tribe in two of the towns and the blue tribe in the other two.

Analytic framework

Using topology (mathematics) and topography (geology), we can formalize our analysis by attaching a few requirements:

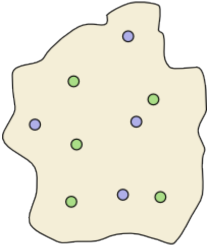
- The wall must be continuous, it must not intersect itself, it must not split, it must not pass through a town, and each end of the wall must be at the coast.
- Let us call such a wall a dividing wall.

Dividing Walls Theorem



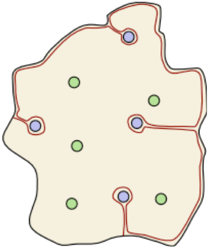
Theorem 1. Given any configuration of blue and green towns, there is a dividing wall that separates blue towns from green towns.

Is there a dividing wall?



Theorem 1. Given any configuration of blue and green towns, there is a dividing wall that separates blue towns from green towns.

A Dividing Wall



Theorem 1. Given any configuration of blue and green towns, there is a dividing wall that separates blue towns from green towns.

Measurement Model: Dissimilarity Index

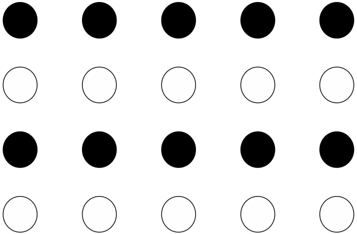
$$D = \frac{1}{2} \sum_{i=1}^{n} \left| \frac{b_i}{B} - \frac{w_i}{W} \right|$$

Where:

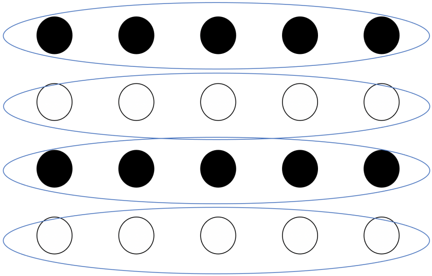
- b_i = Black population in tract i
- B = Total Black population in city
- w_i = non-Black population in tract i
- W = non-Black population in city

This census data model, however, only provides a single dimension of a neighborhood.

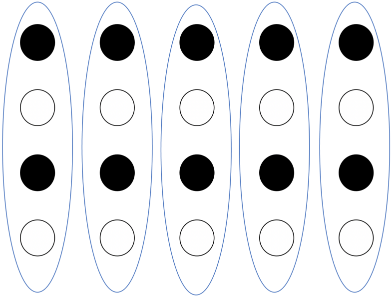
Modeling Antiblackness



Modeling Antiblackness



Modeling Antiblackness



Community Resilience Estimates

- The CRE are a measure of the capacity of individuals and households within a community to absorb, endure, and recover from external stresses.
- The CRE data combine American Community Survey (ACS) and the Population Estimates Program (PEP) data to identify social and economic vulnerabilities by geography.
- There is a nice CRE Interactive Tool that allows for a quick overview of local contexts.

Community Resilience Estimates

- Households with an income-to-poverty ratio less than 130%
- Less than one individual living in the household is aged 18–64
- Household crowding, defined as more than 0.75 persons per room
- Households with limited education
- No one in the household is employed full-time year-round
- Individual with a disability posing a constraint to significant life activity
- Individual with no health insurance
- Individual aged 65 or older
- Households without a vehicle
- Households without broadband internet access

```r
cre_correlates_dc <- get_acs(
  geography = "tract", state = "DC",
  year = 2023, survey = "acs5",
  variables = c(
    median_income = "B19013_001",
    poverty_rate = "B17001_002",
    unemployment_rate = "B23025_005",
    no_health_insurance = "B27010_033",
    educ_less_than_hs = "B15003_002",
    median_age = "B01002_001",
    housing_cost_burden = "B25070_010",
    no_vehicle = "B08201_002",
    black_population = "B02001_003",
    median_rent = "B25058_001"),
  summary_var = "B02001_001",
  output = "wide", geometry = FALSE)
```

Spatial Model

– Base spatial model formulation:

$$\mathbf{y} = \boldsymbol{X}\beta + \tau + \epsilon$$

- $\boldsymbol{y}$ is a n x 1 response vector
- $\boldsymbol{X}$ is a design matrix that contains explanatory variables
- β represents fixed effects coefficients
- τ denotes spatially dependent random errors
- ϵ represents independent random errors

Dimensionality in Spatial Models

Response vector structure ($\mathbf{y}$):

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- Each element, y_i, represents the observed response at a neighborhood's location i
- These are ordered by adjacency relationships to preserve the geographical context
- Review of distributions, spatial autocorrelation (i.e., $Cov(y_i, y_j)$), and decomposition

Dimensionality in Spatial Models

Design matrix of explanatory variables structure ($\boldsymbol{X}$):

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{1,1} & \dots & x_{1,p} \\ 1 & x_{2,1} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,p} \end{bmatrix}$$

- First column is the intercept term
- Subsequent columns represent p explanatory variables
- Each row corresponds to a specific neighborhood's covariates

Dimensionality in Spatial Models

Sample design matrix of explanatory variables

$$\boldsymbol{X} = \begin{bmatrix} 1 & 65,000 & 0.62 & 3,200 \\ 1 & 28,000 & 0.32 & 5,100 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 127,000 & 0.75 & 6,840 \end{bmatrix}$$

- Column 1 is the $E(\mathbf{y})$ when all other predictors are zero
- Variable 1 (col 2) as median income
- Variable 2 (col 3) as proportion of residents with HS diploma
- Variable 3 (col 4) as population density (residents/sq. mi)

Findings: Information and spatial segregation

There are multiple models to be considered:

- Spatial regression using intersectional interactions
- Structural Equation Modeling (SEM) with CRE components
- Multilevel Analysis of Individual Heterogeneity and Discriminatory Analysis (MAIHDA)
  - Evans et al. (2024). A Tutorial for Conducting MAIHDA. *Population Health*, Vol. 26, 101664
  - Combines intersectional stratification with neighborhood-level clustering
  - Models individuals nested within: Intersectional strata (e.g., low-income Black men), community typologies from framework (e.g., Patchwork Nation) classifications

Special Thanks

Research assistants: Bayowa Onabajo (Howard University), Jibek Gupta (Howard University), Myles Ndiritu (Morehouse College), Zoe Williams (Howard University)

Lab manager: Lyrric Jackson (Spelman College)

Funding: Alfred P. Sloan Foundation, AUC Data Science Initiative, Data.org

Partners: The Carpentries

References

Chinni, D., & Gimpel, J. (2010). Our Patchwork Nation: The Surprising Truth about the “Real” America. Gotham Books.

Evans, C. R., Leckie, G., Subramanian, S. V., Bell, A., & Merlo, J. (2024). A tutorial for conducting intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). SSM - Population Health, 26, Article 101664. https://doi.org/10.1016/j.ssmph.2024.101664.

U.S. Census Bureau. (2024). Community Resilience Estimates. Retrieved March 26, 2025, from https://www.census.gov/programs-surveys/community-resilience-estimates/about.html.