도시에서의 이동데이터 활용 및 연구 사례

발표일: 2025년 8월 4일 월요일, 8월 5일 화요일

발표자: 서울시립대학교 도시공학과 도시문화공간계획연구실 석사과정 김혜인

hye2in@uos.ac.kr / hyein2kim@gmail.com

목차

- 1. 이동데이터와 도시
 - 1) 이동데이터의 의의
 - 2) 집계방식에 따른 이동데이터의 종류
 - 3) 이동데이터의 적용 사례
- 2. 연구사례: 생활이동데이터를 통해 바라본 서울시의 사회적 혼합
 - 1) 분석개요
 - 2) 분석방법 및 데이터
 - 3) 분석결과 및 종합
- 3. 참고문헌

이동데이터의 의의

Pandemics

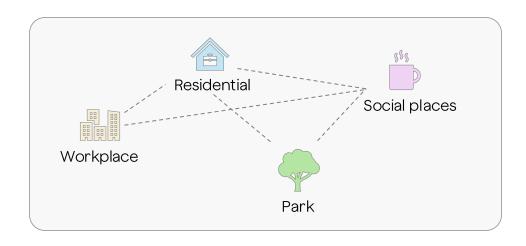
Emergency events

Natural disaster

Transit system

Personal preference

Human Mobility



- 다양한 사회적, 개인적 요인들과 상호 영향을 주고 받는 사람들의 이동
- 이동 양상을 관찰하고, 분석하고, 예측함으로써 이동을 통해 잠 재적으로 변화할 도시 공간의 변화를 미리 파악하고 계획할 수 있음 (Xu et al., 2023).
- 이동 데이터는 정적인 인구 데이터와 달리, 시공간적 특징을 지니고 있으며,
- 사람들의 활동 공간 및 활동 패턴을 추론 가능하도록 하여 실제 사람들의 행태를 기반으로 한 의사결정에 도움을 줄 수 있음

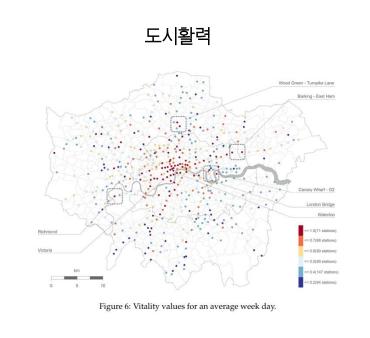
* 활동 공간 (Activity space): 사람들이 일상생활을 하는 동안 이동하는 지리적 범위 (Ren, 2016).

이동데이터가 담고 있는 도시맥락

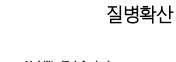
- 도시의 공간구조, 연결성, 활력, 전염병 및 여러 자원의 전파 등 사람들의 움직임을 통해 형성되는 현상에 대한 대리변수(proxy)로서 활용됨
- 직접적인 이동데이터 대신에, 경우에 따라 교통수단 데이터, 신용카드 데이터, SNS 위치정보 데이터 등의 지리공간데이터가 사용되기도 함
- 대표 예시 3가지

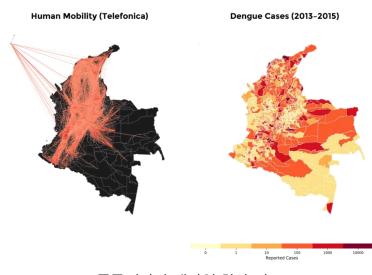
Shenzhen 10,722.487 people/km² 10,722.487 people/km² 10,722.487 people 1,956 km² Polycentric Monocentric Bogota 7,605,691 people 458 km² Polycentric Monocentric Bogota 7,605,691 people 4,827,518 people 10,865 km² Lisbon 3,399,479 people 12,409 km² Metropolitan area Main center Sub-center Non-center

Urban dynamics through the lens of human mobility Xu, Y., Olmos, L. E., Mateo, D., Hernando, A., Yang, X., & González, M. C. (2023)



Using mobility data as proxy for urban vitality
Sulis, P., Manley, E., Zhong, C., & Batty, M. (2018)





<mark>콜롬비아의 댕기열 확산 지도</mark> UNICEF. (n.d.)

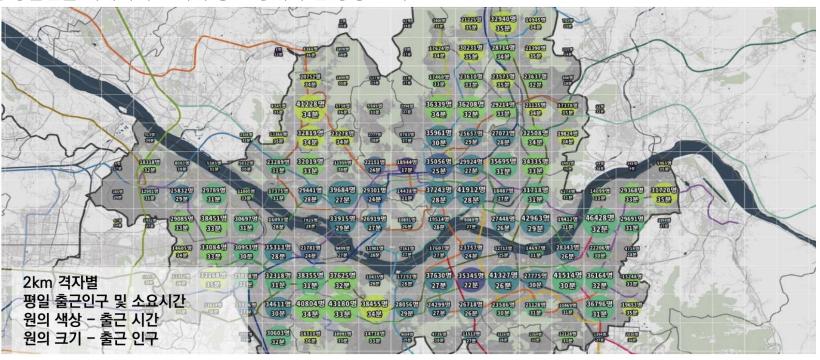
이동데이터의 집계방식에 다른 차이

- 과거에는 사람들의 이동행태를 파악하기 위해서 설문조사를 통해 데이터를 수집
- 대표적인 예로, 교통계획 목적으로 진행되는 가구통행실태조사가 있음
 - 교통수단, 통행시간, 통행거리, 통행목적 개인에 대한 자세한 정보를 제공하지만 표본수가 적으며 (김승준 외, 2023), 그 주기 또한 5년마다 진행되어 빠르게 변화하는 이동 행태를 연속적으로 관찰하기에는 어려움이 있음
- 통신네트워크가 발달되면서, 모바일 폰을 통해 생성된 수많은 양의 데이터가 도시계획을 비롯한 다양한 분야에 활용되기 시작 (Calabrese et al., 2014).

| | 궤적 (Trajectory) 「 | OD(Origin-Destination) Flow 데이터 | | | | | | |
|---------------|--------------------------|---------------------------------|--------|----|--------|-----------|-----------|------------|
| activity_type | end_location | end_time | | | | 출발 행정동 코드 | 도착 행정동 코드 | 이동인구(합) |
| Movement | geo:37.568694,126.980433 | 2024-03-05T07:40:00.009+ | | | 0 | 1101053 | 1101053 | 28113.60 |
| Movement | geo:37.568694,126.980433 | 2024-03-05T08:13:00.009+0 | | | 1 | 1101053 | 1101054 | 1699.65 |
| Movement | geo:37.568502,126.980043 | 2024-03-05T08:32:20.009+ | | | 2 | 1101053 | 1101055 | 1545.06 |
| | | | | | 3 | 1101053 | 1101056 | 1105.87 |
| Visit | geo:37.568502,126.980043 | 2024-03-05T08:55:37.692+0 | | | 4 | 1101053 | 1101057 | 1732.11 |
| Movement | geo:37.568640,126.979862 | 2024-03-05T08:56:17.000+0 | | | | | | |
| timoling 0.2. | -집된 궤적데이터 데이터프 | 대이 인브 예시 (자체 구최 | · 데이터) | И€ | 2시 해저드 | 등 단위 생활이등 | 트데이터 예시 (| (2023년 3원) |

서울시/KT 분석 사례 – 생활권 계획

- 서울 시민들 중 어디에 사는 사람들이 출근하는데 오래 걸리는가? 규모는 얼마나 되는가?
 - 지역별 생활권을 파악하여 근거리 통근 정책 추진 방향 모색



수도권 생활이동 데이터 매뉴얼 서울특별시, KT (2024)

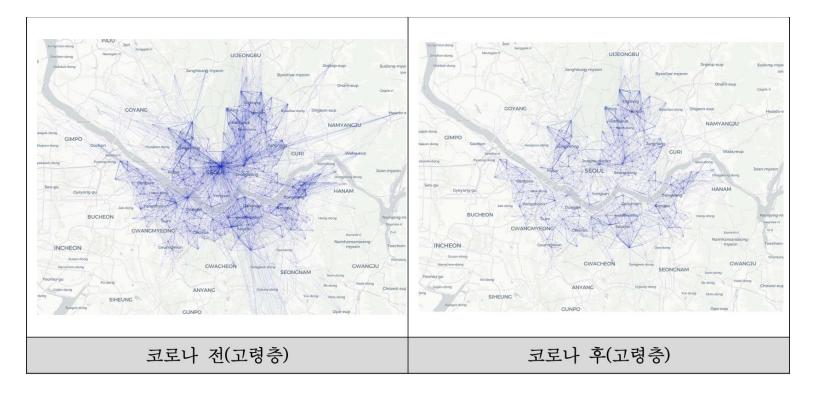
서울시/KT 분석 사례 - 광역계획

- 수원 장안구에서 목적지별 교통여건 개선 전략
 - 수혜 인원을 고려하여 지역별 상황에 맞는 대중교통 해결방안 모색



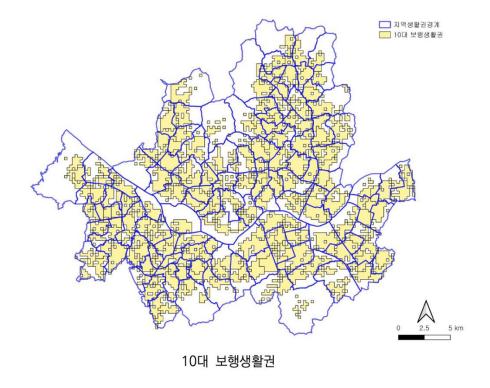
수도권 생활이동 데이터 매뉴얼 서울특별시, KT (2024)

코로나19 전후의 이동패턴 변화 분석 사례



생활이동 데이터를 활용한 코로나19 전후의 연령별 통행패턴변화 및 변화요인 분석 전승호 (2023)

시민들이 걸어서 접근가능한 생활권 분석

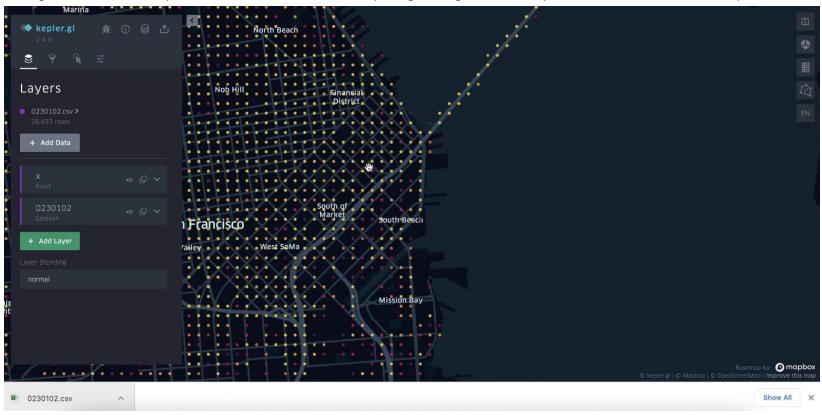


생활이동데이터를 활용한 서울시 연령대별 보행생활권과 지역생활권의 비교 연구 이지원 (2025)

이동데이터의 적용 사례 2: 기타 이동 데이터

Mapbox의 이동데이터 시각화 사례

(https://www.mapbox.com/blog/how-to-utilize-mapbox-movement-data-for-mobility-insights-a-guide-for-analysts-data-scientists-and-developers)

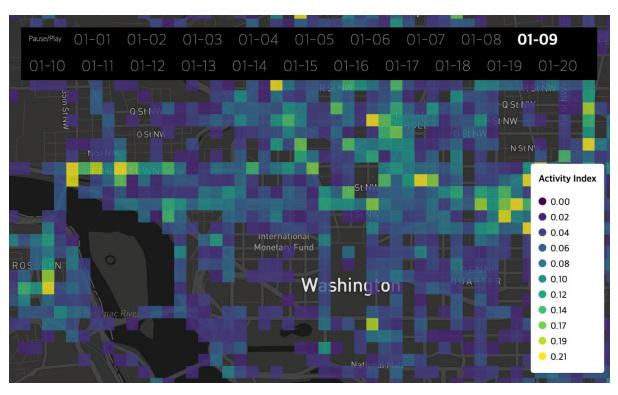


Kepler.gl 을 활용한 이동 데이터 시각화 Mapbox(2021)

이동데이터의 적용 사례 2: 기타 이동 데이터

Mapbox의 이동데이터 적용 사례

(https://www.mapbox.com/blog/inauguration-vs-pandemic-what-can-shut-down-washington-d-c)



일일 보행자 및 차량 데이터를 활용하여 워싱턴 D.C.의 활동량 분석 및 시각화 Mapbox(2021)

연구사례

제목 생활이동데이터를 통해 바라본 서울시의 사회적 혼합 (김혜인, 이희정)

분석개요

연구배경

- 도시민들의 사회적 고립의 지속적인 증가. 이는 COVID-19 이후로 더욱 심화됨 (NASEM, 2024).
- 사회적 고립은 개인적 차원에서 신체적, 정신적 건강, 삶의 질에 영향을 미치며, 이는 나아가 공동체와 전체 사회에도 영향을 미침 (WHO, 2023).
- 특히, 사람들 간의 경제적 격차는 이러한 양상을 더욱 촉진시킴 (Lee & Han, 2024).
- 도시의 상호작용을 촉진하는 역할로, 제3의 장소들 (카페, 도서관, 공원 등)이 주목받고 있음
- 다양한 사회경제적배경의 사람들이 공존하는 공간을 조성하기 위해, 제3의 장소들의 역할을 검토하는 것이 필요
- 이러한 장소들이 다양한 사람들의 공존을 유도하는지 정량적으로 파악하는 것은, 사회적 혼합을 고려한 공간 전략 수립에 기초 자료를 제공할 수 있음

연구질문 및 분석목표

- 서울시에서 사회경제적 배경이 다른 사람들이 모이는 공간은 어디인가?
- 해당 공간의 공공문화시설이 분포는 어떠한가?

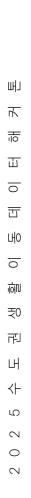
분석의 범위

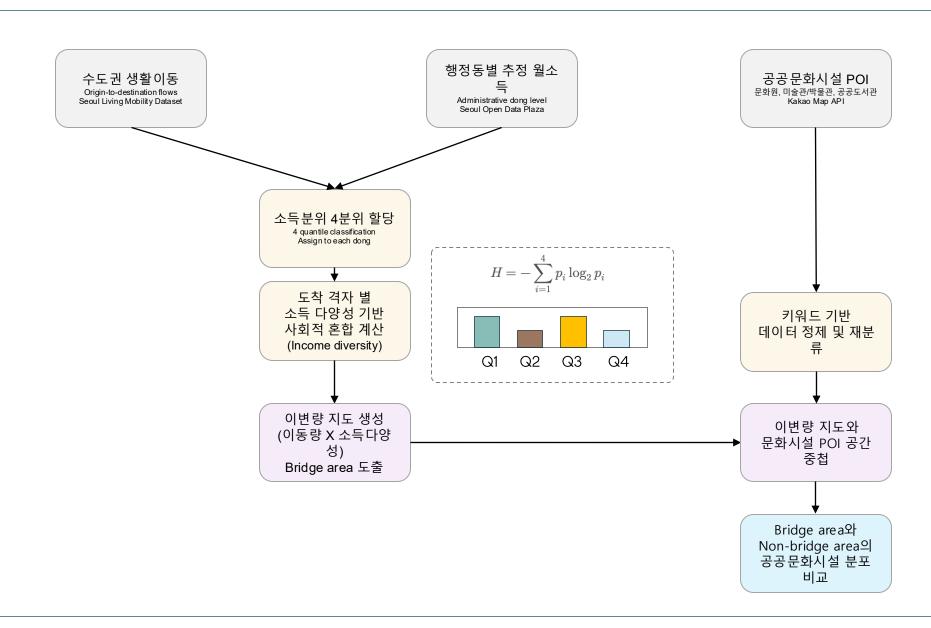
- 공간적 범위: 서울시
- 시간적 범위: 2025년 1월~5월 (5개월)
- 분석의 단위: 250m 격자

데이터 및 분석 방법

분석데이터

| 데이터명 | 출처 |
|------------------------|-------------|
| 수도권 생활이동 데이터 | 서울시 빅데이터캠퍼스 |
| 서울시 행정동별 추정 소득 및 추정 소비 | 서울 열린데이터 광장 |
| 공공문화시설 POI | 카카오맵 API 수집 |





분석과정

- 데이터 용량: 1개월당 약 9GB (5개월 분 총 약 45GB)
- 분석 SLRUM 작업 설정 요건
 - 파티션: CPU1
 - 노드 수: 1
 - 태스크 수: 1
- 분석 언어: Python
- 연산 처리 라이브러리: Polars
- 순수 분석 처리 소요 시간: 전처리 (50분) + 지표계산 (15분) + 시각화(6분) = 총 약 1시간 10분



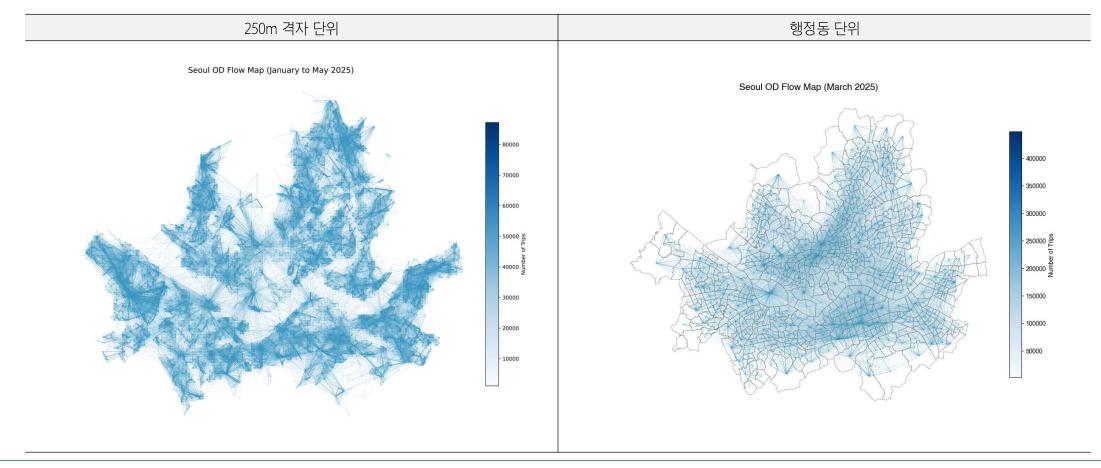
- 즉시 연산을 실행하지 않고, LazyFrame을 통해 여러 연산을 한번에 계획한 후 collect() 시점에서 쿼리를 최적화하여 일괄 실행하는 지연 평가 방식을 사용. 메모리 사용량을 최소화 하면서도 병렬 처리를 통해 Pandas 대비 훨신 빠른 대용량 데이터 처리가 가능

| # display(| base_ldf.h | nead().co | ollect()) | | | | | | | | | | | | | | | |
|----------------|--------------------|-----------|-----------|--------------------|----------|----------|-----------|-----------|-----------|-----------|-----------|-------------|--------------|-------------|------------|---------|----------|----------|
| shape: (5, 19) | | | | | | | | | | | | | | | | | | |
| date | o_cell_id | o_cell_x | o_cell_y | d_cell_id | d_cell_x | d_cell_y | o_admi_cd | d_admi_cd | move_dist | move_time | total_cnt | adolescents | young_adults | middle_aged | late_adult | st_hour | fns_hour | day_type |
| date | str | i64 | i64 | str | i64 | i64 | i64 | i64 | i64 | i64 | f64 | f64 | f64 | f64 | f64 | i64 | i64 | i8 |
| 2025- 01-01 | "다사 53005375" | 953125 | 1953875 | "다사 52505350" | 952625 | 1953625 | 11110515 | 11110515 | 585 | 3 | 3,25 | 0,0 | 0,0 | 0,0 | 3,25 | 0 | 0 | 0 |
| 2025- 01-01 | "ር⊦ለት 53005375" | 953125 | 1953875 | "다사 53255350" | 953375 | 1953625 | 11110515 | 11110515 | 304 | 2 | 2,7 | 0.0 | 0,0 | 2,7 | 0.0 | 0 | 0 | 0 |
| 2025- 01-01 | "ር⊦ሉት 52755325" | 952875 | 1953375 | "ር⊦ሉት 52755325" | 952875 | 1953375 | 11110515 | 11110515 | 50 | 1 | 3,69 | 0.0 | 0,0 | 0,0 | 3,69 | 0 | 0 | 0 |
| 2025- 01-01 | "다사 52505350" | 952625 | 1953625 | "ር⊦ሉት 53005450" | 953125 | 1954625 | 11110515 | 11110515 | 1118 | 2 | 2,6 | 0,0 | 0,0 | 2,6 | 0,0 | 0 | 0 | 0 |
| 2025- 01-01 | "ር⊦ሉት 52755350" | 952875 | 1953625 | "ር⊦ለት 52755325" | 952875 | 1953375 | 11110515 | 11110515 | 200 | 0 | 2,79 | 0,0 | 0,0 | 2,79 | 0,0 | 0 | 0 | 0 |

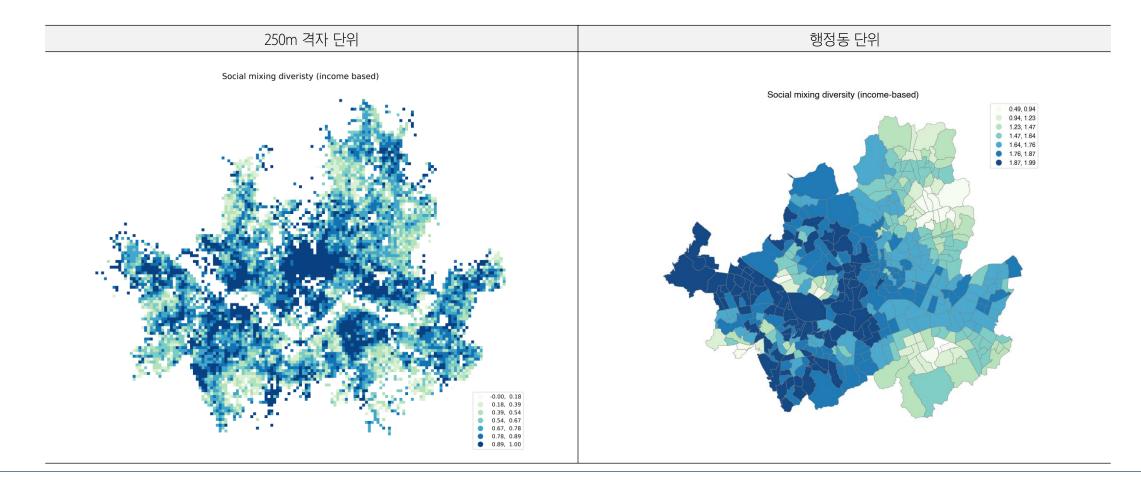
연산에 포함한 전처리 조건

- 출발 격자 타입: 야간 상주지
- 연령 그룹화
- 서울 내 이동으로 한정
- 날짜 데이터 Date 타입 변환
- 주중 주말 컬럼 생성
- 20분단위 시간 1시간 단위로 조정
- 주중 주말 컬럼 생성
- · 전처리 전후의 총 이동 수 (total_cnt) 합산
 - 전: 약 69억건
 - 후: 약 9억건

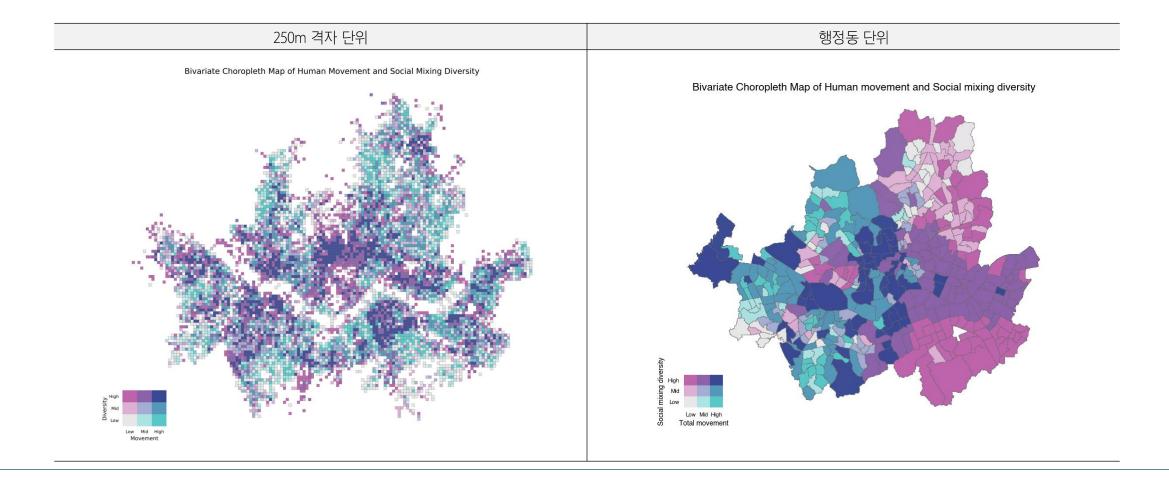
- OD Matrix
 - 생활이동으로 형성되는 공간구조 파악을 위해 출발지-도착지 간 네트워크 시각화
 - 노드: 출발지, 도착지링크 가중치: 이동 건수



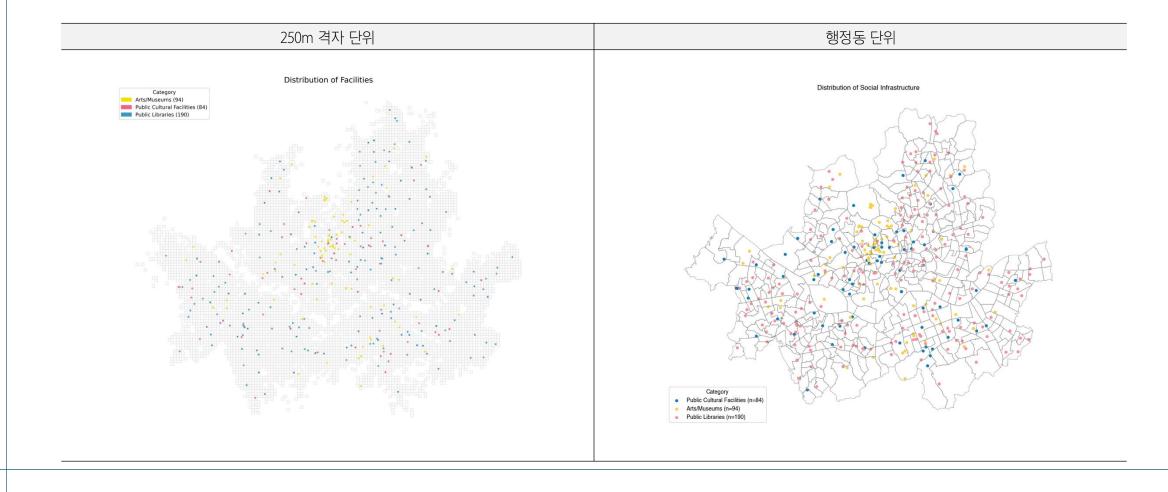
- 소득 기반의 사회적 혼합 지수 계산
 - 이동 데이터에 포함된 거주지 정보 (야간상주지)에 소득수준을 매칭하여, 각 지역을 방문하는 사람들의 경제적 배경이 얼마나 다양한지 (사회적 혼합)을 계산



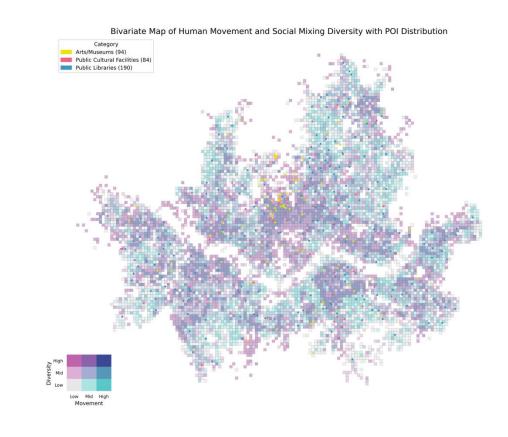
- Bridge area 도출
 - 사회적 혼합도(다양성)과 이동한 사람의 수(밀도)를 함께 고려한 Bridge area 도출을 위해 이변량 지도를 생성

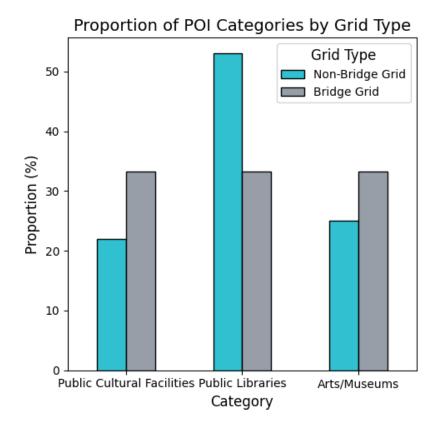


- 공공문화시설 POI 공간 분포 확인
 - Bridge area와 공공문화시설과의 관계 분석을 위해, 카카오맵 POI로 수집한 공공문화시설의 공간적 분포 확인



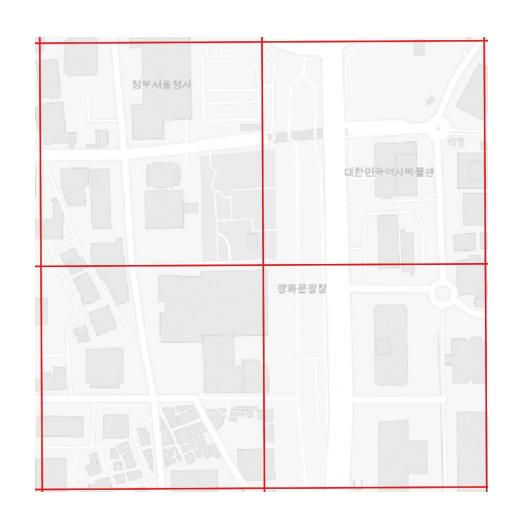
- Bridge area와 Non-bridge area의 공공문화 시설 분포 개수 비교





종합

- 격자 단위 분석의 장점:
 - 1. 행정동 단위 분석과 비교하였을 때, 동일한 지수를 보다 세부적인 스케일에서 계산 및 도출하는 것이 가능하였음 (예: 사회적 혼합 도는 동일한 행정동 내에서도 서로 다른 분 포를 보인다)
 - 2. 행정동 단위 데이터와 비교했을 때, 보행으로 이동 가능한 근린 단위 스케일로 실제 일상생활 반경과 부합하는 측면
- 제안 사항
 - 대용량의 데이터인 만큼, 담고 있는 정보가 방대함 (성연령, 주간상주지, 야간상주지, 이 동목적, 이동시간, 이동거리 등) 연산 시간 단 축 및 분석 결과의 간결함을 위해 전처리가 가장 중요
 - 이동 데이터는 최신 데이터이기 때문에, 다른 데이터 결합시 행정동 불일치가 있을 수 있음 (통폐합 등)



- 김승준, 양재환, 박세현. (2023). 인공지능 활용한 교통데이터 통행목적과 이용자특성 추정. 서울연구원.
- 이지원. (2025). 생활이동데이터를 활용한 서울시 연령대별 보행생활권과 지역생활권의 비교 연구. 석사학위논문. 서울시립대학교 대학원.
- 전승호. (2023). 생활이동데이터를 활용한 코로나19 전후의 연령별 통행패턴변화 및 변화요인 분석. 석사학위논문. 서울대학교 대학원.
- Francesco Calabrese, Laura Ferrari, and Vincent D. Blondel. (2014). Urban Sensing Using Mobile Phone Network Data: A Survey of Research. ACM Comput. Surv. 47, 2,
- Lee, Y., & Han, S. (2024). Perceptions of inequality and loneliness as drivers of social unraveling: Evidence from South Korea. Scientific Reports, 14(1)
- Mapbox. (2021). Inauguration vs. Pandemic: What Can Shut Down Washington, D.C.? Mapbox Blog. https://www.mapbox.com/blog/inauguration-vs-pandemic-what-can-shut-down-washington-d-c
- Mapbox. (2023). How to Utilize Mapbox Movement Data for Mobility Insights: A Guide for Analysts, Data Scientists, and Developers. Mapbox Blog. https://www.mapbox.com/blog/how-to-utilize-mapbox-movement-data-for-mobility-insights-a-guide-for-analysts-data-scientists-and-developers
- National Academies of Sciences, Engineering, and Medicine [NASEM]. 2024. Addressing the Impact of COVID-19 on Social Isolation and Loneliness. Washington, DC: The National Academies Press.
- Ren, Fang (2016). Activity Space. obo in Geography.
- Sulis, P., Manley, E., Zhong, C., & Batty, M. (2018). Using mobility data as proxy for measuring urban vitality. Journal of Spatial Information Science, 16, 137–162.
- UNICEF. (n.d.). Fighting epidemics. UNICEF. https://www.unicef.org/innovation/fighting-epidemics
- Xu, Y., Olmos, L. E., Mateo, D., Hernando, A., Yang, X., & González, M. C. (2023). Urban dynamics through the lens of human mobility. Nature Computational Science, 3(7), 611–620.

| 권생활이동데이터해거톤 | 감사합니다 |
|-------------|-------|
| 버 | |
| ≺⊢ | |
| 5 | |
| 2 | |
| 2 0 | |
| 23 | |