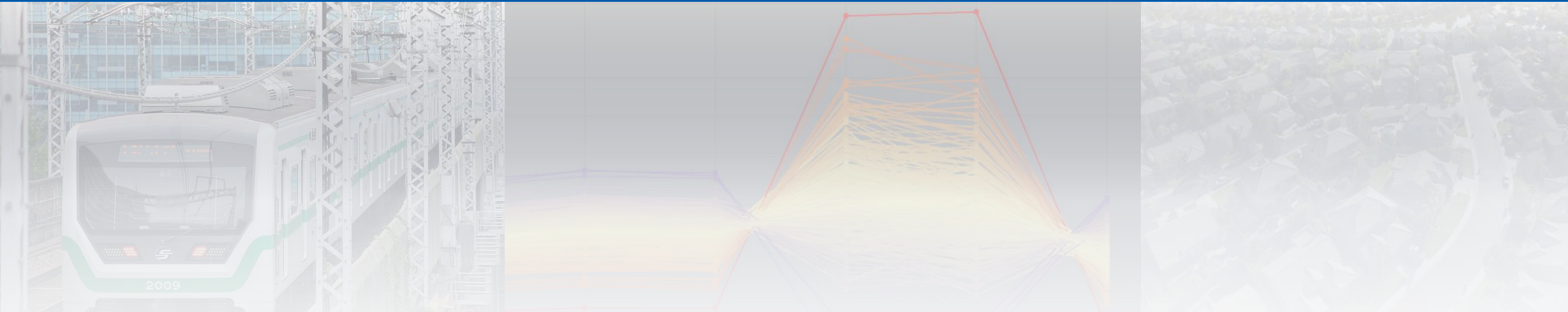


Examining Longitudinal Neighborhood Change and Human Mobility Patterns with Functional Data Analysis Approach

UK-South Korea Geospatial Data Science Seminar
UBAI, University of Seoul / June 9th 2023

Paul Hayoung Jung
Inha University & UC Riverside
Paul.Jung@inha.ac.kr

Bio: Paul H. Jung

- (2022–) Assistant Professor, Asia Pacific School of Logistics, Inha University
- (2023–) Assistant Professional Researcher, UC Riverside School of Public Policy
- (2021) Ph.D. in Geography and Urban Regional Analysis, UNC Charlotte

- Research Interests: Economic / Transportation Geography
  - International Trade, Freight Transportation, Transportation System
  - Neighborhood Dynamics, Population Degrowth
  - Spatial Flow Data Analysis, Spatiotemporal Modeling

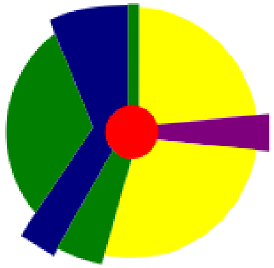
- (2023–) Vice Chair, AAG Transportation Geography Specialty Group
- (2022–) Advisory Panel Member, Republic of Korea Urban Transportation Policy Working Committee
- (2020–2022) Geographer, U.S. Census Bureau Geographer
  - U.S. Metropolitan area delineation, World population mapping, Population risk analysis

Exploring Longitudinal Neighborhood Change with the Functional Data Analysis Approach

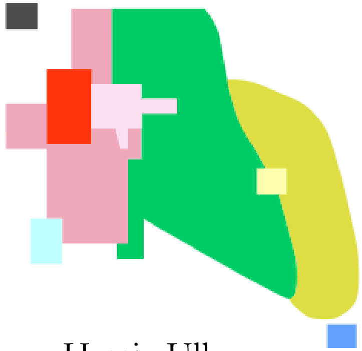
Introduction: Neighborhood Dynamics

- An old, classic but still new question
  - **How are urban neighborhoods spatially organized?**
  - Theoretical models: Concentric Model, Sector Model, Multiple Nucelli model
  - Effective cross-sectional framework

Burgess Concentric model

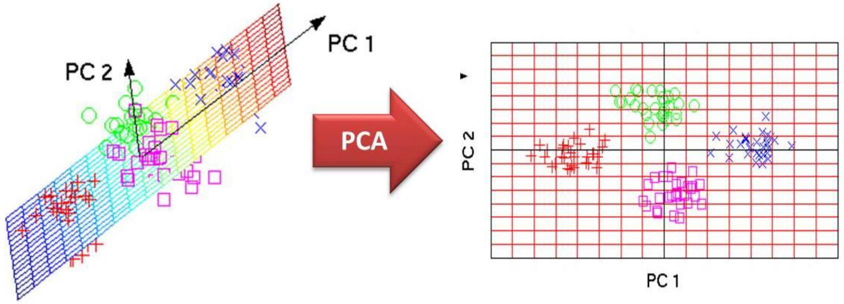


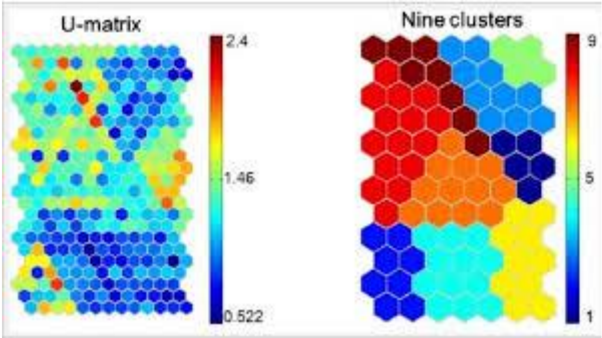
Hoyt Sector model



Harris-Ullman Multiple Nuclei model

- Spatial analysis models for empirical studies
  - Mainly using multidimensional census data (census tracts)
  - Dimension reduction: PCA, SOM
  - Cluster analysis: k-means, hierarchical clustering
  - Neighborhood classification
  - Very effective in addressing racial and income segregation

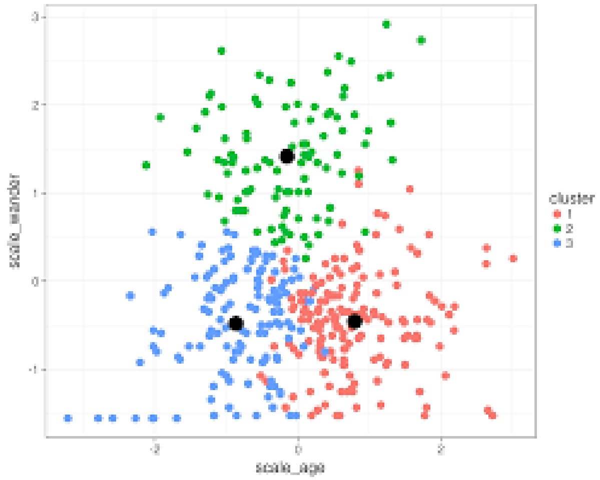




Introduction: Neighborhood Dynamics

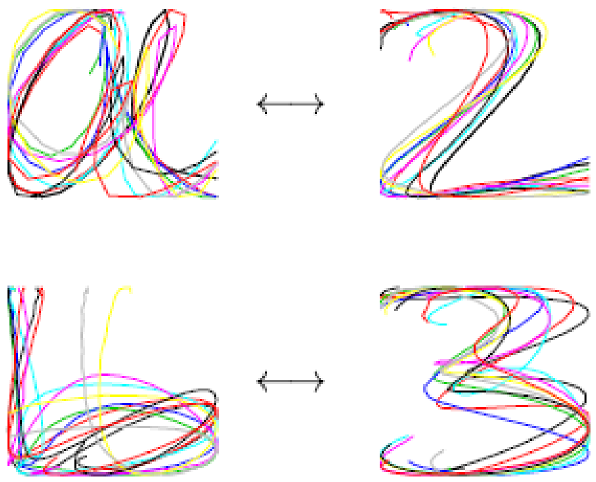
- **How are urban neighborhoods spatially organized over time?**
  - Focus: Cross-sectional => Temporal dimension
  - Many neighborhood phenomena are developed across time (e.g., White flights in US cities, income/racial segregation, gentrification, transit development)
  - (Northern UK cities) Urban deprivation
  - (South Korea) Small and medium cities experiencing challenges with depopulation and population aging.
  - Adding time dimension increases complexity in data analysis

- Existing neighborhood trajectory models: Discrete Change
  - **Markov chain analysis** (Rey et al., 2018)
    Neighborhood sequence analysis (Delmelle, 2015, 2016)
  - Neighborhood changes are represented by the stepwise, discrete, categorical changes in these two models



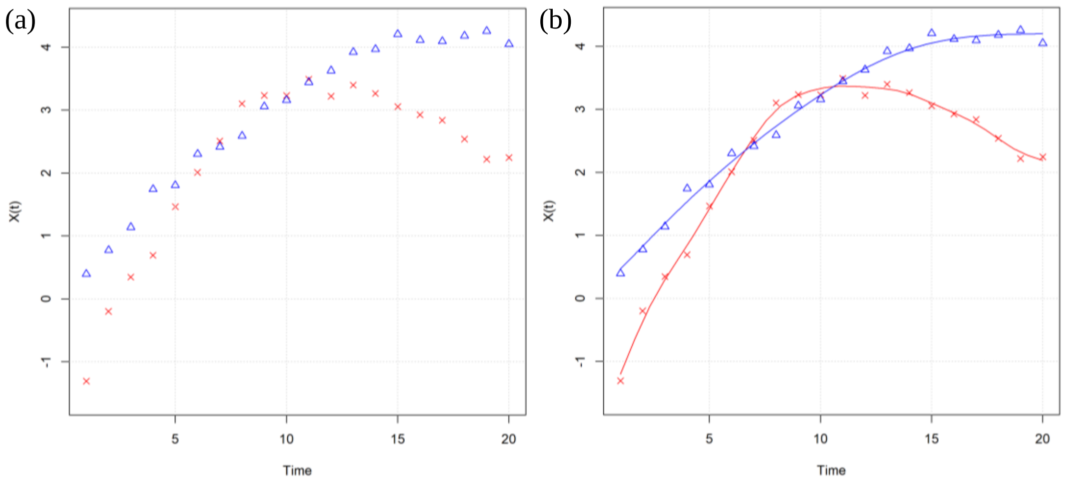
Neighborhood Changes as Continuous Process

- Neighborhood change as a continuous process
  - Describing neighborhood changes as time-dependent, continuous growth process would be more natural



- **FDA can incorporate the time dimension**
  - Variables are treated as functions of continuum (time, space)
  - Time is embedded in variables
  - A time-dependent curve becomes a unit of analysis

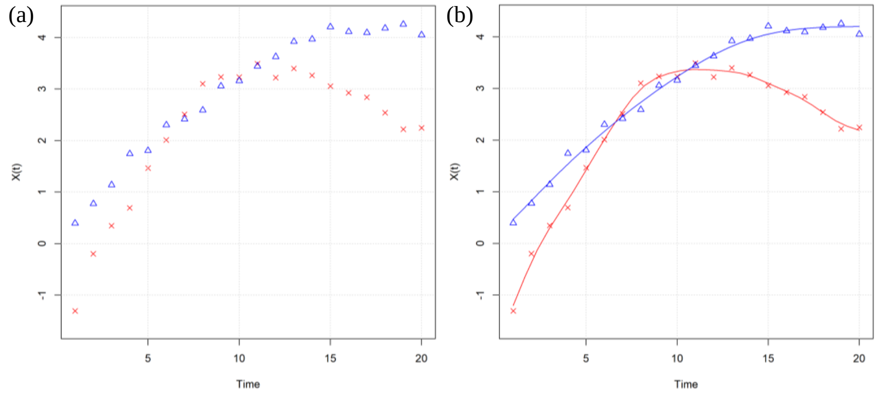
- Each neighborhood has a signature of the trajectory of changes
  - FDA can quantify and compare the shape of neighborhood change curves



Functional Data Analysis: Multivariate Functional PCA

$$\boldsymbol{X_i} = \begin{bmatrix} x_i^1 & y_i^1 & z_i^1 & w_i^1 & \cdots \\ x_i^2 & y_i^2 & z_i^2 & w_i^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ x_1^\tau & y_i^\tau & z_i^\tau & w_i^\tau & \cdots \end{bmatrix}$$

Neighborhood panel data



Curve-fitting (B-spline)

$$\boldsymbol{X_i}(t) = [x_i(t) \quad y_i(t) \quad z_i(t) \quad w_i(t) \quad \cdots]$$

Neighborhood functional data

Multivariate functional PCA

Functional decomposition: All functions are infinite linear combination of basis functions

$$\boldsymbol{X_i}(t) \approx \sum_m^\infty \rho_m^i \boldsymbol{\psi}_m(t)$$

Using first few principal components, each tract's curve is simply represented as (functional) coordinates!

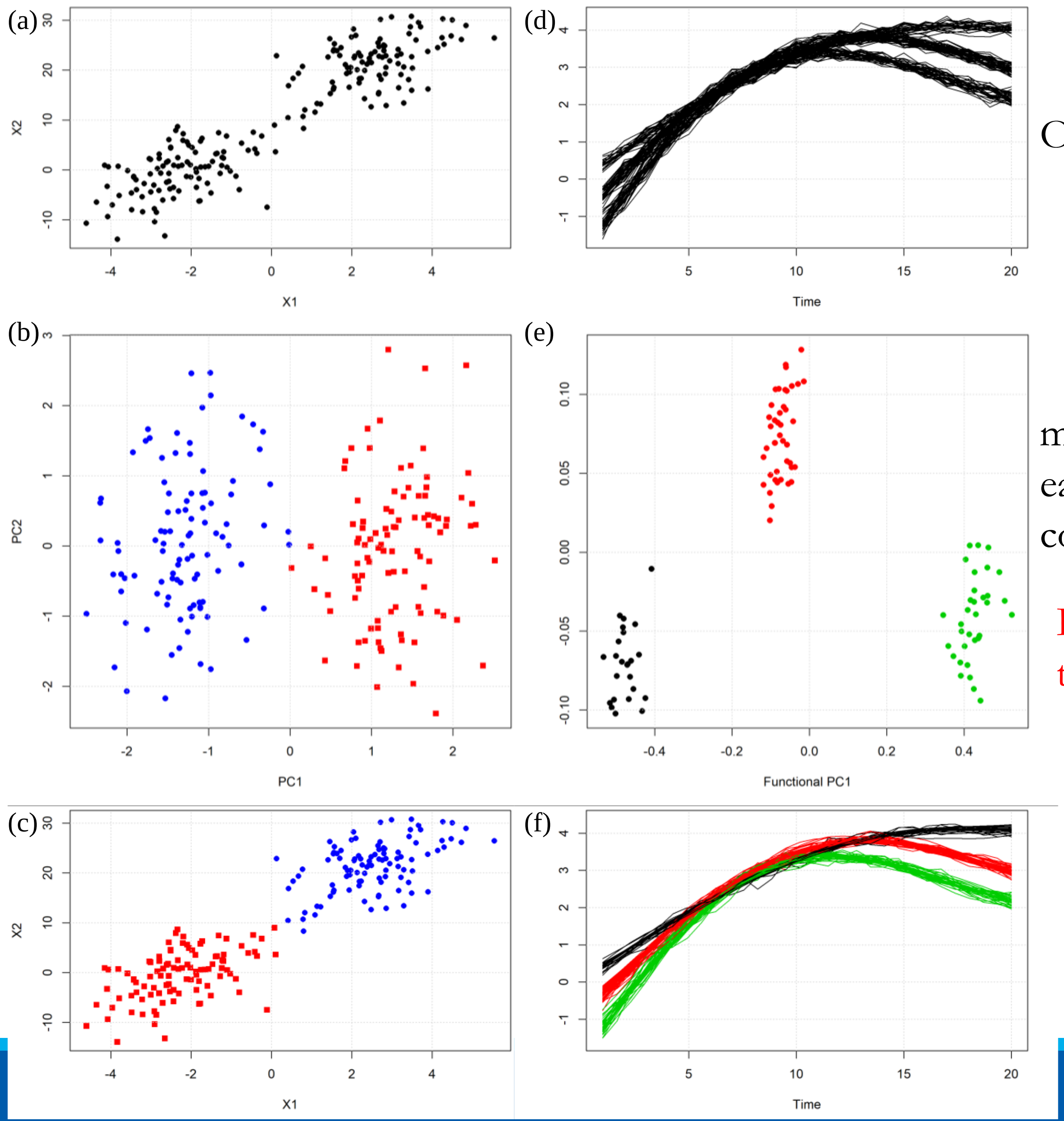
$$(\rho_1^i, \rho_2^i, \rho_3^i, \rho_4^i)$$

Classical PCA vs Functional PCA

Original data

PCA and
each point represented as
coordinates of 2 PCs

PCA+
k-means Clustering



Original data

mfPCA and
each curve represented as
coordinates of 2 mfPCs

Each point corresponds
to each curve!

mfPCA +
Clustering

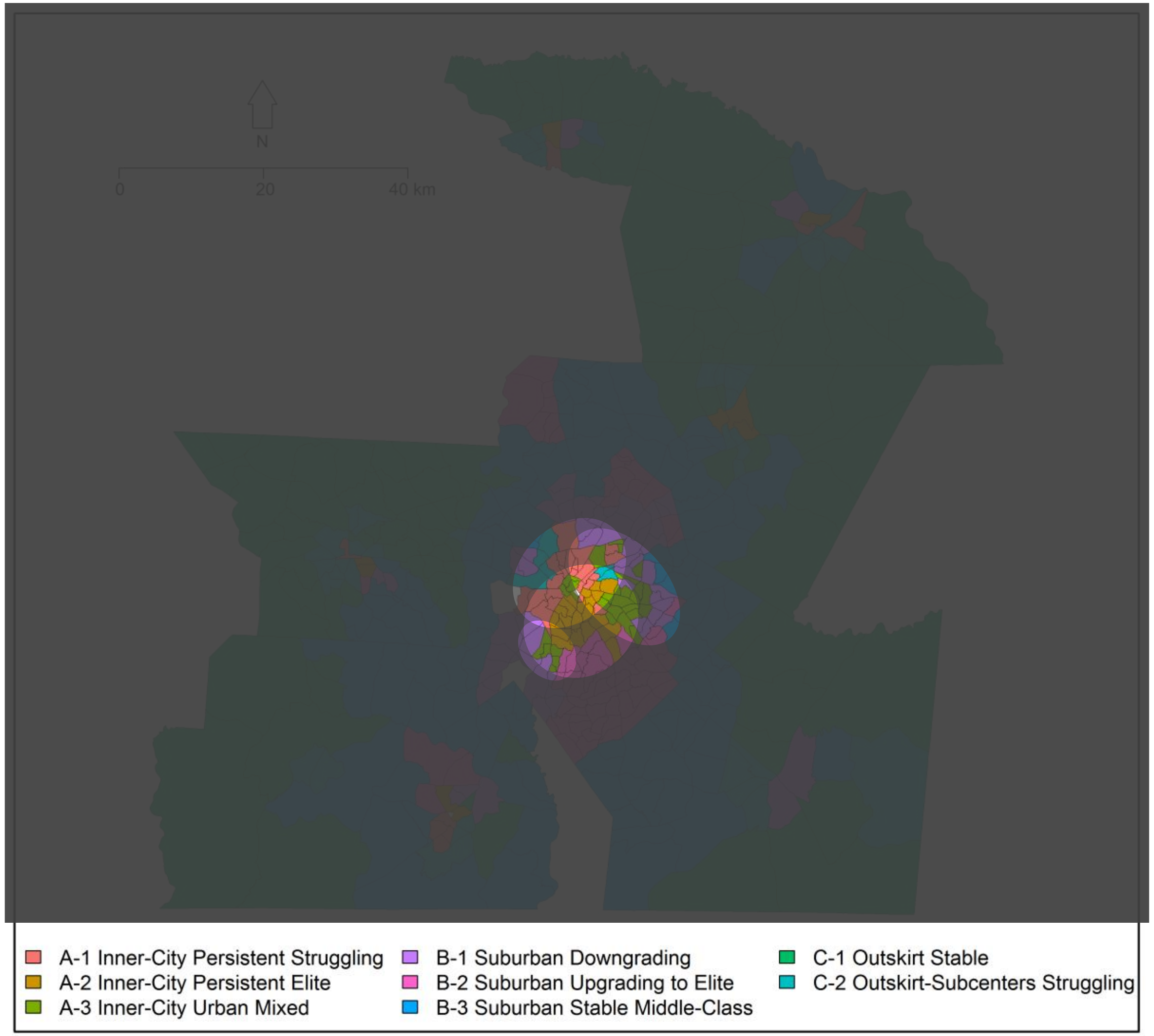
Applications: Neighborhood Trajectory Analysis

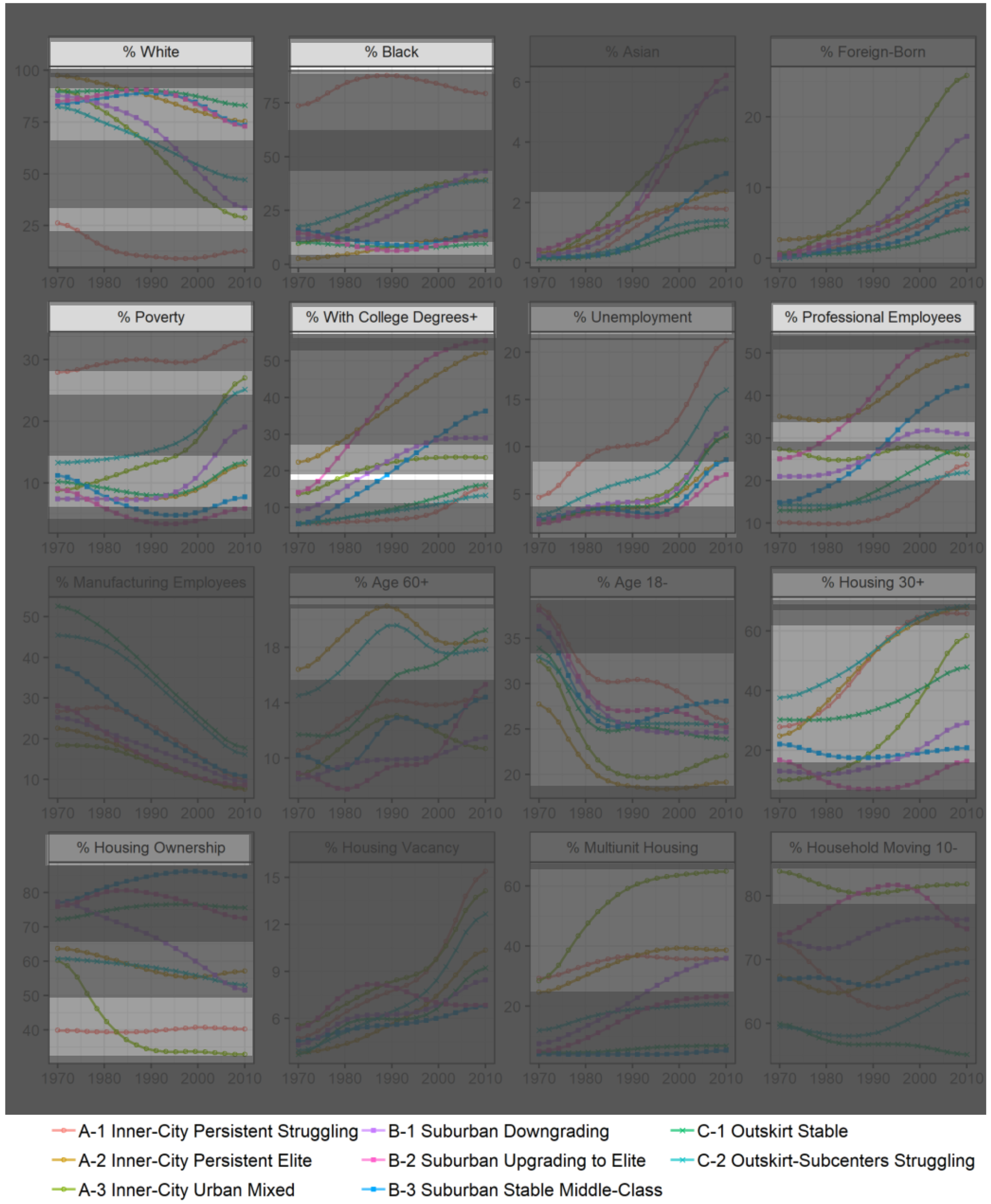
- Charlotte, Detroit and Dallas-Fort Worth Metropolitan Areas
- Data: Longitudinal Tract Database (LTDB) 1970-2010
- 16 census variables: Based on Shevky-Bell hypothesis

Category	Percent Variable	Note
Racial and Ethnic Segregation	Black residents	
	White residents	
	Asian residents	
	Foreign-born residents	
Socioeconomic Status	Unemployed	
	Below poverty level	
	Persons with at least 4-year degree	
	Manufacturing employees	
	Service industry employees (Professional)	
Lifecycle / Lifestyle	Persons aged 60 years an above	Housing
	Persons aged 18 and under	Housing
	Owner-occupied housing	Housing
	Vacant housing	Housing
	Multiunit structures	Housing
	Structures built more than 30 years ago	Age
	Household heads move into a unit less than 10 years ago	Age

Applications: Neighborhood Trajectory Analysis

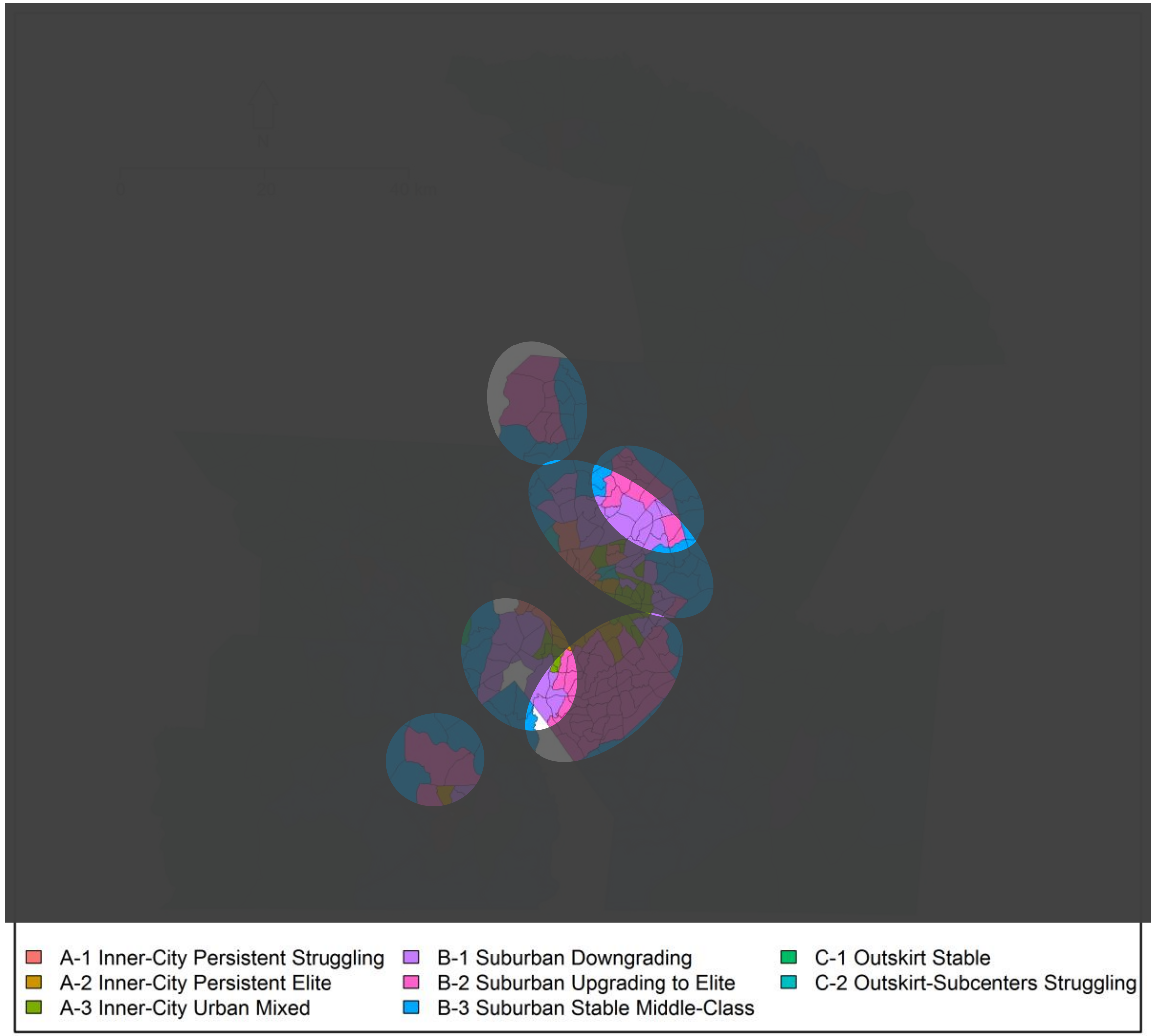
- Charlotte: 8 clusters (5 fPCs)

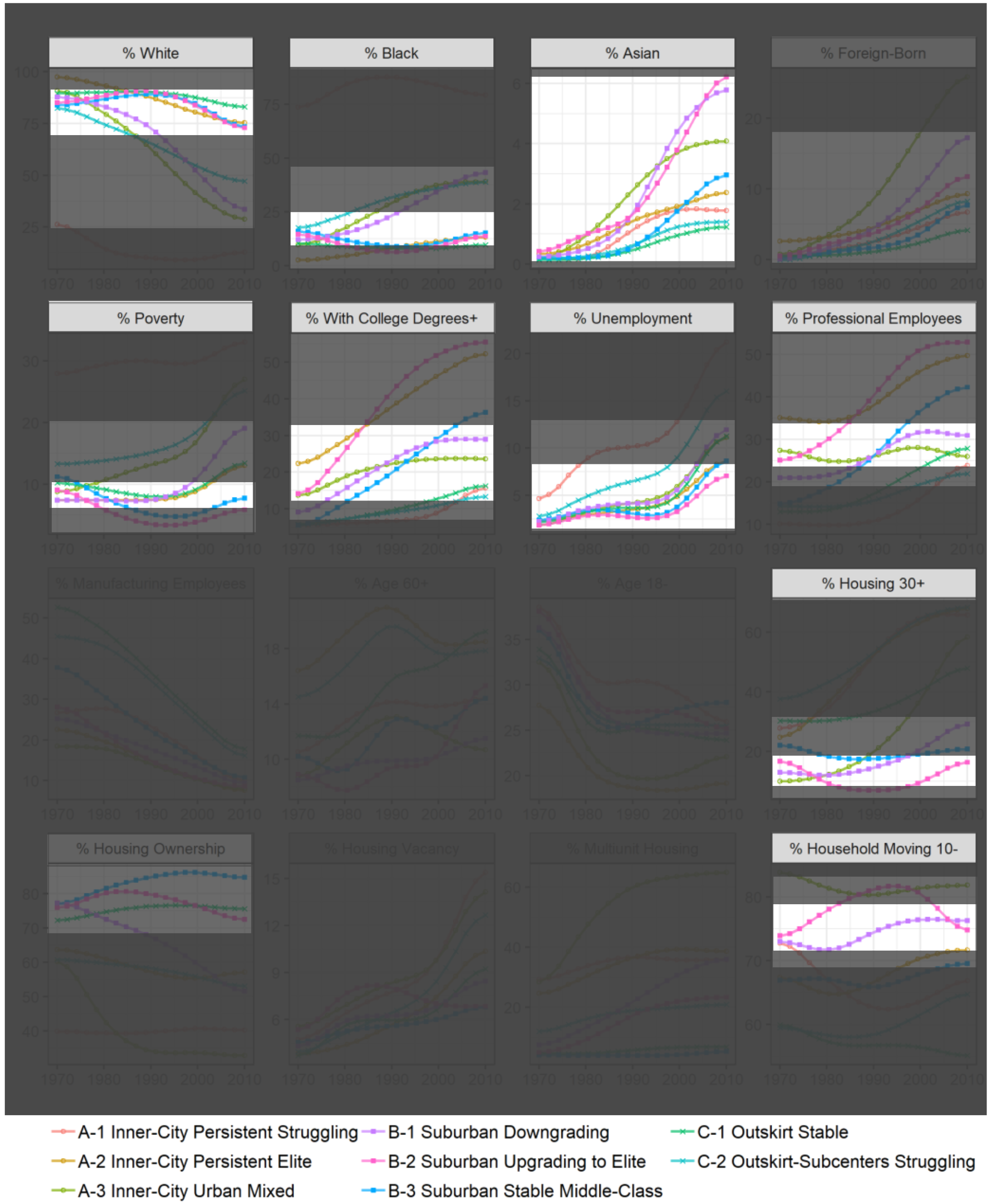




Applications: Neighborhood Trajectory Analysis

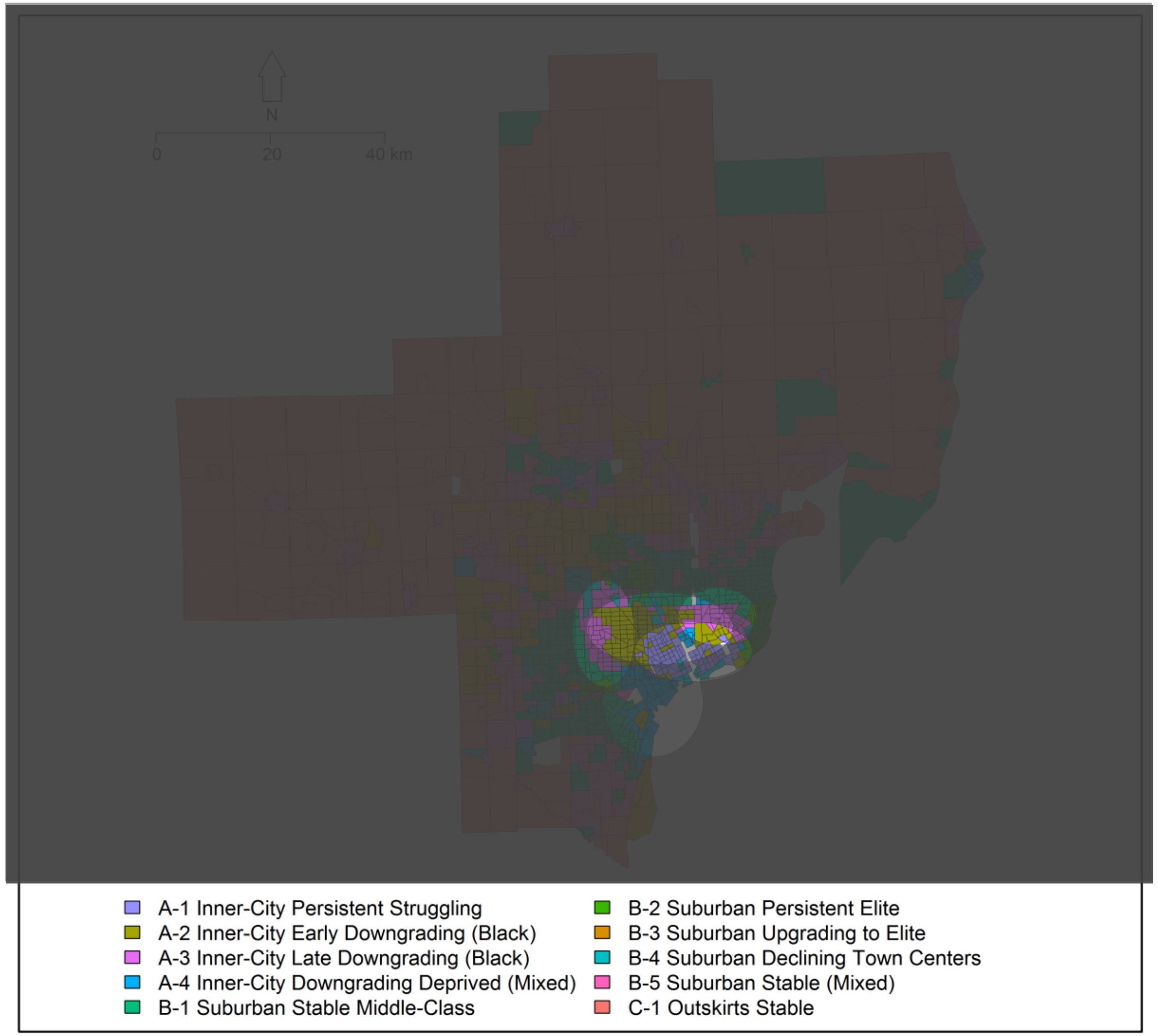
- Charlotte: 8 clusters (5 fPCs)

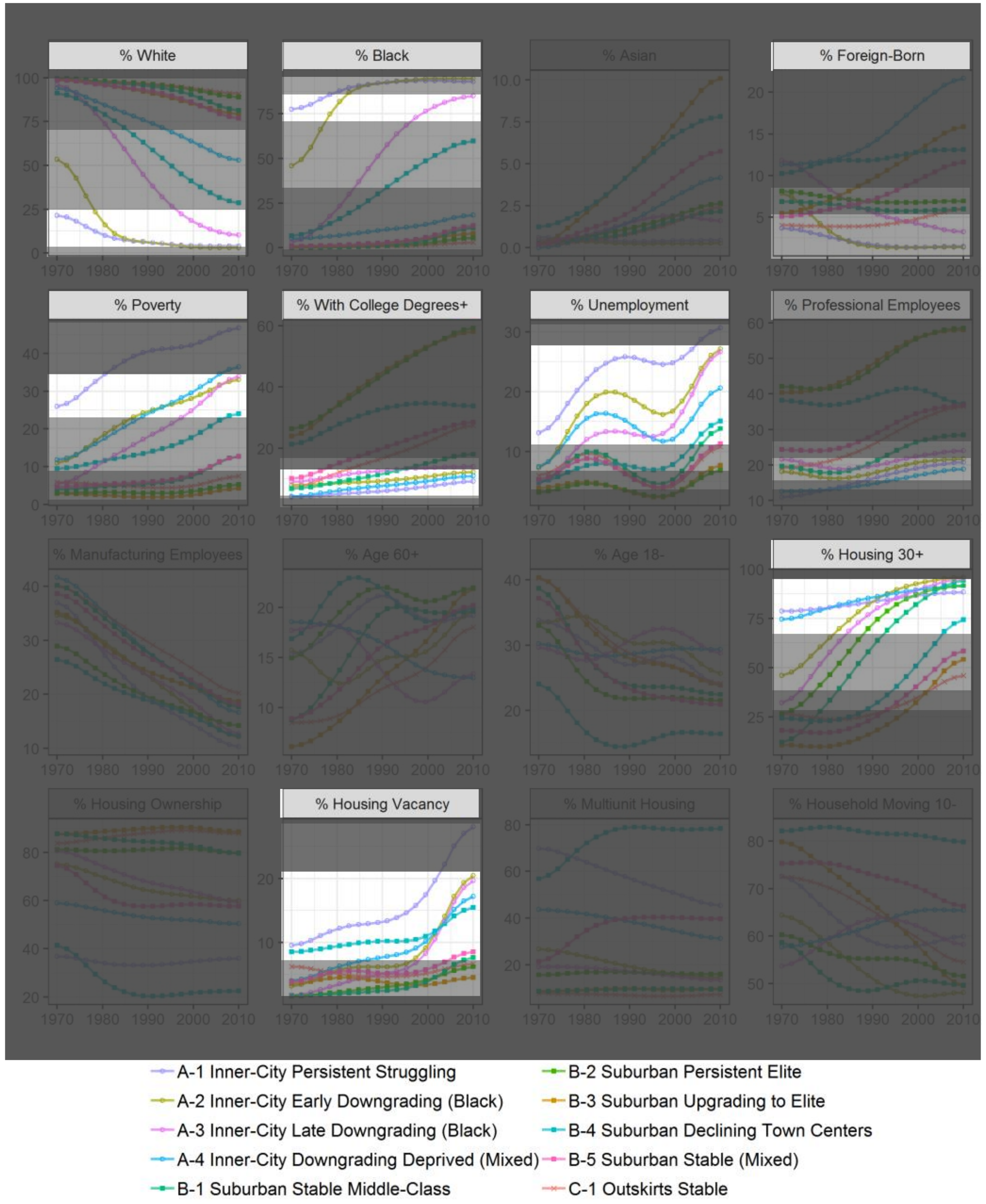




Applications: Neighborhood Trajectory Analysis

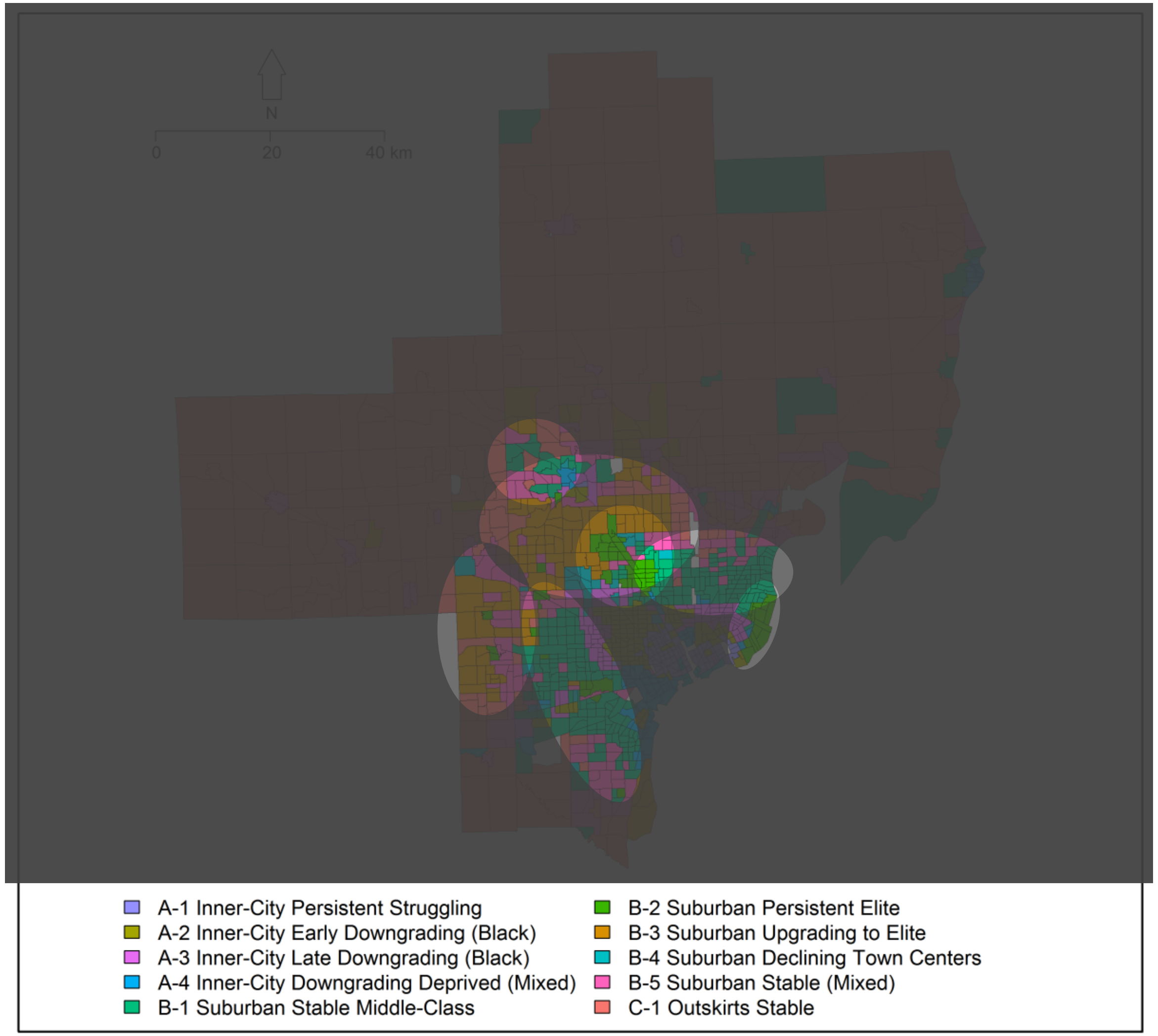
- Detroit: 10 clusters (6 fPCs)

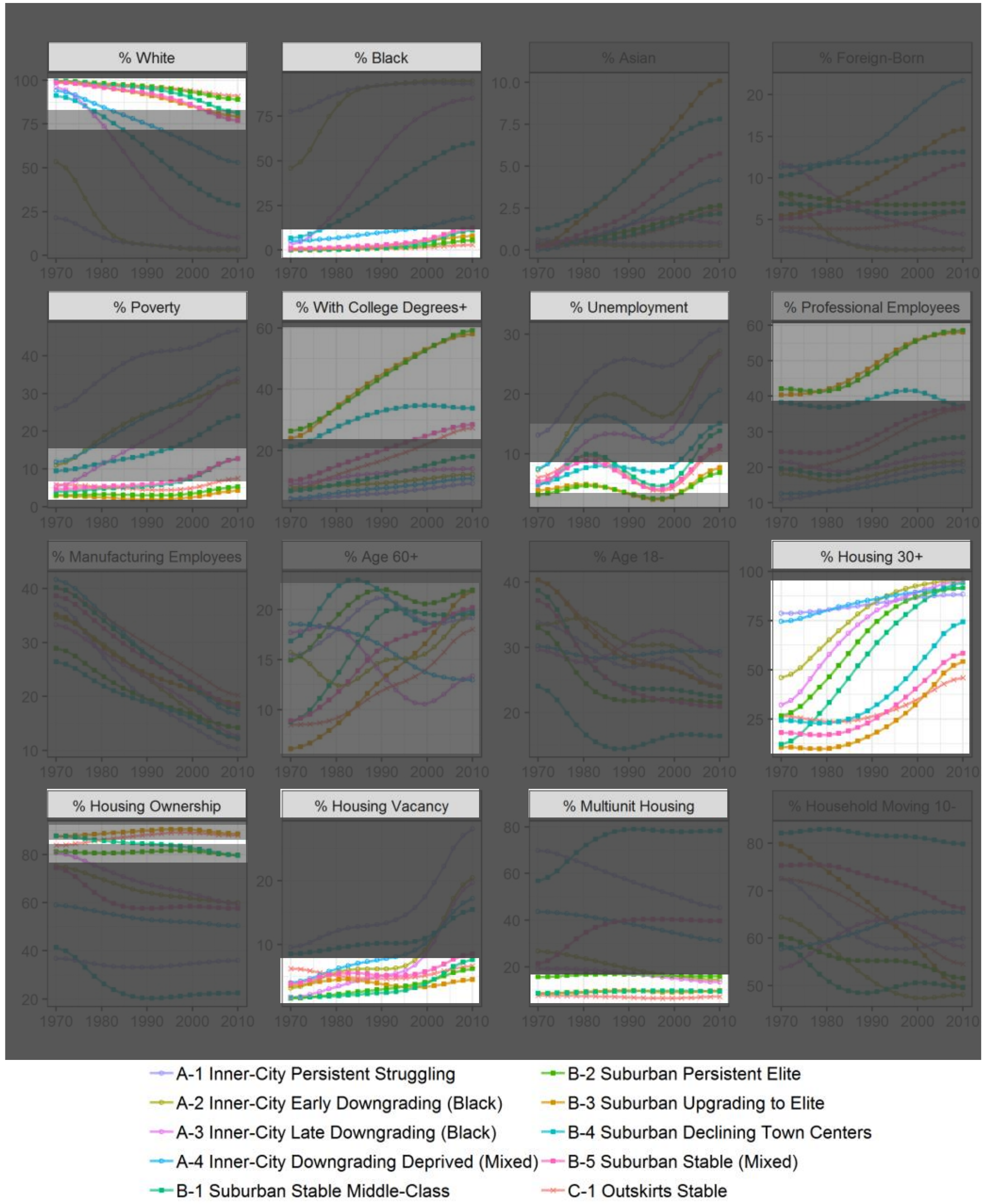




Applications: Neighborhood Trajectory Analysis

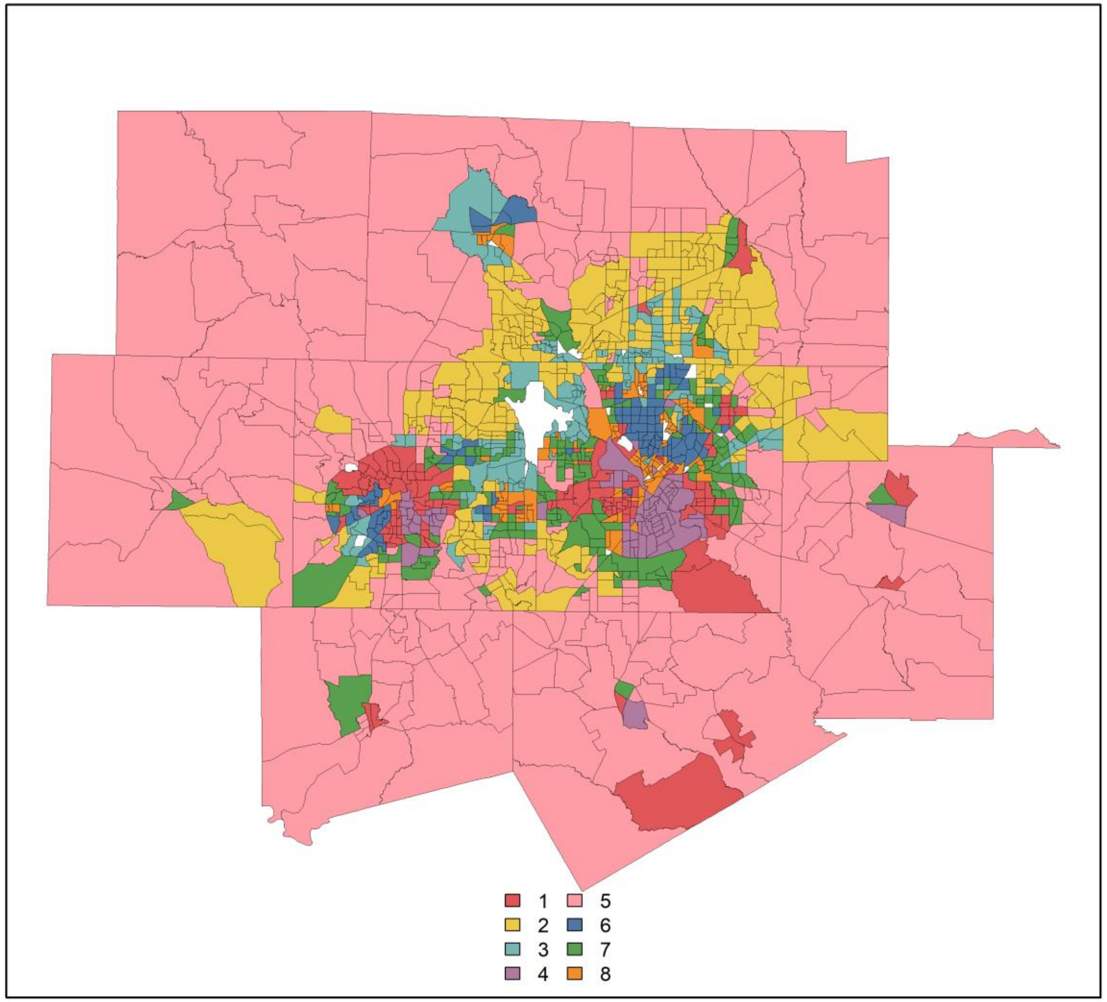
- Detroit: 10 clusters (6 fPCs)

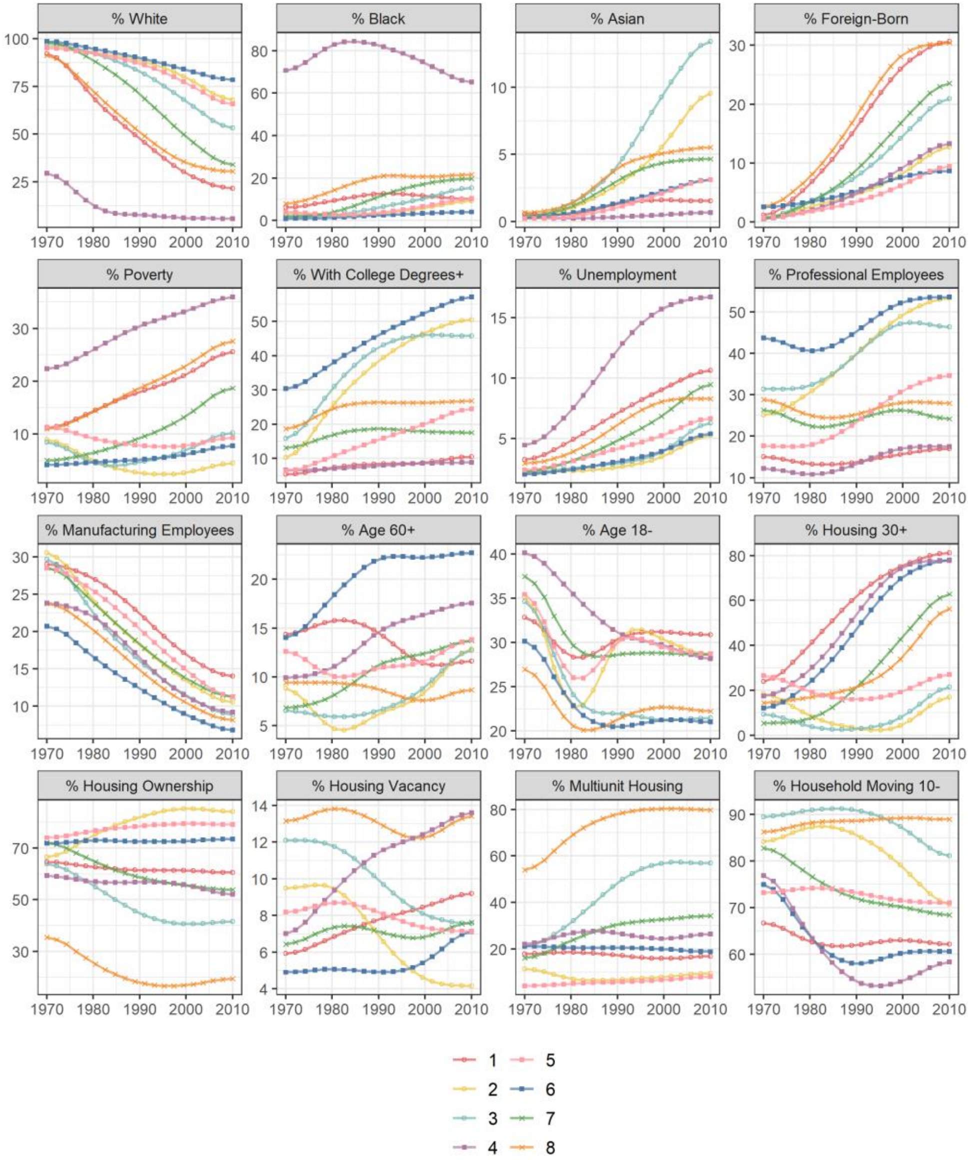




Applications: Neighborhood Trajectory Analysis

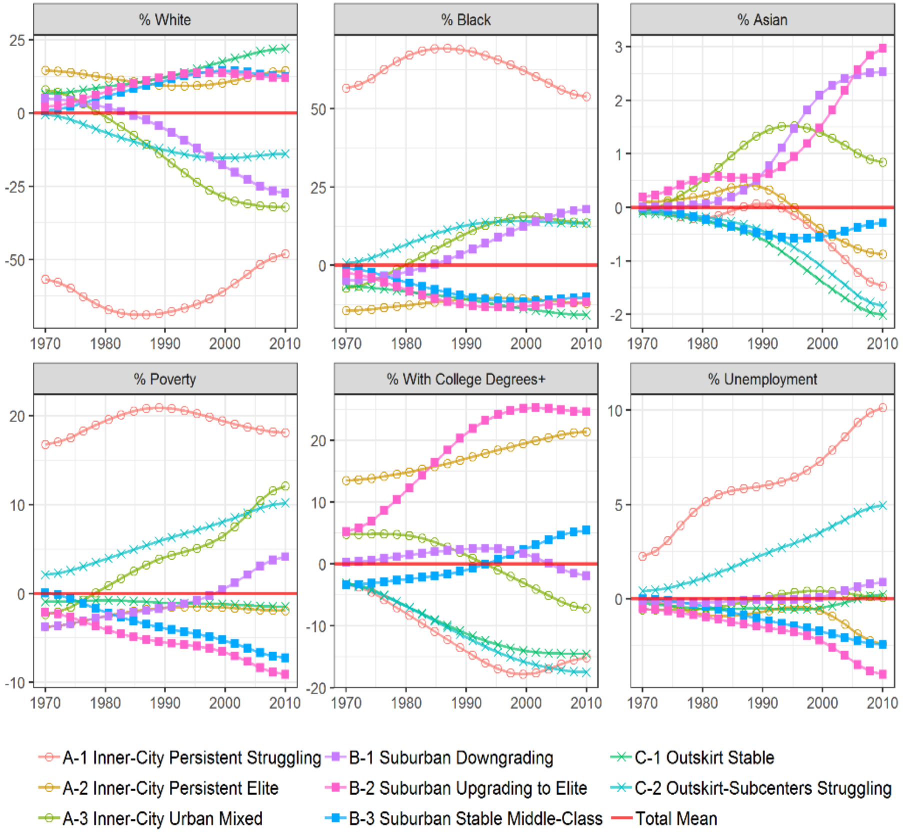
- Dallas-Fort Worth: 8 clusters

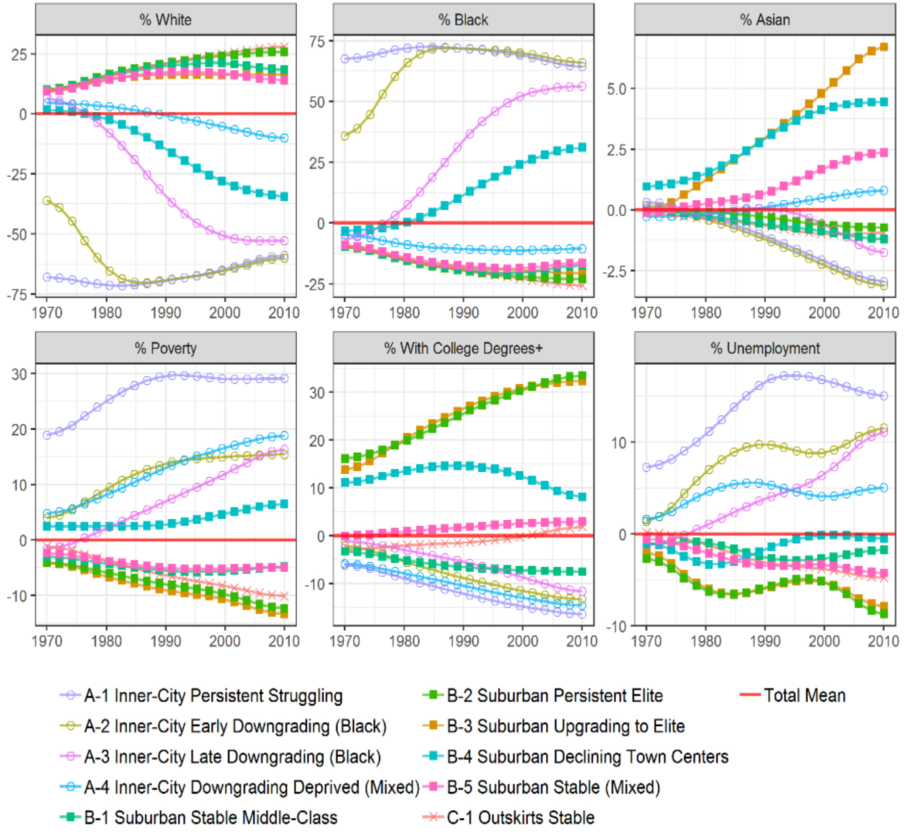




Applications: Neighborhood Trajectory Analysis

- General findings
  - Increasing racial segregation
  - Increasing socioeconomic segregation between inner-city and suburbs and between suburbs
  - Accelerated deindustrialization





Day-to-day Dynamics of Transportation Systems and Dynamic Time Accessibility

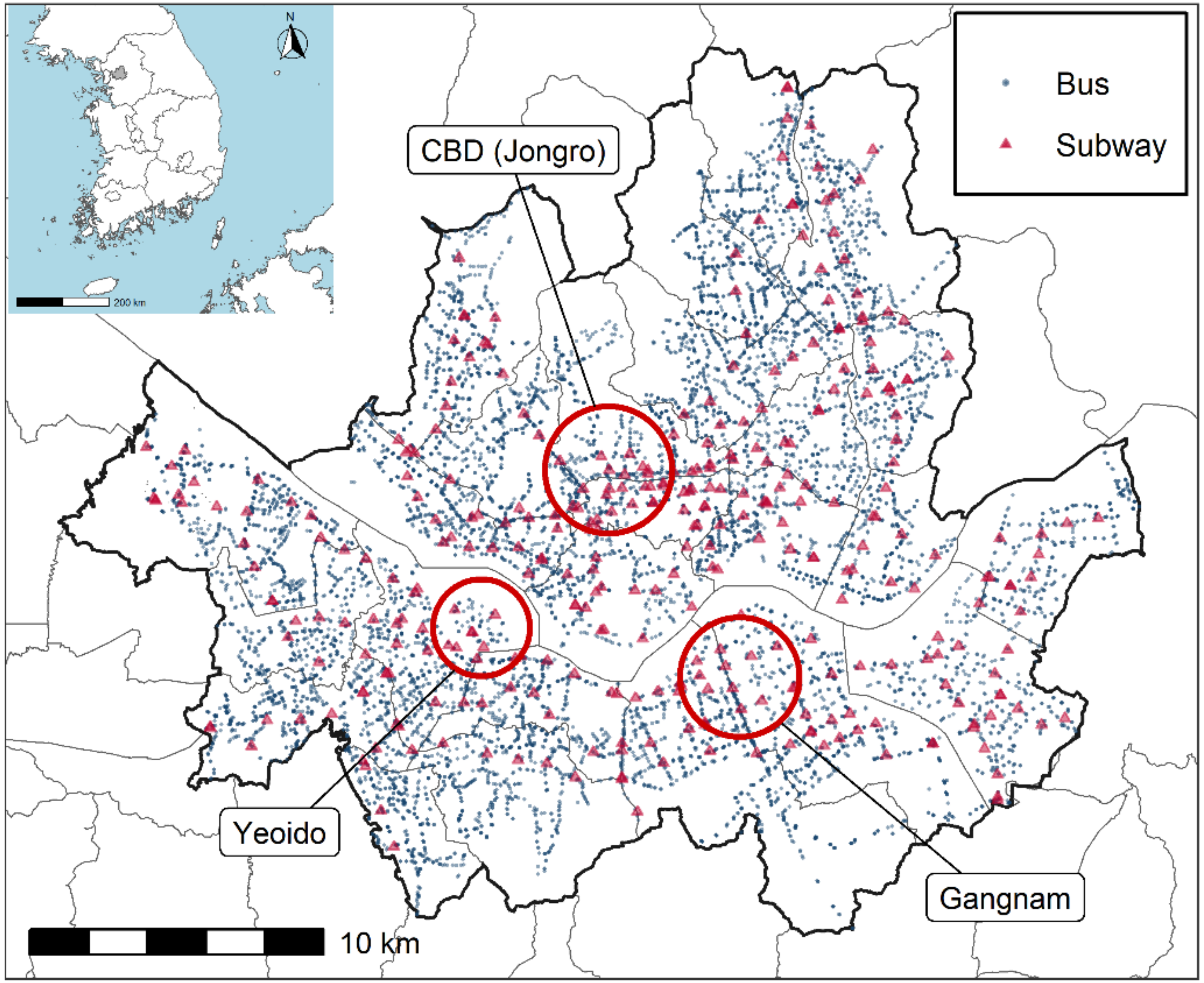
Introduction: Dynamic Accessibility

- Accessibility
  - Accessibility has been a major issue in urban policymaking
  - Measuring accessibility indicates spatial disparity in transit service in urban areas
  - Accessibility has been measured in a static way by using cross-sectional data which were collected in a certain time point (snapshot)

- Time Dynamics of Transportation Systems: Why is time neglected in accessibility measurement?
  - Urban transportation systems follow the daily and weekly rhythm of urban activities
    Traffic volume, road congestion, and transit-time schedules
  - Cross-sectional measurement of accessibility level had been dominant because of the limitations in data collection methods
  - Availability of real-time transit data with higher temporal and spatial granularity
  - More attention to "Dynamic Time Accessibility"

Background and Data: Seoul Integrated Transit System

- **How do neighborhoods vary by accessibility?**
  Day-to-day fluctuation in travel patterns
  Across transit modes
- Data: Seoul Smart Card Data (2015)
  - 5/17 – 5/23 2015 (A week data)
  - 149,330,464 trips
  - 13,518 bus stops, 364 subway stations
- Seoul, South Korea
  - 10M in city and 26M pop in the entire metro area
  - 3 CBD: Jongro, Yeoido, Gangnam
- Seoul Integrated Transit System
  - 65% of passengers served by the transit system
  - Last-mile accessibility not covered by the subway
  - Free-of-charge intermodal transfer
    Within the same mode
    Between bus and subway



Application 2: Transportation Studies

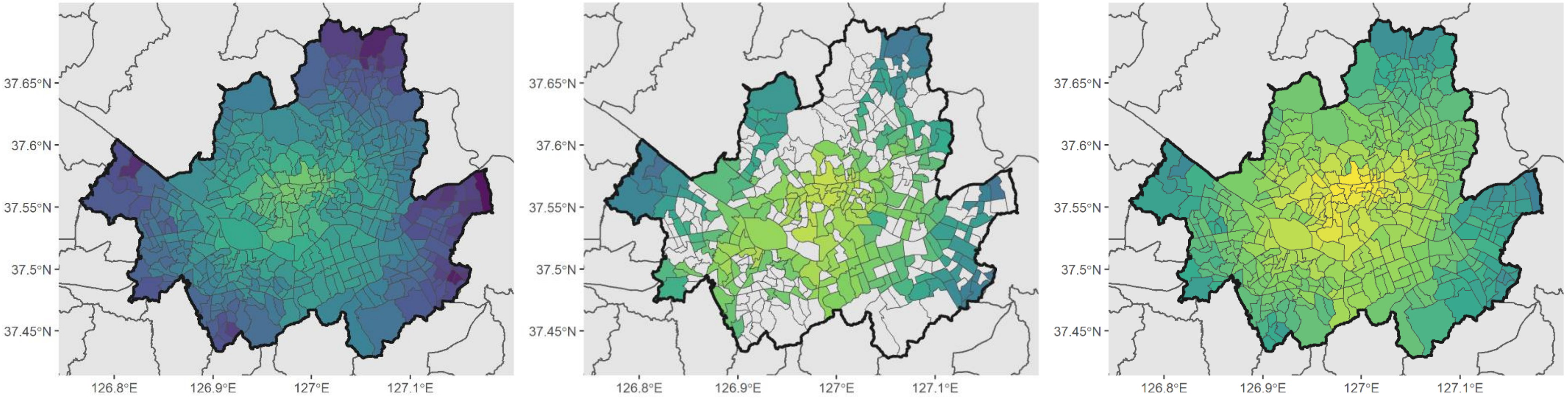
- Static Accessibility

$$a_{i,t}^{m} = \frac{\sum_{i \neq j}^{j} a_{ij,t}^{m}}{|J_m|}$$

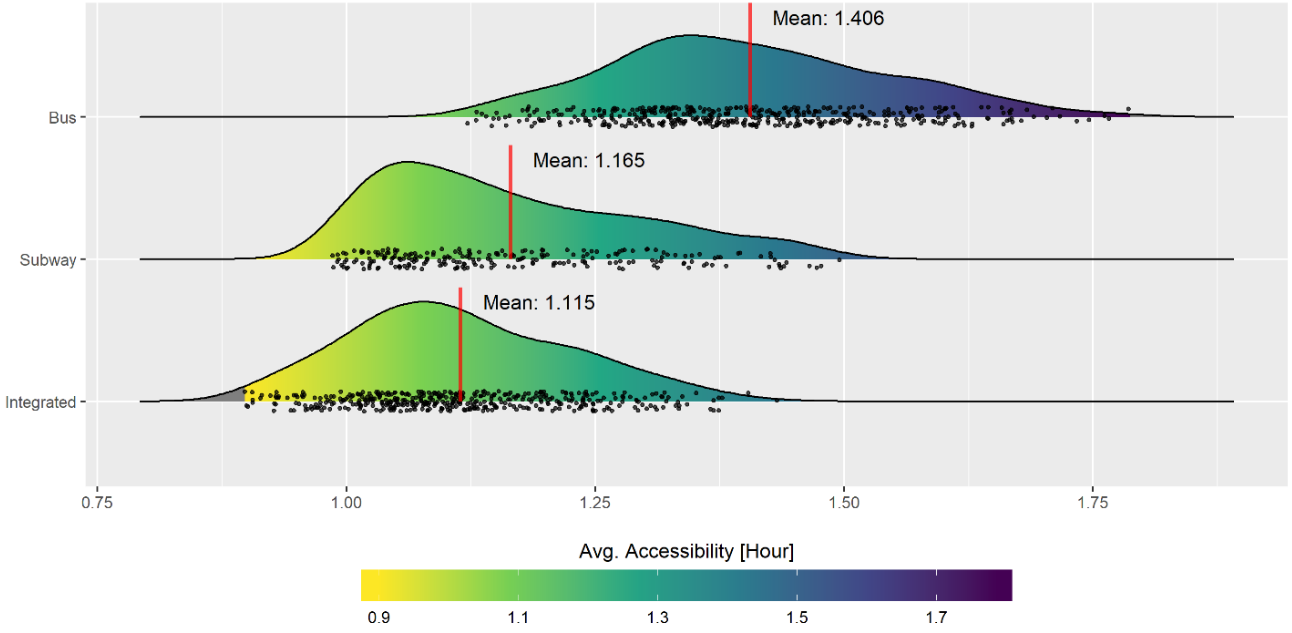
Node-level average travel time to *i*

$$A_{n,t}^{m} = \frac{\sum_{i \in N_m} a_{i,t}^{m}}{|N_m|}$$

Tract-level average travel time to *i*



Bus Subway Integrated



Day-to-Day Dynamic Time Accessibility: Seoul Integrated Transit System

- Dynamic Time Accessibility

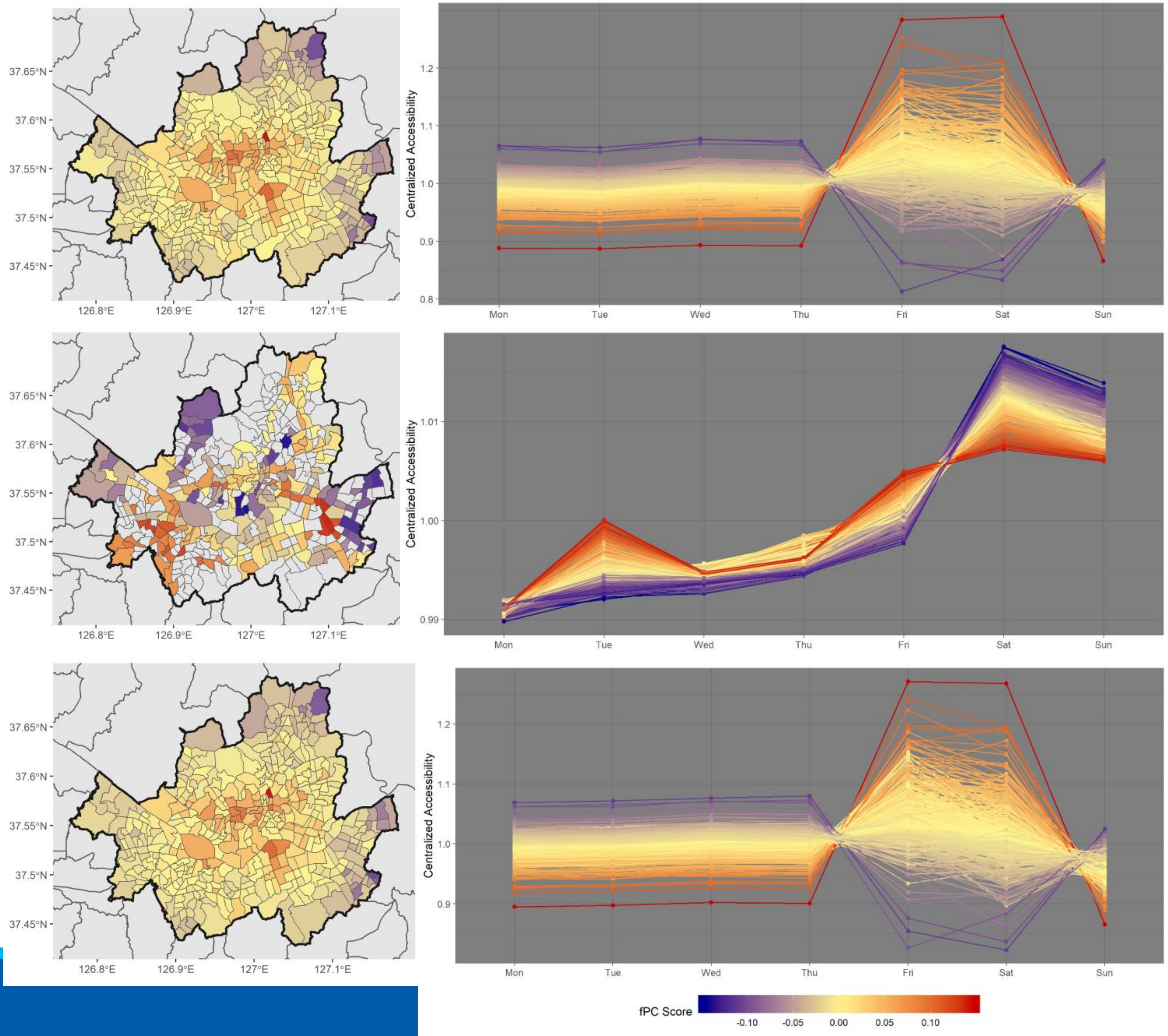
$$A_n(t) = \bar{A}_n \times f_n(t)$$

$$f_n(t) = \mu(t) + \sum_{k=1}^{K^*} \gamma_{nk}\xi_k(t) + \varepsilon_n(t),$$

Bus

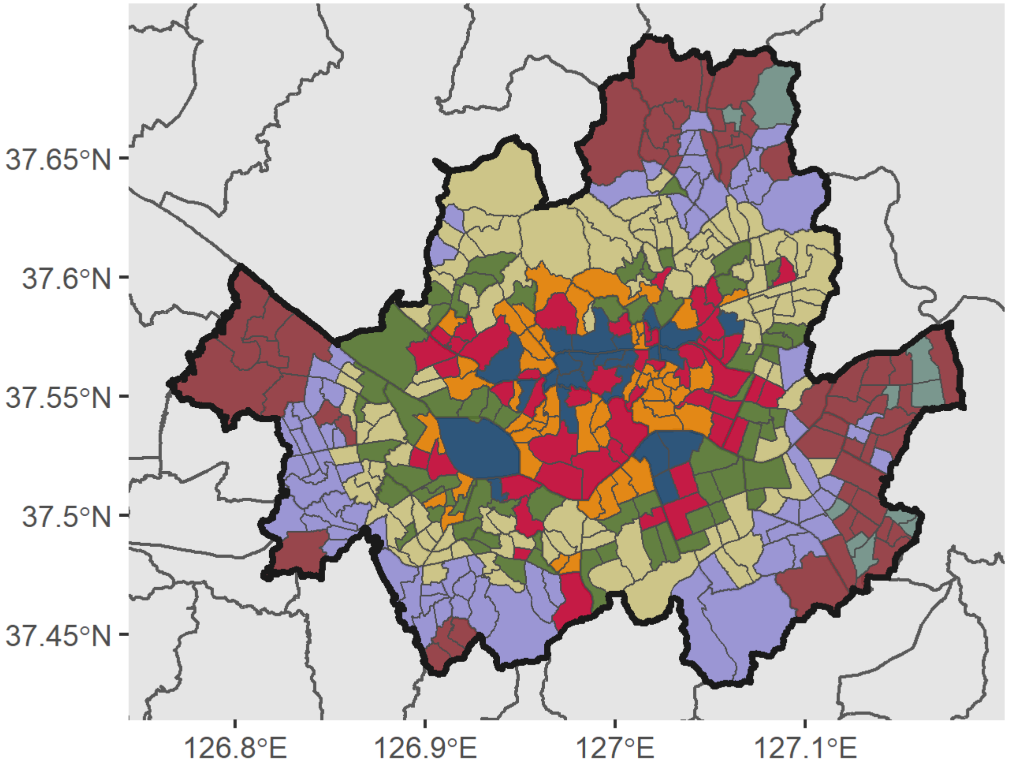
Subway

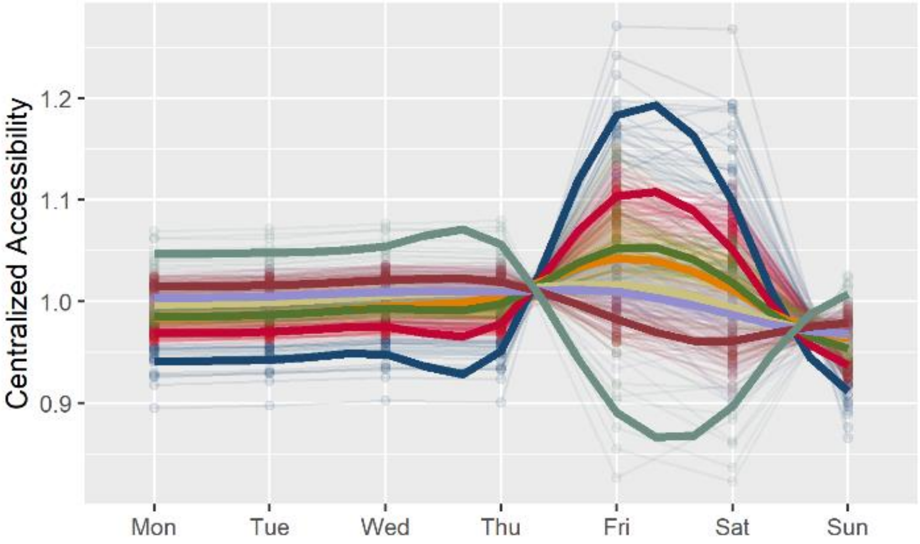
Integrated

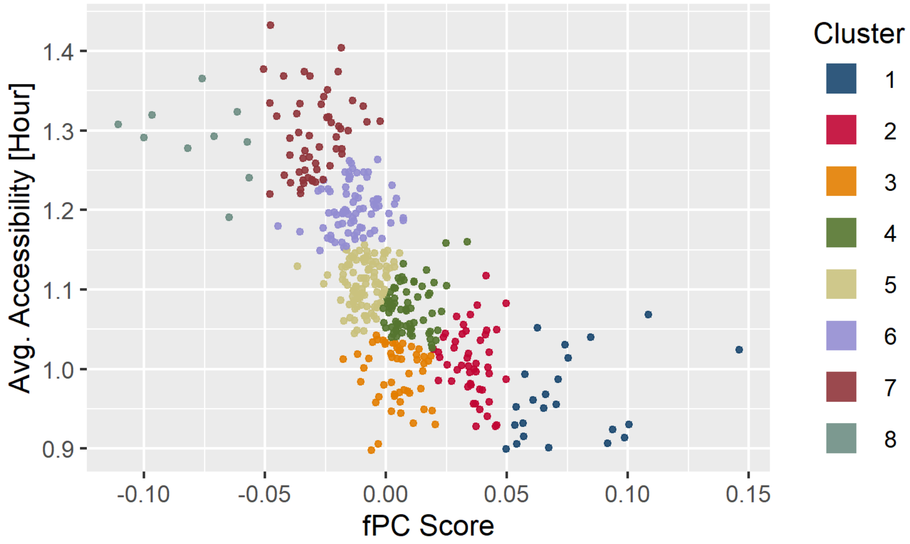


Day-to-Day Dynamic Time Accessibility: Neighborhood Clusters

- Cluster by Dynamic Time Accessibility: Integrated







Conclusions

- FDA-based approach addresses a neighborhood growth process as a time-dependent and continuous curve, by taking a multivariate curve as a unit of analysis
- FDA-based approach easily deals with spatiotemporal data to explain neighborhood changes and dynamics of transportation systems
- Applications to U.S. metropolitan areas clearly address racial and socioeconomic segregation patterns in longitudinal perspective
- mfPCA also explains spatial disparity in accessibility and its fluctuation across neighborhoods
- By comparing curve shapes, mfPCA and clustering can identify neighborhoods which have similar neighborhood trajectories and neighborhoods with similar dynamic time accessibility

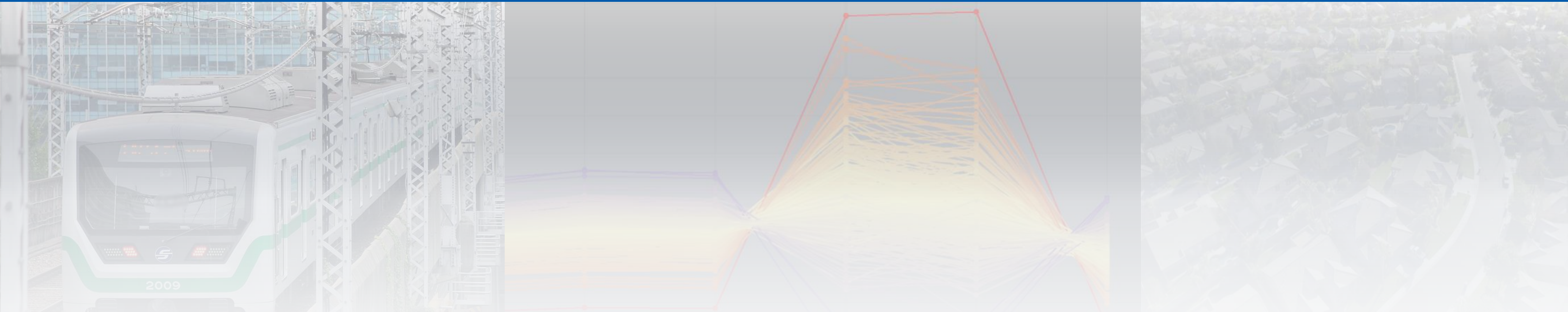
Future Research

- FDA can combine different data (census + satellite image data + health data)
  - With different data frequency, FDA can combine them together and address more complex spatial phenomena

- FDA-based neighborhood index can be further developed
  - FDA-based spatial statistical inferences are now under development
  - Spatial features of the neighborhood process can be considered (e.g., spatial diffusion)

- Further applications to transportation systems
  - Finer temporal granularity (Within-day travel patterns)

- Wider applications: Long-term neighborhood changes in other countries
  - (U.S.) Adding 2020 and pre-1970 census data
    Using 5-year ACS data with FDA can explain neighborhood changes with higher temporal granularity
  - (UK) Regional decline with deindustrialization in the Northern England industrial cities
  - (South Korea) Population degrowth issues in small and middle metro areas



Examining Longitudinal Neighborhood Change and Human Mobility Patterns with Functional Data Analysis Approach

UK-South Korea Geospatial Data Science Seminar
UBAI, University of Seoul / June 9th 2023

Paul Hayoung Jung

Inha University & UC Riverside

Paul.Jung@inha.ac.kr