

Calculating Residential Segregation Indices in A Reproducible Pipeline

Boyi Guo

Department of Biostatistics
University of Alabama at Birmingham

Last Updated 2022/02/17

Outline

Outline

- ▶ Motivation
- ▶ Pipeline Demonstration
 - ▶ Brief introduction to GitHub, Rstudio,
 - ▶ Brief introduction to reproducible pipeline targets, renv
 - ▶ How to run the pipeline
 - ▶ How to fetch data
- ▶ Some Preliminary Findings
 - ▶ Deciding
 - ▶ Recap of segregation measures

Motivation

Motivation

- ▶ Since the publication of xxx manuscript, the inquiries about social segregation indices for REGARDS are growing.
- ▶ These request contains calculating different geographic units, different years
- ▶ Why a reproducible pipeline, instead of a packaged function/
 - ▶ I think it is a great time to have more systematic knowledge about reproducible research
 - ▶ Why reproducible research. we defer to Peng. Hicks
 - ▶ How to start with reproducible research

Pipeline Demonstration

Preparation

- ▶ Software & Package Installation
 - ▶ Git & GitHub <https://git-scm.com/book/en/v2/Getting-Started-Installing-Git>
 - ▶ R & RStudio
<https://www.rstudio.com/products/rstudio/download/>
 - ▶ R package `renv`: <https://rstudio.github.io/renv/index.html>

Preliminary Findings & Remarks

Residential Segregation Indices Recap

According to Massey and Denton (1988), there are five major dimensions of segregation

- ▶ *Evenness*: spatial distribution of different groups among *units* in a metropolitan area
- ▶ *Exposure*: possibility of interaction between minority and majority group members
- ▶ *Concentration*: relative amount of physical space occupied by a minority group in the metropolitan area
- ▶ *Centralization*: how a group spatially located near the center of an urban area
- ▶ *Clustering*: which areal *units* inhabited by minority members adjoin one another, or cluster, in space

Total of 19 indexes. By default, the majority refers to non-Hispanic Whites.

Focus

- ▶ Dissimilarity index for Evenness: the percentage of population would have change residence to have the same percentage overall
 - ▶ 0.0 (complete integration) to 1.0 (complete segregation)
- ▶ Interaction index for Exposure: probability that a minority person shares a unit area with a majority person
 - ▶ 0.0 (complete segregation) to 1.0 (complete integration)
- ▶ Isolation index for Exposure: probability that a minority person shares a unit area with a minority person
 - ▶ 0.0 (complete integration) to 1.0 (Complete segregation)

Remarks

- ▶ Define the geographic units: in our calculation, we aggregated tract level statistics to reflect county level information
- ▶ When there are only two groups, interaction index and isolation index sum up to 1
- ▶ “Indexes of evenness and exposure are correlated but measure different things: exposure measures depend on the relative sizes of the two groups being compared, while evenness measures do not.”
- ▶ Isolation index only require one group, a minority group

References

References I

Massey, Douglas S, and Nancy A Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67 (2): 281–315.