

2023-1 데이터사이언스캡스톤프로젝트 Final Presentaiton

서울시 공동육아나눔터 입지 선정

포동포동

김보현 김민



CONTENTS

01

서론

02

연구 내용

- EDA
- 행정동 필터링
- 최적 입지 선정
- 결과 분석

03

결론

주제 선정 배경

서론

데이터별 EDA

행정동 필터링

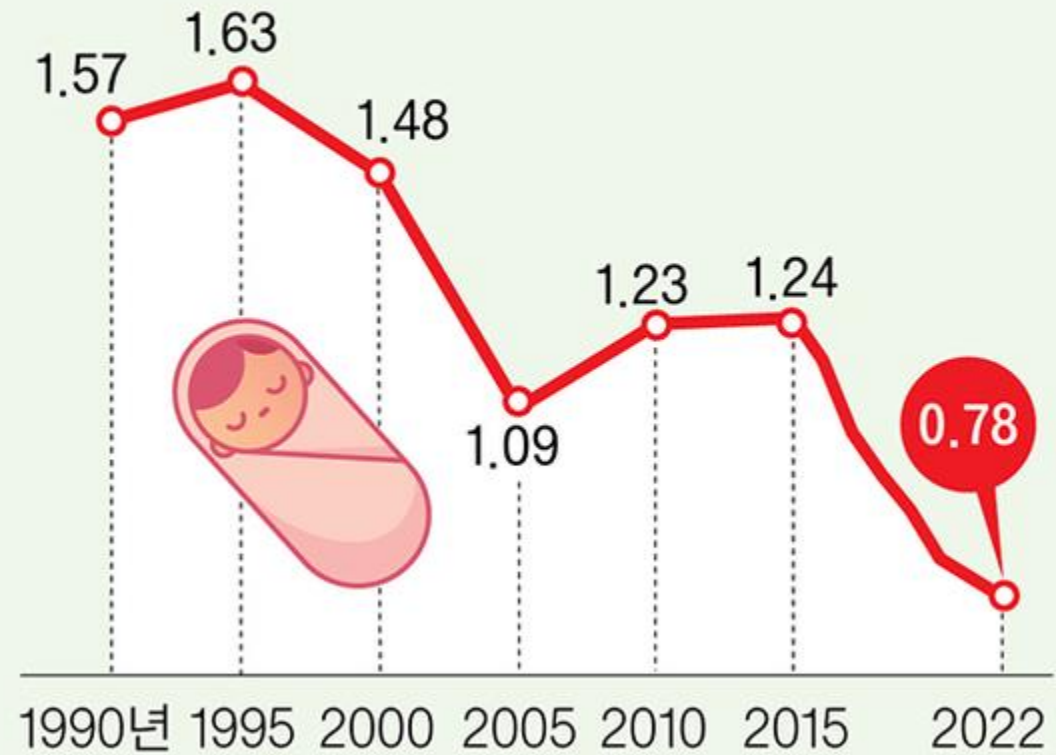
최적 입지 선정

결과 분석

결론

연간 합계출산율 추이

단위:명



자료:통계청

우리 사회에서 점차 심각해지고 있는 저출산 문제

어떻게 해결할 수 있을까?

출산과 육아에 대한 개인의 부담이 줄어들어야



국가 차원의 제도적 노력이 필요

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

서울시 공동육아나눔터 입지 선정

가족 돌봄 기능을 보완해 부모의 육아부담을 줄이고,
양육 친화적인 사회 환경을 조성하고자 운영되는 공간



서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

서울시 공동육아나눔터 입지 선정

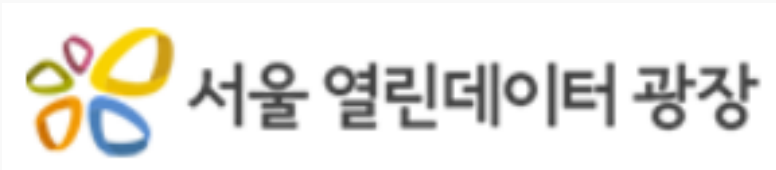
서울시 시민들이 공동육아나눔터를 최대한 활용할 수 있도록, 최적 입지 선정을 진행

-> 현재 서울시 내 설치된 공동육아나눔터의 위치를 고려해, 추가 입지 제안



- 서론
- 데이터별 EDA
- 행정동 필터링
- 최적 입지 선정
- 결과 분석
- 결론

출처:



| 카테고리 | 데이터셋 |
|------|-------------------------------|
| 아동 | 서울시 어린이집 정보 |
| | 서울시 행정동별 어린이 보호구역 지정 통계 |
| | 서울시 공동육아나눔터 정보 |
| 인구 | 행정동 단위 생활인구 |
| | 서울시 우리마을가게 상권분석 서비스 (상권-직장인구) |
| | 서울시 혼인이혼 (동별) 통계 |
| | 서울시 저소득 한부모 가족 통계 |
| 예산 | 서울시 예산결산총괄 통계 |

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

최종 데이터셋 생성

1. 각 데이터셋 병합

- 행정동 & 자치구 기준으로 병합
- 행정동 정보가 없는 데이터 - 주소 데이터를 정제하여 위도 & 경도로 변환 > 카카오 API로 행정동 추출
- 자치구별 데이터 - 자치구 내 모든 데이터에 같은 값 할당

2. 결측치 보간

- 각 데이터의 특성에 맞는 다양한 방법을 활용
 - 직장인구 - 생활인구 데이터 기준으로 클러스터링 진행, 클러스터링의 평균으로 결측치 보간
 - 어린이집 - 네이버 지도 검색 결과 기반, 직접 입력

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

| | 통계청 행정동 코드 | 행정부행 정동코드 | 시 군 구 명 | 행 정 상 영 영 | 총생활인구수 | 남자0세부터 9세생활인구 수 | 남자10세부 터14세생활 인구수 | 남자30세부 터34세생활 인구수 | 남자35세부 터39세생활 인구수 | 남자40세부 터44세생활 인구수 | ... | 2021_예 산현액 | 2021_세 입 | 2021_세 출 | 총_직장_인구 _수 |
|-----|------------------|--------------|------------------|-----------------------|--------------|-----------------------|-------------------------|-------------------------|-------------------------|-------------------------|-----|---------------|-------------|-------------|---------------|
| 0 | 1101053 | 11110530 | 중 구 | 사 직 영 | 23050.525410 | 376.461406 | 167.316184 | 1036.331230 | 1290.318078 | 1199.036465 | ... | 662770 | 695108 | 540729 | 10012.710690 |
| 1 | 1101054 | 11110540 | 중 구 | 상 정 영 | 5056.038923 | 62.542217 | 47.922133 | 185.886491 | 225.446885 | 233.844939 | ... | 662770 | 695108 | 540729 | 837.927419 |
| 2 | 1101055 | 11110550 | 중 구 | 부 영 | 15235.972430 | 428.735398 | 332.435485 | 363.460493 | 487.424951 | 513.138530 | ... | 662770 | 695108 | 540729 | 733.290323 |
| 3 | 1101056 | 11110560 | 중 구 | 원 영 | 18042.284850 | 583.821039 | 422.296752 | 376.874308 | 519.554757 | 553.489355 | ... | 662770 | 695108 | 540729 | 2227.870968 |
| 4 | 1101057 | 11110570 | 중 구 | 무 영 | 6575.949430 | 231.906654 | 148.312047 | 145.121536 | 200.780400 | 195.878479 | ... | 662770 | 695108 | 540729 | 548.537174 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 419 | 1125070 | 11740690 | 강 구 | 안 영 | 4599.955456 | 241.866921 | 183.310289 | 110.437675 | 159.174426 | 173.434588 | ... | 1170691 | 1189507 | 995053 | 548.537174 |
| 420 | 1125071 | 11740700 | 강 구 | 안 영 | 27795.065810 | 1454.971673 | 733.966765 | 830.498214 | 1092.286804 | 1019.898946 | ... | 1170691 | 1189507 | 995053 | 545.500000 |
| 421 | 1125072 | 11740570 | 강 구 | 안 영 | 28347.804830 | 1237.421860 | 643.528617 | 897.732191 | 1075.315778 | 1073.704786 | ... | 1170691 | 1189507 | 995053 | 315.258064 |
| 422 | 1125073 | 11740610 | 강 구 | 안 영 | 44432.109490 | 1573.124435 | 703.279815 | 1641.266842 | 2151.273657 | 1768.883368 | ... | 1170691 | 1189507 | 995053 | 561.370656 |
| 423 | 1125074 | 11740685 | 강 구 | 안 영 | 55549.385130 | 1923.757141 | 930.263555 | 1992.566143 | 2203.675301 | 2017.696132 | ... | 1170691 | 1189507 | 995053 | 680.511521 |

424 rows x 31 columns

424 rows x 31 columns

최종 데이터셋 완성!



연구 방법

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

데이터 수집 및 전처리



EDA를 통한 행정동 및 자치구별 특징 분석



행정동 필터링을 위한 클러스터링 변수 생성



클러스터링 결과를 통한 행정동 필터링



LSCP와 MCLP 알고리즘을 통한 최적 입지 선정



최적 입지 선정지 결과 분석

서론

데이터별 EDA

행정동 필터링

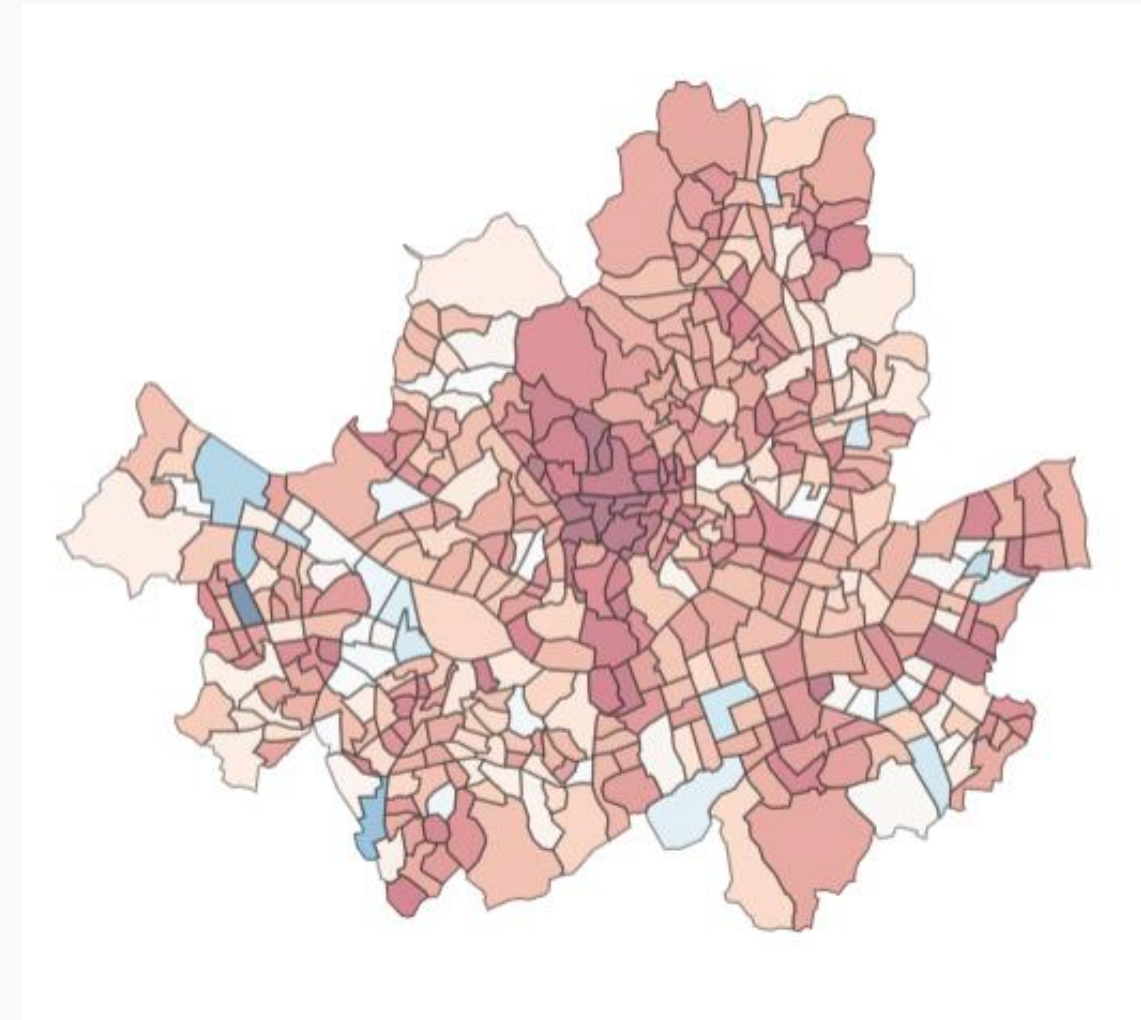
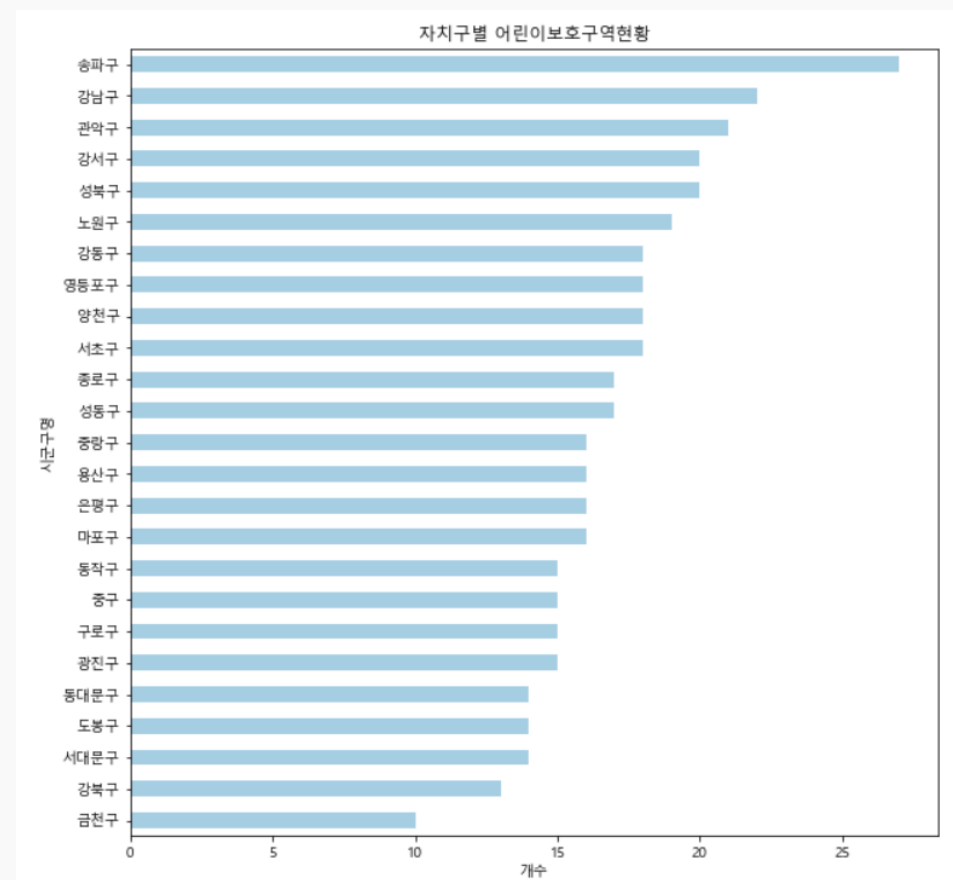
최적 입지 선정

결과 분석

결론

본 발표에서는,
변수 몇 개에 대해서만
살펴봐요

입지 선정의 기준을 명확히 하고자 행정동 및 자치구별 특징 분석 & 인사이트 추출
지도 시각화 및 상관분석 등을 진행



생활 인구

서론

데이터별 EDA

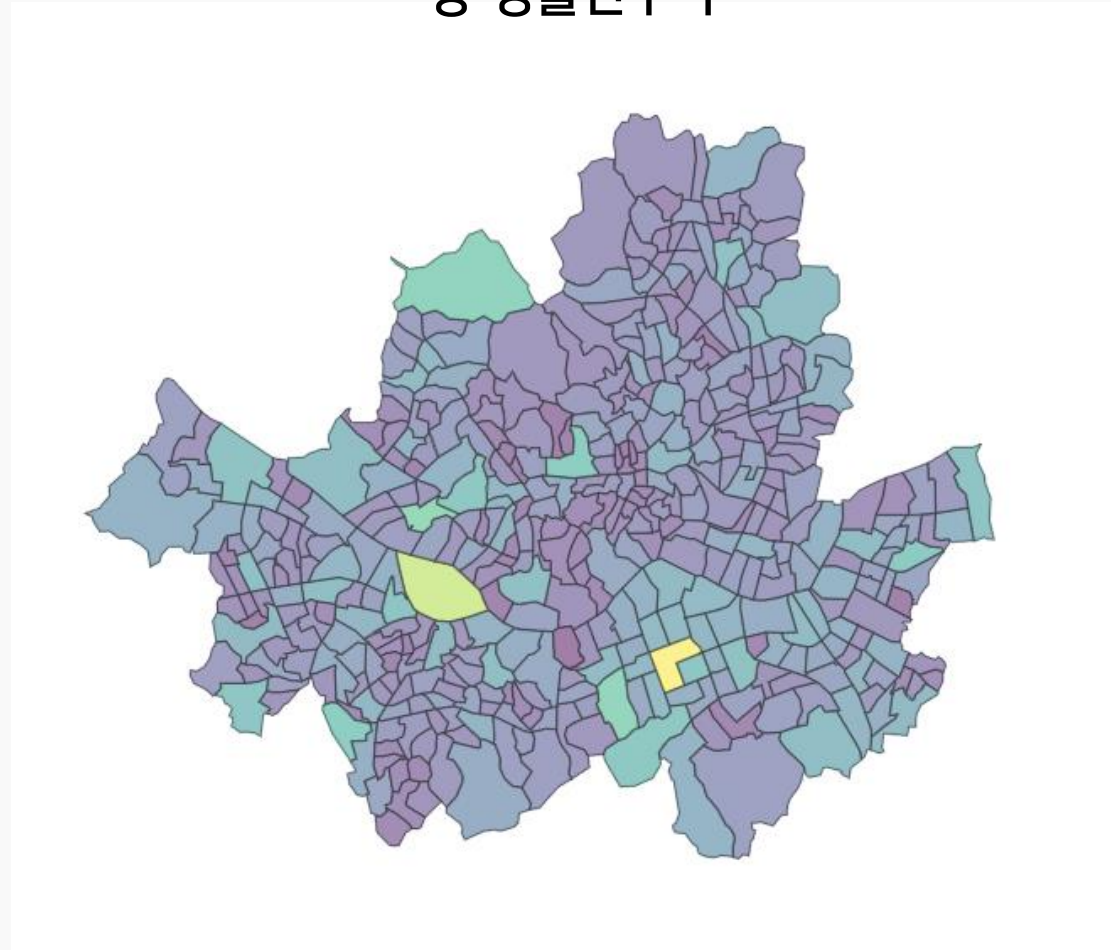
행정동 필터링

최적 입지 선정

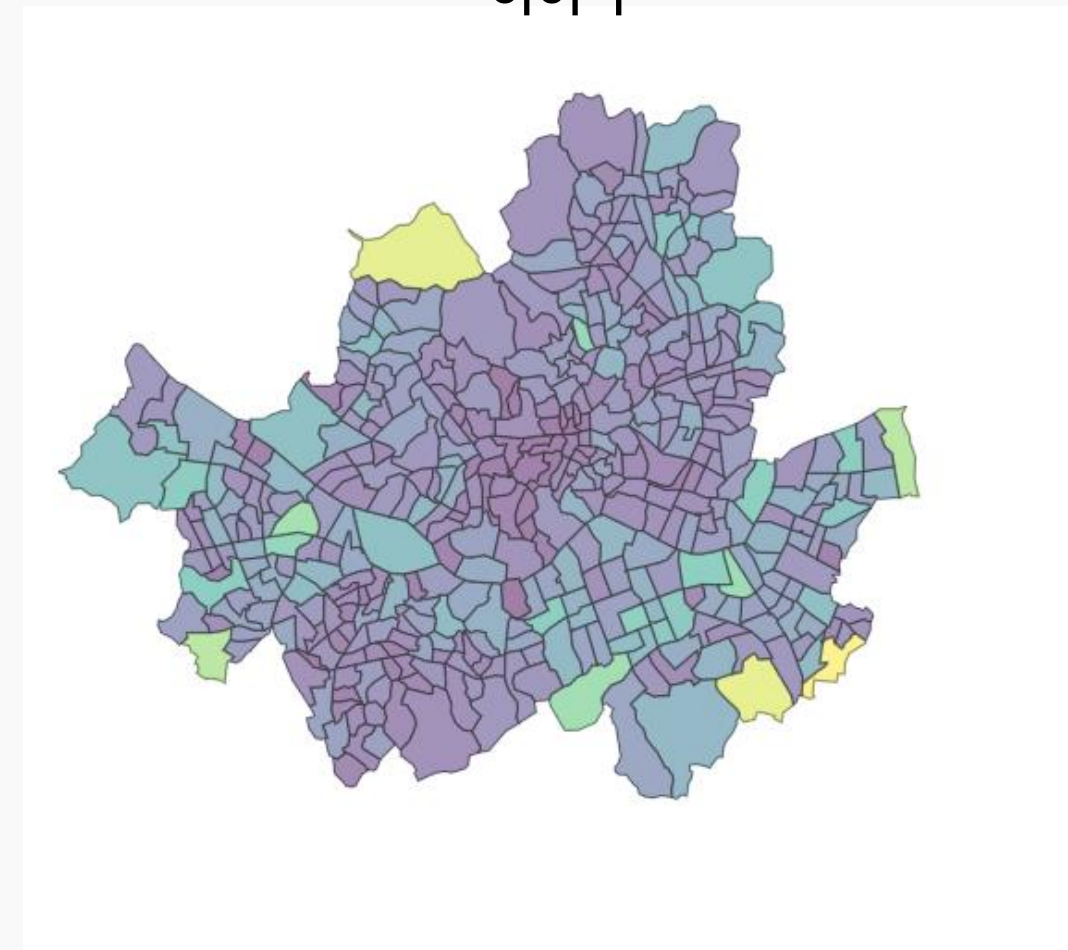
결과 분석

결론

총 생활인구 수



아이 수



총 생활인구가 서울 중심 지역에서 높게 나타난 것과는 달리,
아동 인구는 주거 지역이 많은 서울 외곽 지역에서 더 높게 나타남

* 위례동, 세곡동, 진관동, 강일동, 오류2동 등에서 아동 인구가 특히 높았음

20-40 직장인구

* 전체 직장인구 대비 신혼부부 연령대인 20-40대 직장인구의 비율

서론

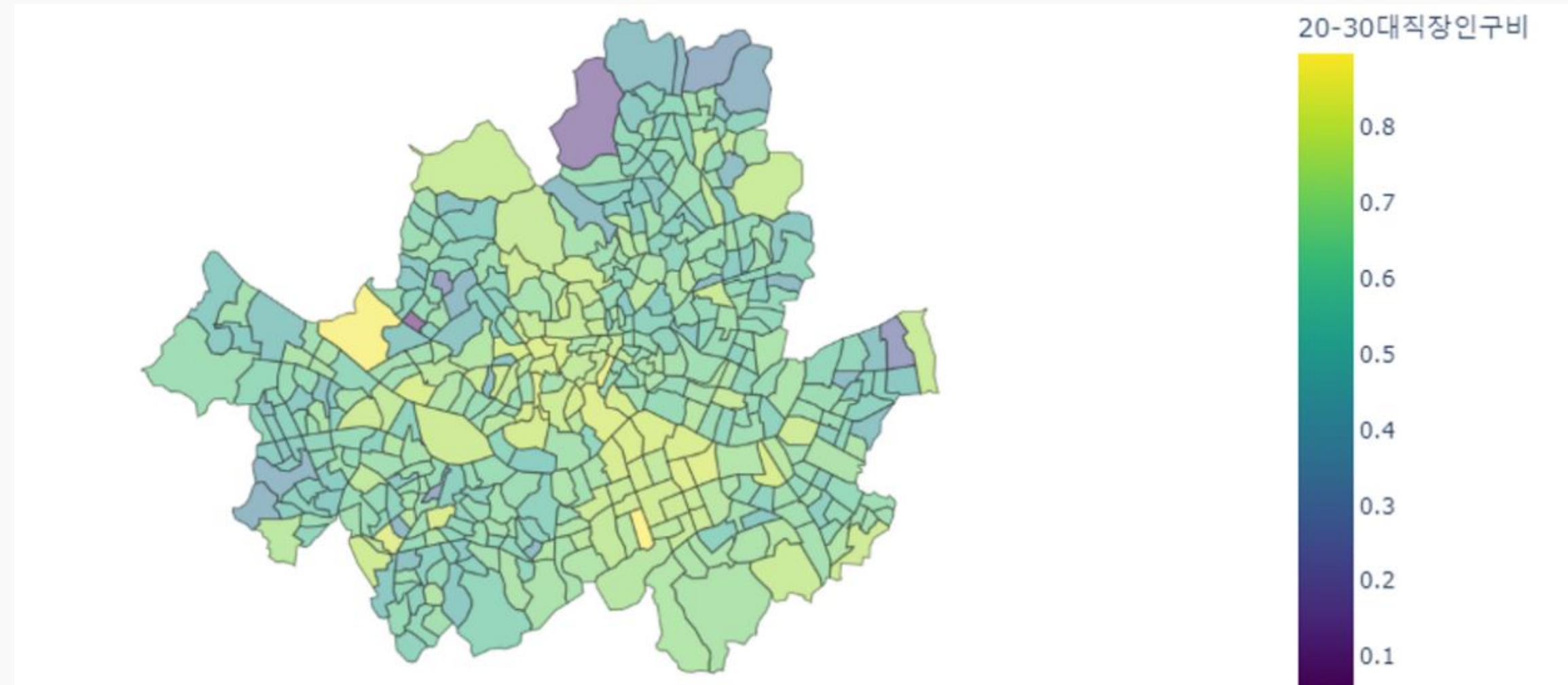
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



20-40대 직장인구의 경우, 마포구 상암동 및 용산구, 서초구, 강남구 등
한강 부근에서 높게 나타남

한부모/저소득 가정

서론

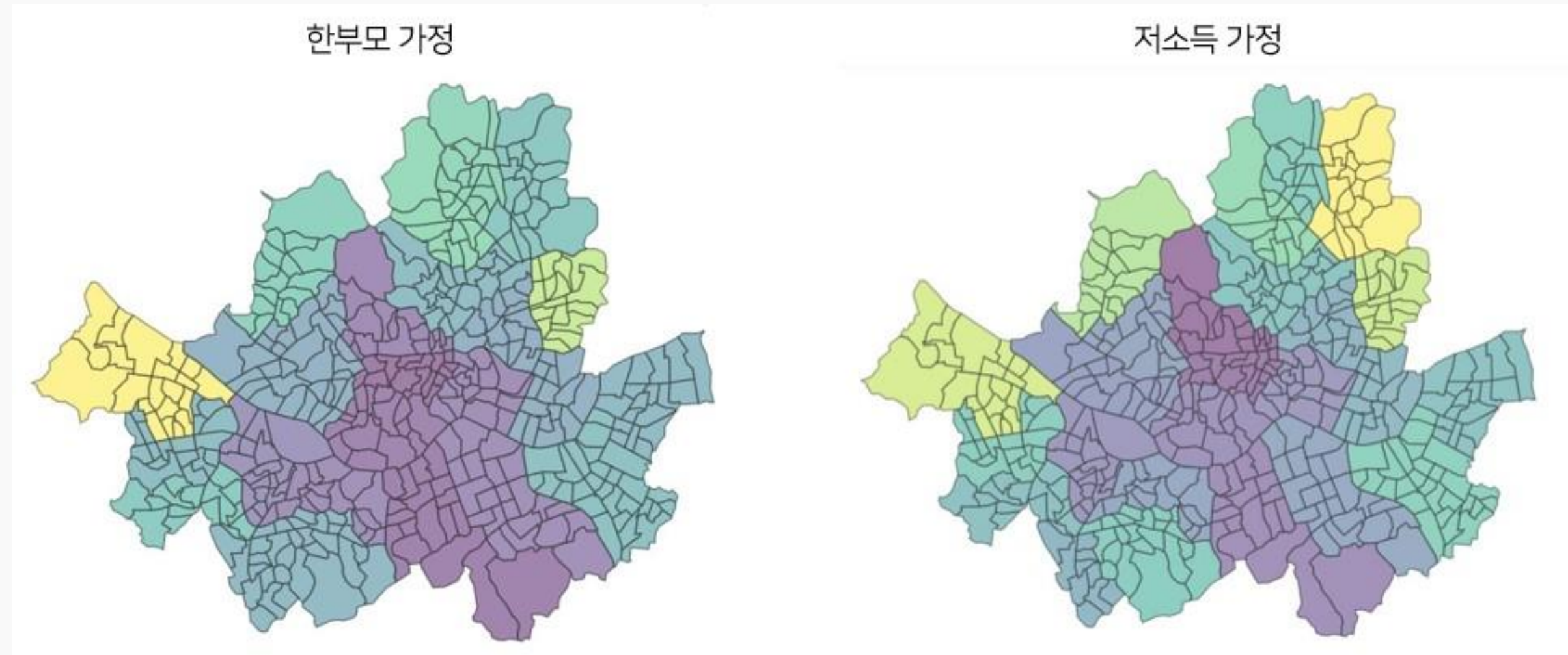
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



서울 외곽의 주거 지역, 특히 서울 북부 및 서부에서 높게 나타남

전체 변수 상관관계

서론

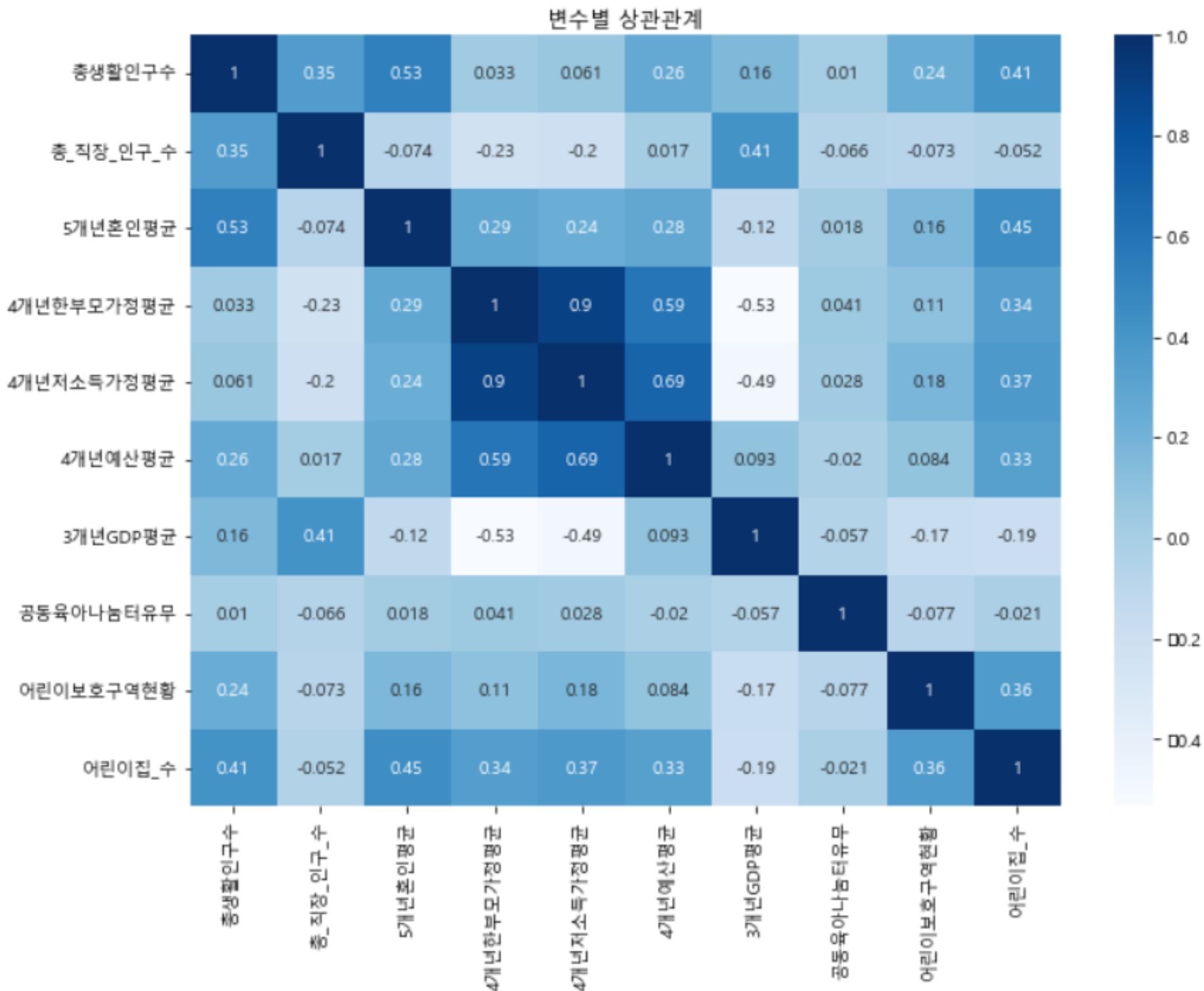
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



전체 변수 상관관계

서론

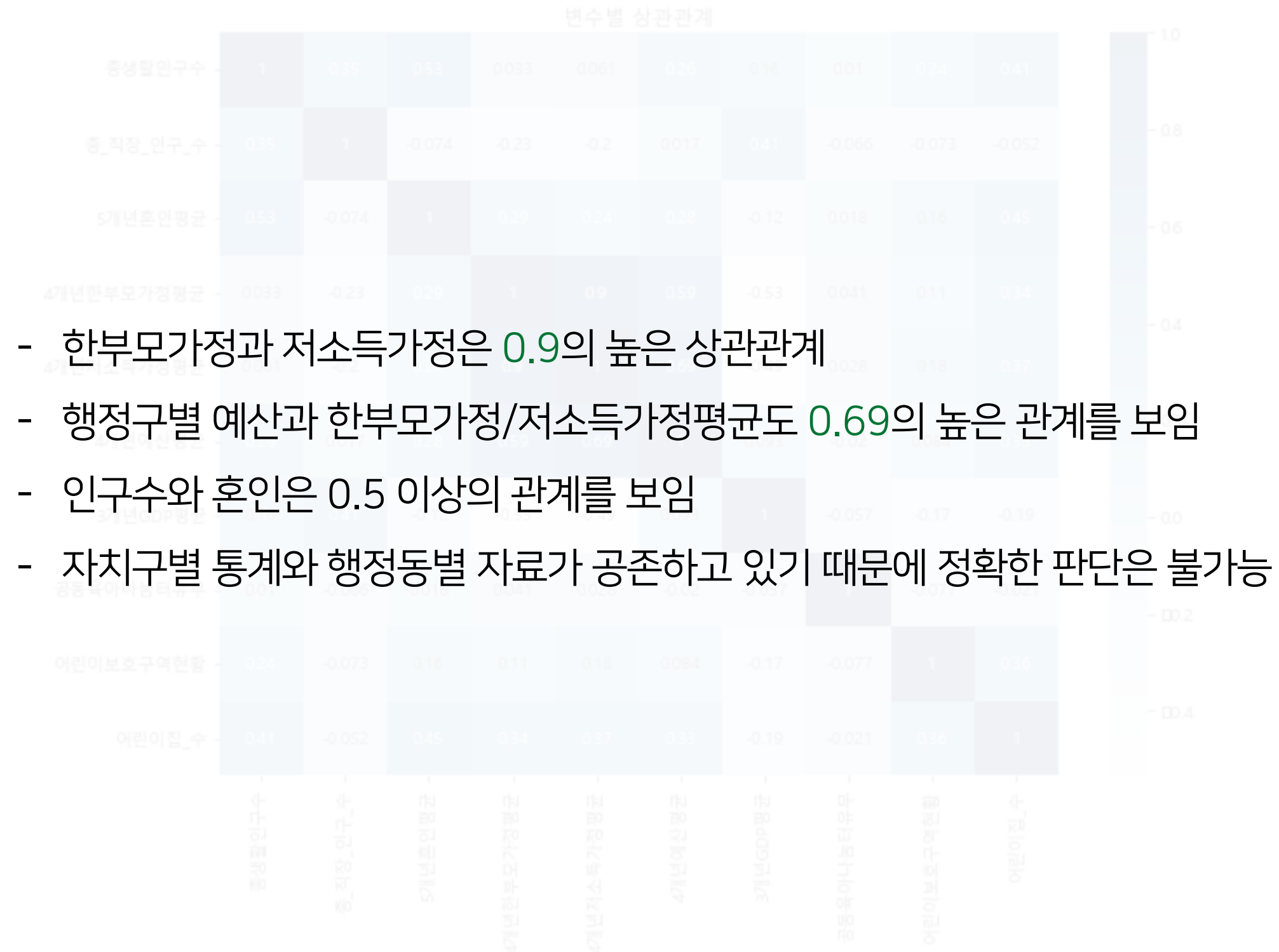
데이터별 EDA

행정동 필터링

최적 입지 선정

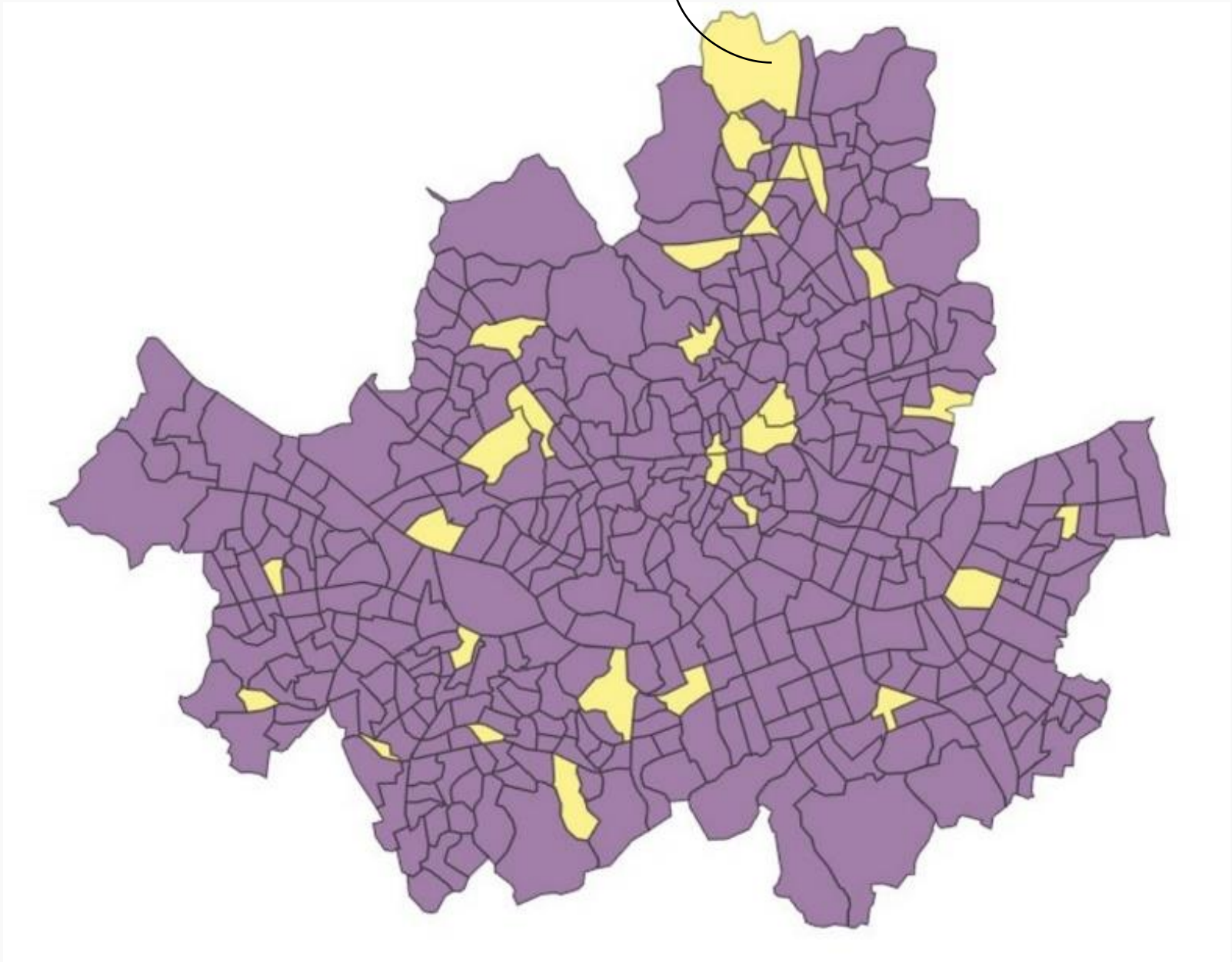
결과 분석

결론



공동육아나눔터

공동육아나눔터 有



현재 서울시 내에 설립된 공동육아나눔터는 총 30곳

- 서론
- 데이터별 EDA
- 행정동 필터링
- 최적 입지 선정
- 결과 분석
- 결론

공동육아나눔터

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

| | 2021_한부모_가구수 | 2021_저소득_가구수 | 총생활인구수 | 2021_예산현액 | 총_직장_인구_수 | 2022혼인 | 어린이보호구역현황 | 키즈카페_수 | 어린이집_수 | 아동인구 | 아동인구비 | 20-30대직장인구 |
|-------|--------------|--------------|--------------|--------------|------------|-----------|-----------|-----------|-----------|-------------|-----------|------------|
| count | 30.000000 | 30.000000 | 30.000000 | 3.000000e+01 | 30.000000 | 29.000000 | 17.000000 | 10.000000 | 30.000000 | 30.000000 | 30.000000 | 30.000000 |
| mean | 194.733333 | 1064.066667 | 24948.295846 | 1.061853e+06 | 759.558143 | 81.689655 | 3.235294 | 1.600000 | 10.766667 | 2275.559945 | 0.088421 | 504.953011 |
| std | 70.297856 | 454.762836 | 9559.198653 | 1.992704e+05 | 954.696790 | 32.601603 | 2.194579 | 1.074968 | 5.130593 | 1353.240675 | 0.035179 | 742.406549 |
| min | 72.000000 | 214.000000 | 8542.106721 | 6.534420e+05 | 109.310811 | 22.000000 | 1.000000 | 1.000000 | 2.000000 | 225.324190 | 0.022910 | 72.229730 |

공동육아나눔터가 있는 행정동 기초통계량

| | 2021_한부모_가구수 | 2021_저소득_가구수 | 총생활인구수 | 2021_예산현액 | 총_직장_인구_수 | 2022혼인 | 어린이보호구역현황 | 키즈카페_수 | 어린이집_수 | 아동인구 | 아동인구비 |
|-------|--------------|--------------|--------------|--------------|-------------|------------|------------|-----------|------------|-------------|------------|
| count | 424.000000 | 424.000000 | 424.000000 | 4.240000e+02 | 424.000000 | 422.000000 | 279.000000 | 150.00000 | 424.000000 | 424.000000 | 424.000000 |
| mean | 178.299528 | 1018.080189 | 24498.104517 | 1.073777e+06 | 1619.203171 | 84.099526 | 3.908602 | 1.56000 | 11.379717 | 2530.989266 | 0.103867 |
| std | 72.482996 | 488.041920 | 12256.917916 | 2.354672e+05 | 4132.004323 | 47.096702 | 2.239218 | 0.95172 | 7.030475 | 1662.480376 | 0.040175 |
| min | 72.000000 | 214.000000 | 2710.765852 | 6.534420e+05 | 8.181818 | 5.000000 | 1.000000 | 1.00000 | 1.000000 | 221.735909 | 0.014165 |

공동육아나눔터가 없는 행정동 기초통계량

<공동육아나눔터가 있는 행정동의 특징>

- 한부모/저소득 가구 수가 더 많음 = 복지제도적 특징이 드러남
- 직장인구 수와 아동 인구수가 비교적 낮게 나타남

-> 현존하는 공동육아나눔터의 위치가 인구 분포를 고려하지 못한 배치

공동육아나눔터

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

| | 2021_한부모_가구수 | 2021_저소득_가구수 | 총생활인구수 | 2021_예산현액 | 총_직장_인구_수 | 2022혼인 | 어린이보호구역현황 | 키즈카페_수 | 어린이집_수 | 아동인구 | 아동인구비 | 20-30대직장인구 |
|-------|--------------|--------------|--------------|--------------|------------|-----------|-----------|-----------|-----------|-------------|-----------|------------|
| count | 30.000000 | 30.000000 | 30.000000 | 3.000000e+01 | 30.000000 | 29.000000 | 17.000000 | 10.000000 | 30.000000 | 30.000000 | 30.000000 | 30.000000 |
| mean | 194.733333 | 1064.066667 | 24948.295846 | 1.061853e+06 | 759.558143 | 81.689655 | 3.235294 | 1.600000 | 10.766667 | 2275.559945 | 0.088421 | 504.953011 |
| std | 70.297856 | 454.762836 | 9559.198653 | 1.992704e+05 | 954.696790 | 32.601603 | 2.194579 | 1.074968 | 5.130593 | 1353.240675 | 0.035179 | 742.406549 |
| min | 72.000000 | 214.000000 | 8542.106721 | 6.534420e+05 | 109.310811 | 22.000000 | 1.000000 | 1.000000 | 2.000000 | 225.324190 | 0.022910 | 72.229730 |

공동육아나눔터가 있는 행정동 기초통계량

| | 2021_한부모_가구수 | 2021_저소득_가구수 | 총생활인구수 | 2021_예산현액 | 총_직장_인구_수 | 2022혼인 | 어린이보호구역현황 | 키즈카페_수 | 어린이집_수 | 아동인구 | 아동인구비 |
|-------|--------------|--------------|--------------|--------------|-------------|------------|------------|-----------|------------|-------------|------------|
| count | 424.000000 | 424.000000 | 424.000000 | 4.240000e+02 | 424.000000 | 422.000000 | 279.000000 | 150.00000 | 424.000000 | 424.000000 | 424.000000 |
| mean | 178.299528 | 1018.080189 | 24498.104517 | 1.073777e+06 | 1619.203171 | 84.099526 | 3.908602 | 1.56000 | 11.379717 | 2530.989266 | 0.103867 |
| std | 72.482996 | 488.041920 | 12256.917916 | 2.354672e+05 | 4132.004323 | 47.096702 | 2.239218 | 0.95172 | 7.030475 | 1662.480376 | 0.040175 |
| min | 72.000000 | 214.000000 | 2710.765852 | 6.534420e+05 | 8.181818 | 5.000000 | 1.000000 | 1.00000 | 1.000000 | 221.735909 | 0.014165 |

공동육아나눔터가 없는 행정동 기초통계량

<공동육아나눔터가 있는 행정동의 특징>



종합해 보자면 ...
서울시 서부 외곽 지역과 서초구 및 강남구 부근의 추가적인 설립 필요!

-> 현존하는 공동육아나눔터의 위치가 인구 분포를 고려하지 못한 배치

클러스터링 변수

서론

데이터별 EDA

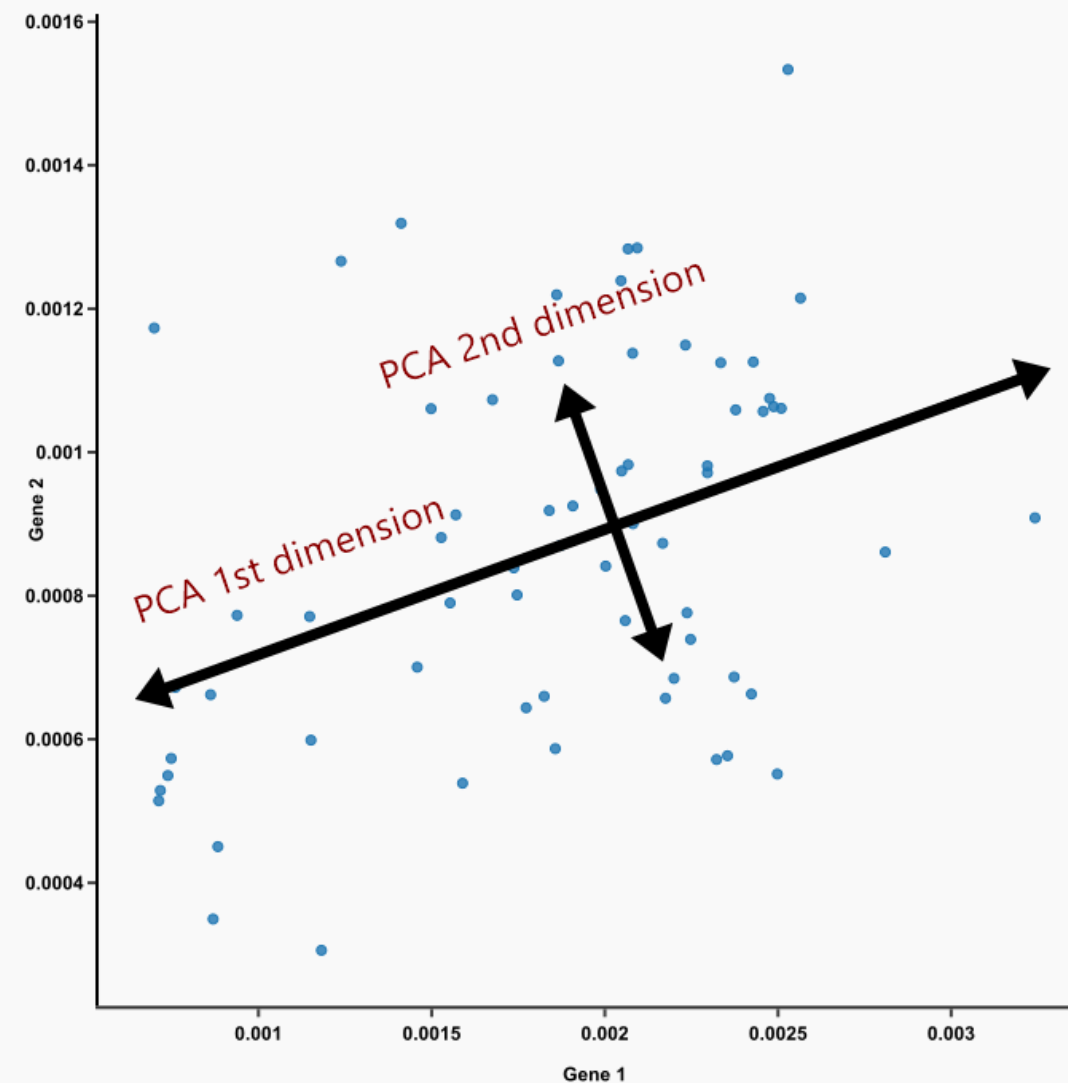
행정동 필터링

최적 입지 선정

결과 분석

결론

행정동 필터링을 위한 클러스터링을 진행하기 앞서, 클러스터링 변수를 생성
주성분 분석(PCA)로 전체 변수를 인구 지수, 예산/복지 지수로 구성된 2차원으로 축소



주성분 분석 (PCA)

변수 간 상관관계를 이용해
기존의 분산을 최대한 보존하게끔
'주성분'을 찾는 차원축소 기법

클러스터링 변수

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

| | |
|-----------------|------|
| 총생활인구수 | 0.28 |
| 남자0세부터9세생활인구수 | 0.26 |
| 남자10세부터14세생활인구수 | 0.30 |
| 남자30세부터34세생활인구수 | 0.16 |
| 남자35세부터39세생활인구수 | 0.19 |
| 남자40세부터44세생활인구수 | 0.22 |
| 남자45세부터49세생활인구수 | 0.25 |
| 여자0세부터9세생활인구수 | 0.30 |
| 여자10세부터14세생활인구수 | 0.31 |
| 여자25세부터29세생활인구수 | 0.16 |
| 여자30세부터34세생활인구수 | 0.18 |
| 여자35세부터39세생활인구수 | 0.24 |
| 여자40세부터44세생활인구수 | 0.32 |
| 총_직장_인구_수 | 0.07 |
| 연령대_20_직장_인구_수 | 0.07 |
| 연령대_30_직장_인구_수 | 0.07 |
| 연령대_40_직장_인구_수 | 0.07 |
| 어린이집_수 | 0.22 |
| 유아인구 | 0.33 |

PC1

인구 관련 변수의 PCA loadings

1. 인구지수

- 생활인구, 직장인구, 아동인구, 어린이집 수 데이터 등 활용
- 아동 관련 변수의 적재량이 높게 나타나, 본 연구에 부합

2. 예산/복지지수

- 한부모/저소득 데이터, 예산 데이터 활용
- 자치구 통계이므로, 자치구에 속하는 모든 행정동은 동일한 값을 가짐

PCA 결과에서, 가장 분산이 큰 축인 PC1의 적재량(loading)의 절댓값을
각 변수의 가중치로 활용해서 지수 생성

- 적재량 (loading): 각 변수가 PC1으로 선형 변환될 때의 계수 값으로,
각 변수가 해당 PC에 미치는 영향을 가리킴

클러스터링 변수

- 서론
- 데이터별 EDA
- 행정동 필터링
- 최적 입지 선정
- 결과 분석
- 결론

| | 시군구명 | 행정동명 | 인구지수 | 예산복지지수 | 공동육아나눔터유무 |
|-----|------|------|--------------|------------|-----------|
| 0 | 종로구 | 사직동 | 10613.930349 | 882439.63 | 0 |
| 1 | 종로구 | 삼청동 | 2075.538390 | 882439.63 | 0 |
| 2 | 종로구 | 부암동 | 6224.970589 | 882439.63 | 0 |
| 3 | 종로구 | 평창동 | 7591.582598 | 882439.63 | 0 |
| 4 | 종로구 | 무악동 | 2795.206491 | 882439.63 | 0 |
| ... | ... | ... | ... | ... | ... |
| 419 | 강동구 | 둔촌1동 | 2053.691100 | 1561745.44 | 0 |
| 420 | 강동구 | 둔촌2동 | 11843.518743 | 1561745.44 | 0 |
| 421 | 강동구 | 암사1동 | 11999.031127 | 1561745.44 | 0 |
| 422 | 강동구 | 천호2동 | 18400.280368 | 1561745.44 | 0 |
| 423 | 강동구 | 길동 | 23028.800269 | 1561745.44 | 0 |

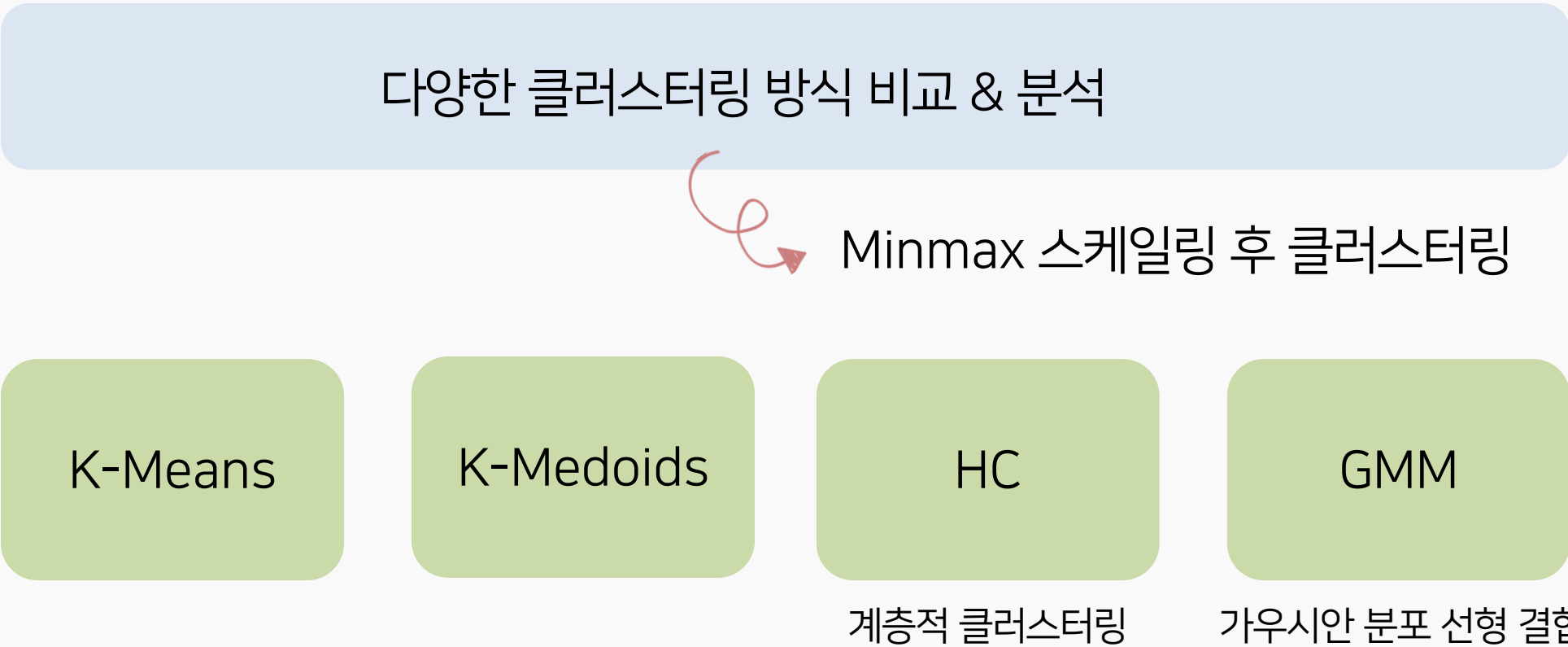
424 rows × 5 columns

최종 클러스터링용 데이터셋 생성

클러스터링



- 서론
- 데이터별 EDA
- 행정동 필터링
- 최적 입지 선정
- 결과 분석
- 결론



클러스터링

서론

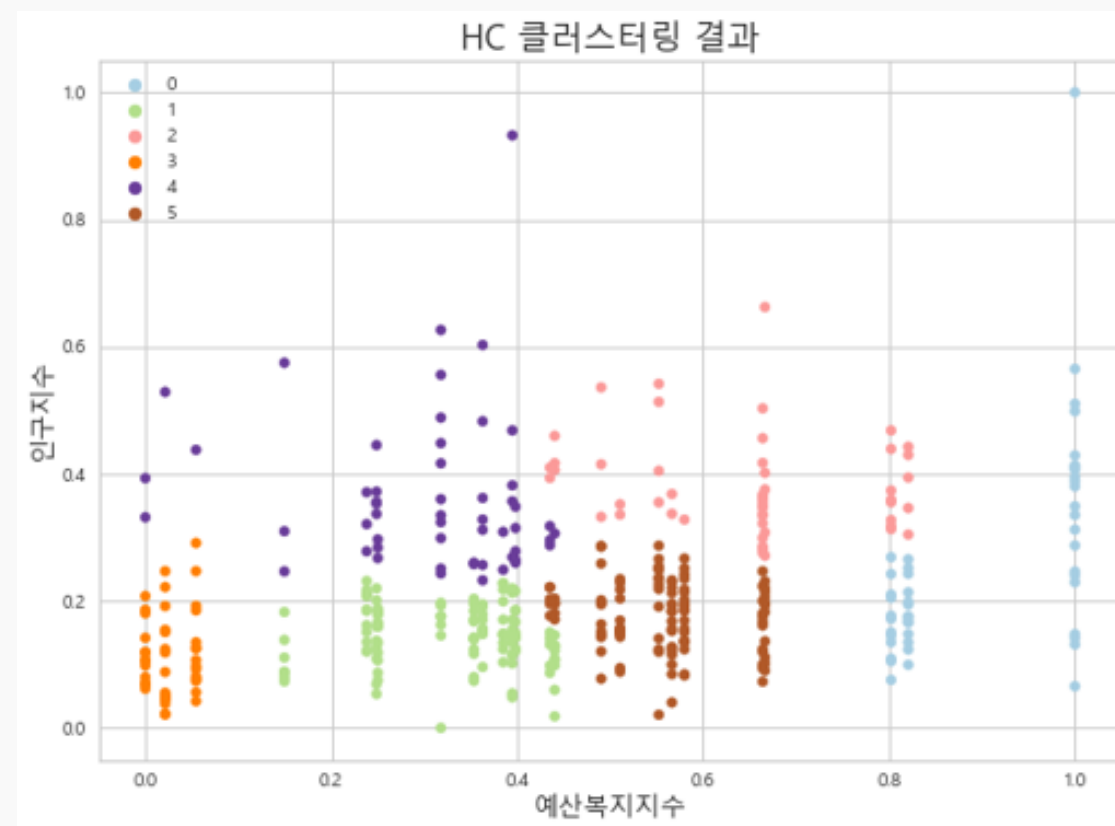
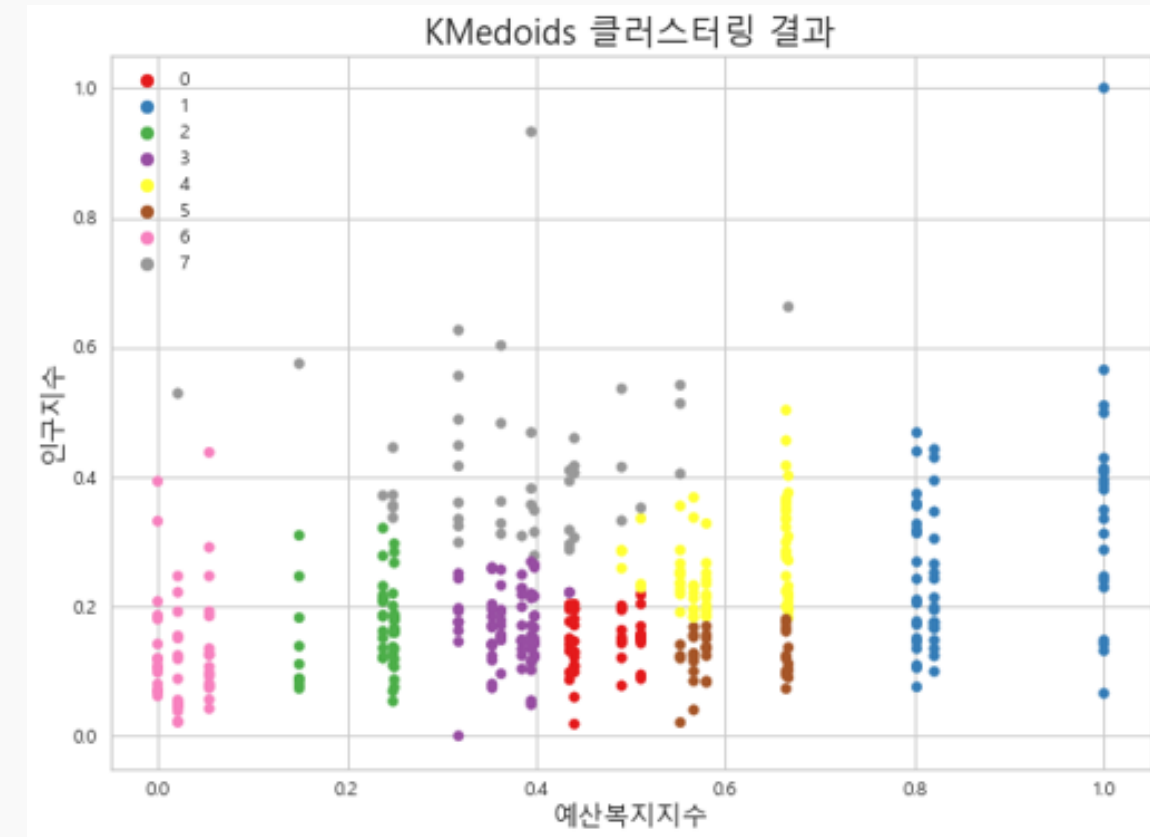
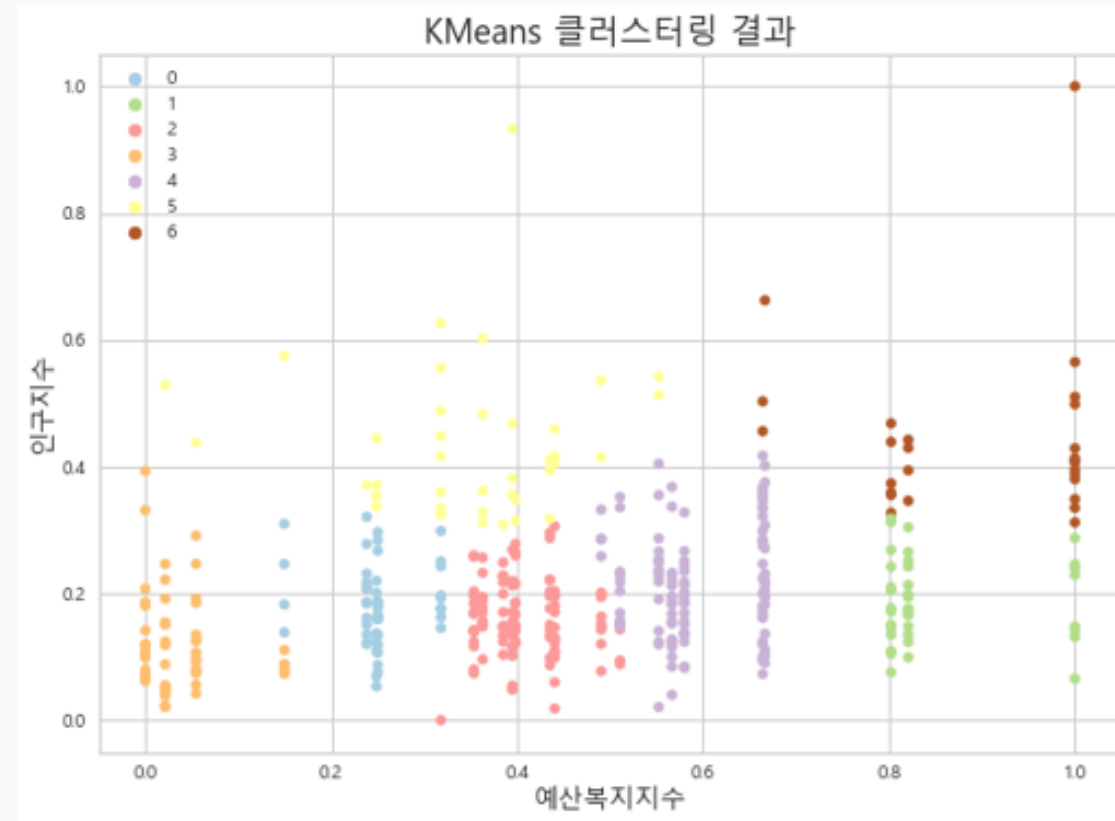
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



군집화가 잘 이뤄진 K-Means, K-Medoids, HC의
결과를 종합하여 행정동 필터링 ✨

클러스터링

서론

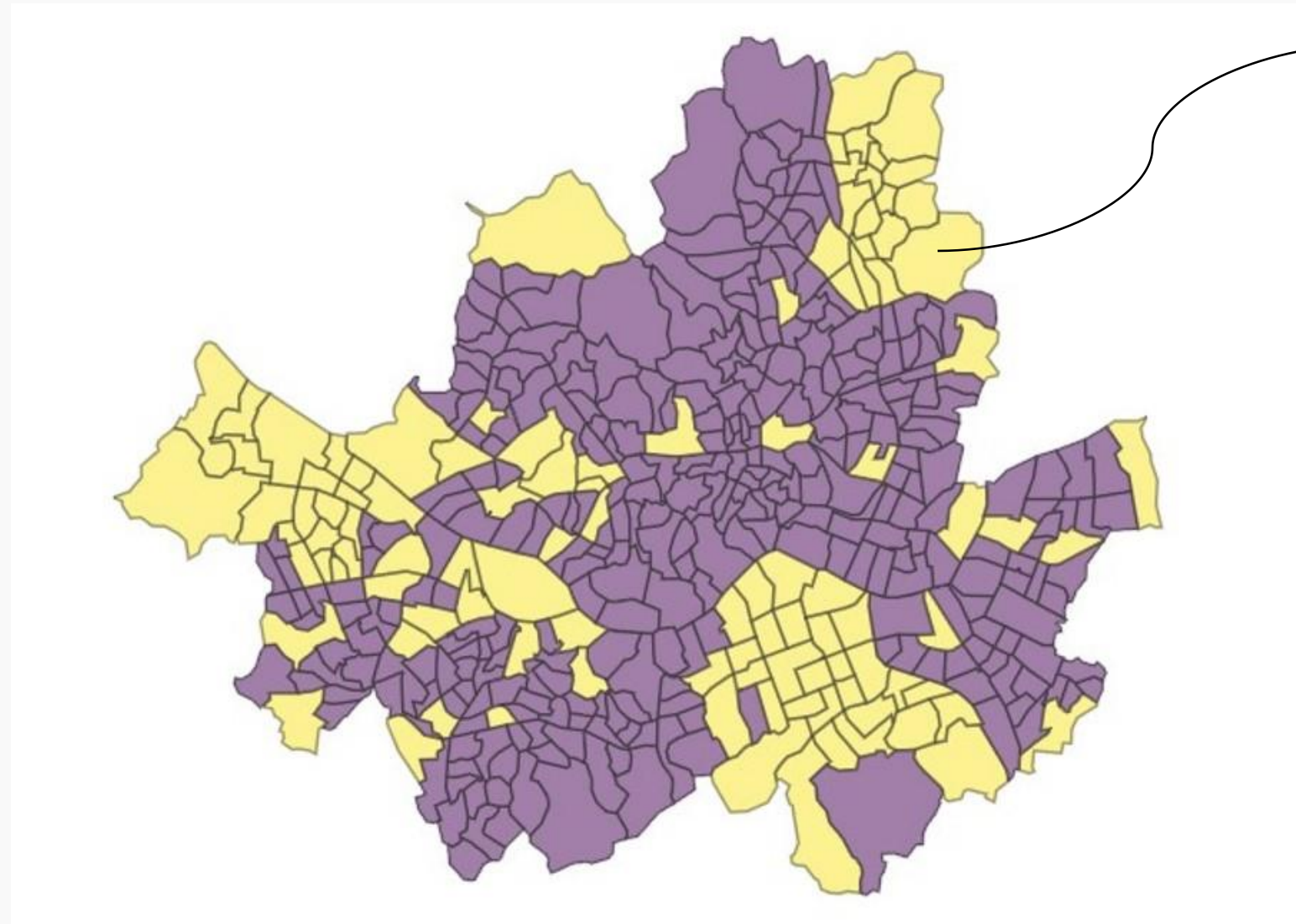
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



필터링 0

클러스터링을 통해 필터링된 행정동

각 클러스터링 결과에서 인구 지수와 예산/복지 지수가 높은 상위 2개의 군집을 선택



3가지의 클러스터링 중 2번 이상 선택된 행정동만 필터링하여, 최종 행정동 105개 추출

입지 선정 개요

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

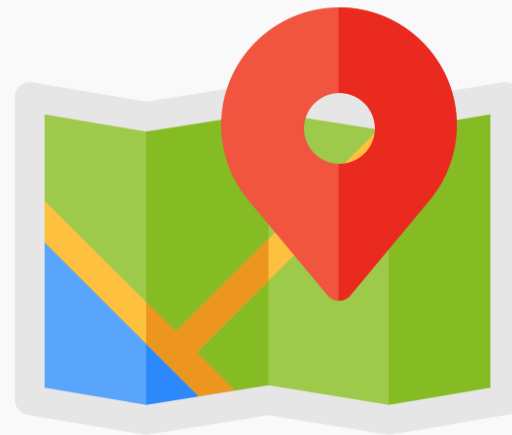
결과 분석

결론

Set Cover

조건 또는 수요를 만족하기 위한 최소의 집합을 찾는 것을 목표
즉, 특정 조건을 만족하는 범위 내에서 시설물의 개수를 최소화하는 것을 목적으로 동작

* LSCP, MCLP 는 Set Cover 알고리즘의 일종



수요: 반경 2.5km 이내의 아동 인구

최소한의 공동육아나눔터 추가 설치를 통해, 최대의 수요(아동 인구)를 충족하자

입지 선정 개요

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

LSCP

Location Set Covering Problem

특정한 지리적 범위 내의 모든 수요를 충족할 수 있는 최소의 시설물의 개수를 찾는 알고리즘

MCLP

Maximal Covering Location Problem

사전에 정해진 시설물의 개수를 기반으로 가능한 모든 수요를 커버할 수 있는 입지를 결정하는 알고리즘

입지 선정 개요

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

수요 계산을 위해 각 행정동의 위도와 경도를 활용해
행정동 간의 거리를 일대일 대응으로 계산



각 행정동의 중심점을 기준으로
반경 2.5km 이내의 유아인구를 수요로 계산!



가능한 모든 수요를 만족하면서 시설물(공동육아나눔터) 개수가 최소가 되도록 설계



- 전체 행정동을 기준으로 수요를 계산한 결과 > LSCP
- 필터링한 행정동(105개)만을 기준으로 수요를 계산한 결과 > LSCP & MCLP

1. 전체 행정동

#LSCP

* 수요(반경 2.5km 이내 아동 인구) 계산 시, 전체 행정동의 아동 인구를 기준으로 합산

설치 위치:
남은 아동수: 1073139

⋮

설치 위치: 248 404 174 348 139 313 221 423 90 364 202 406 261 358 407 256 110 271 8 353 160 189 263 236 288
252 167 378 161 211 352
남은 유아수: 195676.2

32개 행정동에 공동육아나눔터 추가 설치한다면,
서울시 전체 아동 인구의 최대 80% 정도를 충족 가능!



1. 전체 행정동

#LSCP

서론

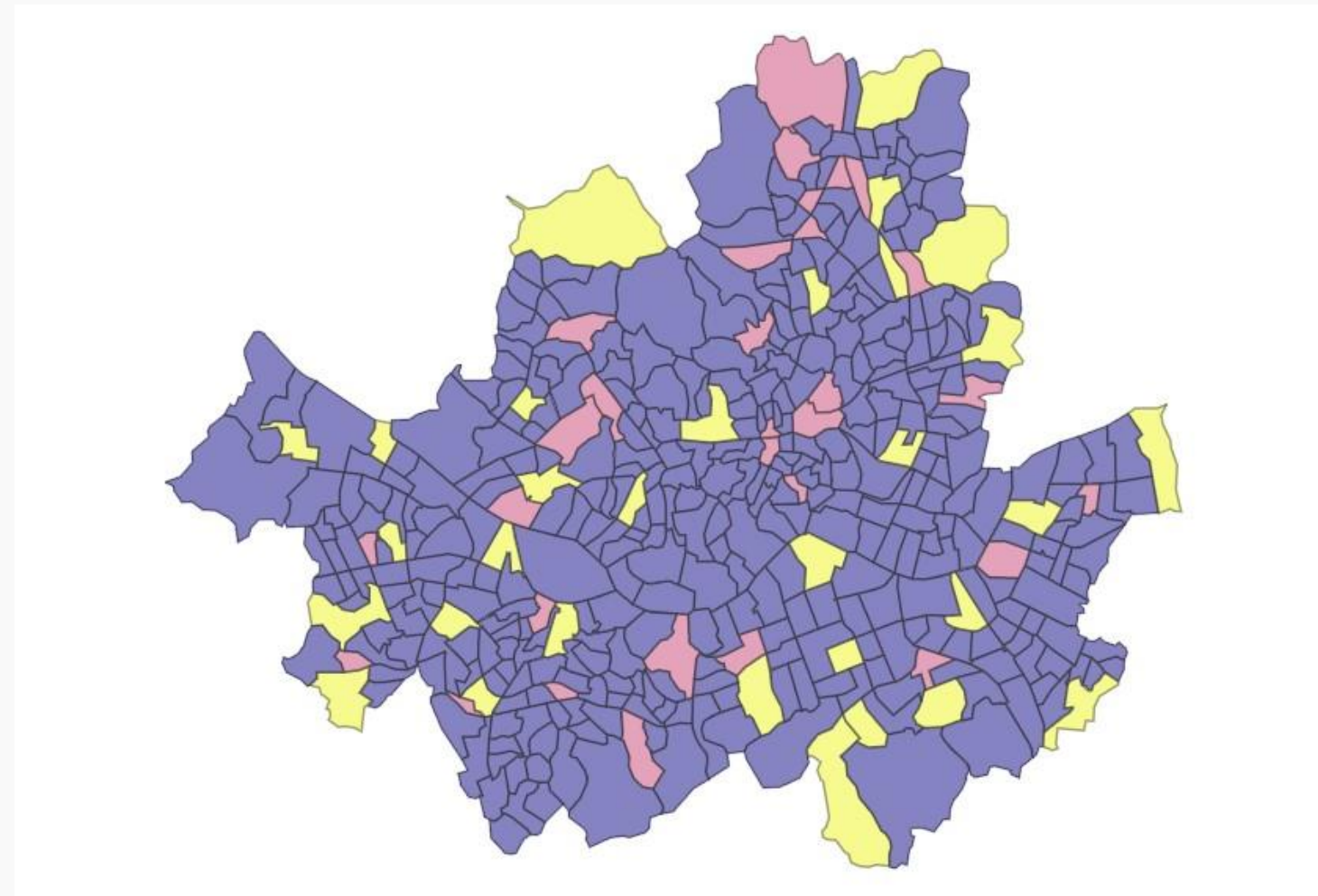
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



■ LSCP를 통해 선택된 입지 선정지
■ 현재 공동육아나눔터

- 서울 외곽의 주거지역
- 한부모/저소득 가정이 많았던 서부와 북부 지역
- 20-40대 직장인구가 많은 마포구, 용산구, 서초구, 강남구 등

목적에 알맞은 결과

2. 필터링 행정동

#LSCP

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

* 수요(반경 2.5km 이내 아동 인구) 계산 시, **필터링된 행정동**만의 아동 인구를 기준으로 합산

설치 위치 :
남은 아동수 : 419744

⋮

설치 위치 : 79 47 11 53 97 66 30 104 85 34 20 102 99 88 26 103 62 61 89 68 5 56 6 4 3 1 52
남은 유아수 : 0

28개 행정동에 공동육아나눔터 추가 설치한다면,
필터링된 행정동 전체의 아동 인구 충족 가능!



2. 필터링 행정동

#LSCP

서론

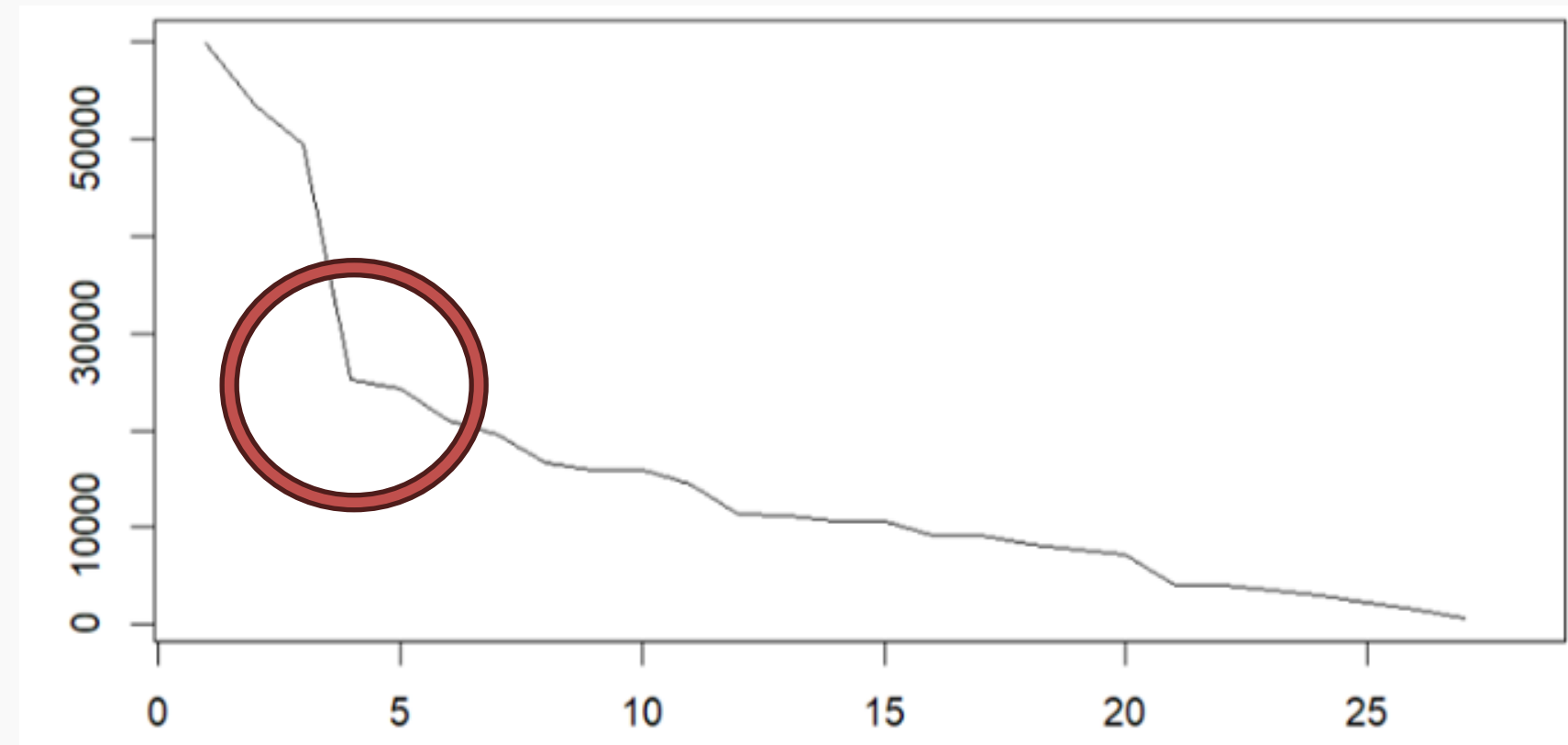
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



시설물 설치 개수에 따른 미충족 수요 그래프

Elbow point인 4를 최적의 공동육아나눔터 설치 개수로 선정

▶ 시설물 개수 = 4 인 MCLP 적용 !

2. 필터링 행정동

#LSCP

서론

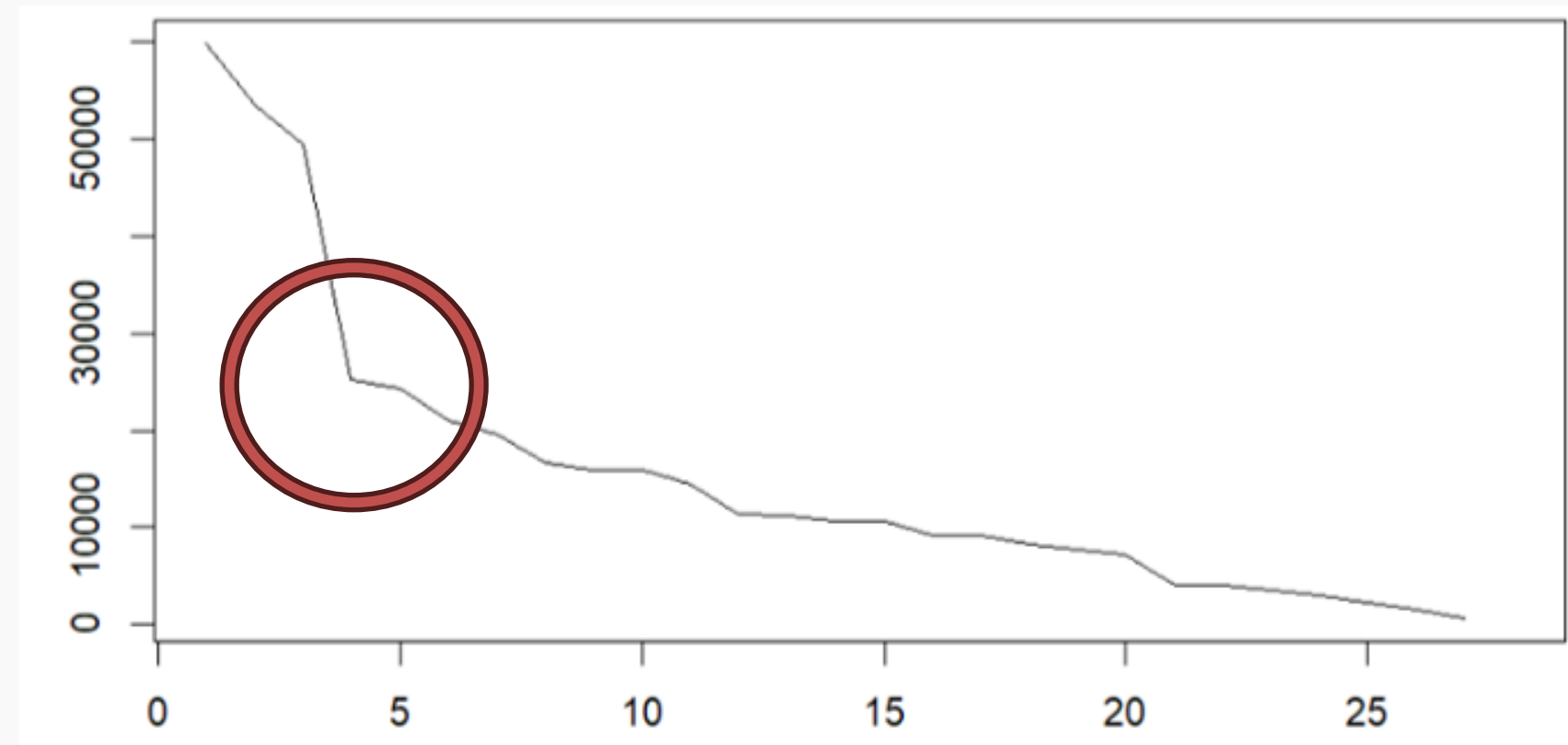
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



시설물 설치 개수에 따른 미충족 수요 그래프



프로젝트의 목적이 '최적 입지선정'이므로
최소한의 비용(공동육아나눔터 설치) 대비 가장 큰 효과(아동 수요 만족)를 얻기 위해
커버 가능한 유아 인구가 급격히 감소하는 elbow point 지점으로 개수 선정

2. 필터링 행정동

#MCLP

서론

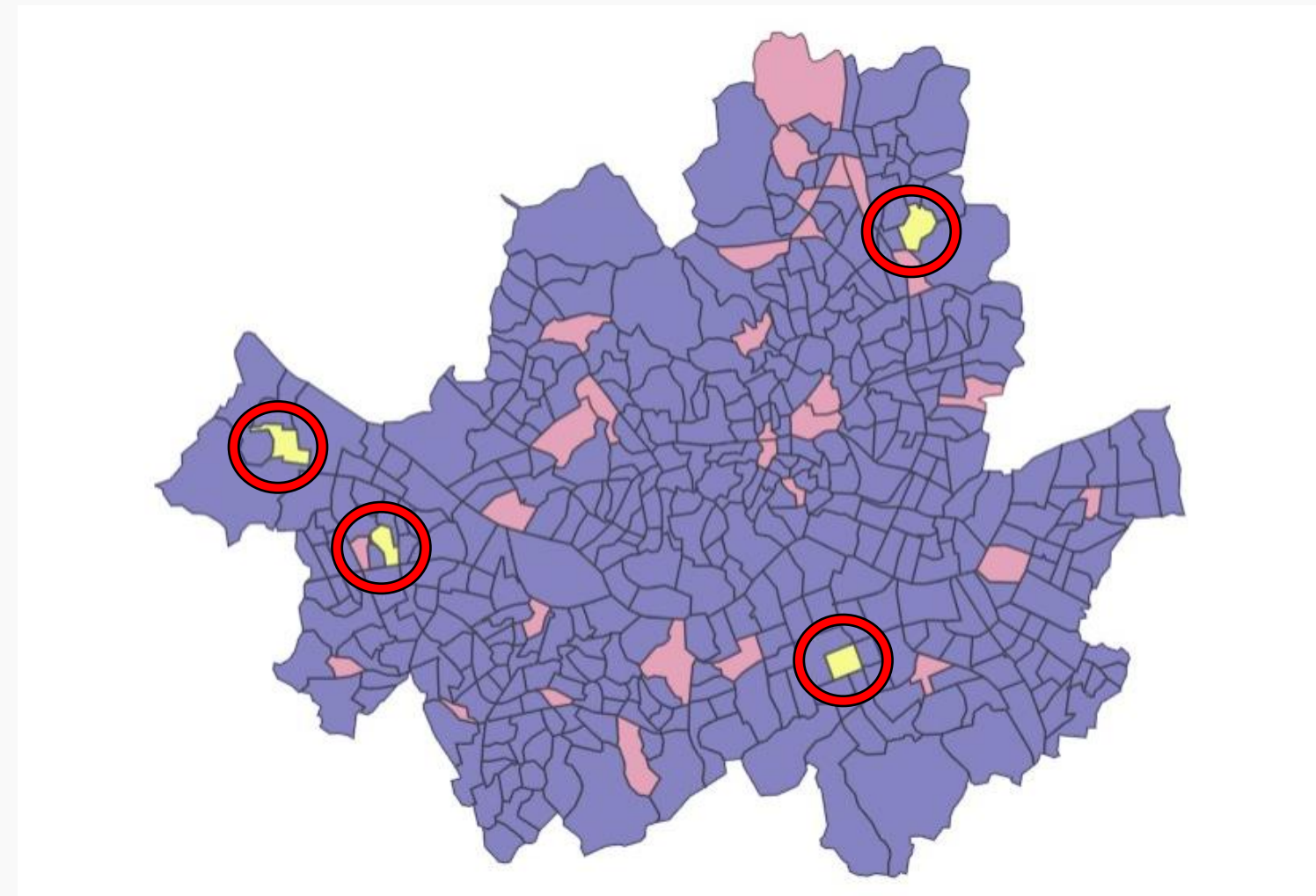
데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론



■ MCLP를 통해선택된 입지 선정지
■ 현재 공동육아나눔터

강남구 역삼2동, 강서구 화곡 4동, 노원구 하계1동, 강서구 방화1동

- 한부모/저소득 가정의 수가 높았던 강서구 및 노원구
- 아동 인구와 20-40대 직장인구가 많은 강남구

복지 제도적 성격과 인구 분포를 모두 고려한 최적의 입지 선정!



서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

| | 유아인구 | 총생활인구수 | 20-40생활인구 | 총_직장_인구_수 | 20-40직장인구 | 2021_한부모_가구수 | 2021_저소득_가구수 | 2021_예산현액 |
|-------|-------------|--------------|--------------|-------------|-------------|--------------|--------------|--------------|
| count | 4.000000 | 4.000000 | 4.000000 | 4.000000 | 4.000000 | 4.0000 | 4.000000 | 4.000000e+00 |
| mean | 4207.438989 | 32294.473467 | 10918.020115 | 3273.958065 | 2505.845161 | 257.0000 | 1599.750000 | 1.445694e+06 |
| std | 2142.446084 | 15056.972683 | 7257.499735 | 6047.445959 | 4725.126032 | 122.3574 | 578.287918 | 8.511471e+04 |
| min | 2344.242265 | 17377.772490 | 5423.830848 | 125.580645 | 54.258065 | 102.0000 | 745.000000 | 1.392518e+06 |

최종 입지 행정동 기초 통계량

| | 유아인구 | 총생활인구수 | 20-40생활인구 | 총_직장_인구_수 | 20-40직장인구 | 2021_한부모_가구수 | 2021_저소득_가구수 | 2021_예산현액 |
|-------|-------------|--------------|-------------|-------------|-------------|--------------|--------------|--------------|
| count | 420.000000 | 420.000000 | 420.000000 | 420.000000 | 420.000000 | 420.00000 | 420.000000 | 4.200000e+02 |
| mean | 2515.023078 | 24423.853384 | 8236.357244 | 1601.804963 | 1174.341641 | 177.55000 | 1012.540476 | 1.070235e+06 |
| std | 1652.354592 | 12225.268763 | 5366.690025 | 4115.679208 | 3297.174251 | 71.67325 | 484.565402 | 2.336432e+05 |
| min | 221.735909 | 2710.765852 | 911.888202 | 8.181818 | 3.000000 | 72.00000 | 214.000000 | 6.534420e+05 |

비입지 행정동 기초 통계량

모든 변수에서 최종 입지 행정동의 수치가 더 높게 나타남

- 직장인구 수 2배 이상 높게 나타남
- 아동 인구 약 1.7배 높게 나타남

공동육아나눔터 최적 입지 선정지로서 적절!

연구 요약

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

EDA

- 아동 인구는 서울 외곽의 주거 지역에서 높게 나타남
 - 한부모/저소득 가정은 서울 북부 및 서부에서 높게 나타남
 - 20-40 직장인구는 한강 부근에서 높게 나타남
 - 현존 공동육아나눔터는 복지 제도적 성격이 강함
 - 현존 공동육아나눔터는 인구 분포를 고려하지 않은 배치로 보임
- > 공동육아나눔터의 복지적인 성격을 유지하면서도, 인구 분포를 고려한 설치가 필요!

클러스터링

- 서울시 424개 행정동 대상 1차 필터링 진행
- 인구 지수, 예산/복지 지수로 클러스터링 변수 생성
- 다양한 클러스터링 알고리즘을 적용해, 최종 105개의 행정동 필터링

입지분석

- LSCP & MCLP 알고리즘을 적용해 입지 분석 진행
 - 전체 행정동 기반 수요 계산 결과) 서울 외곽, 북부 및 서부 일대, 한강 부근의 32개 행정동
 - 필터링 행정동 기반 수요 계산 결과) 강남구 역삼2동, 강서구 화곡 4동, 노원구 하계 1동, 강서구 방화1동
- > 복지 제도적 역할과 인구 분포에 따른 수요 모두 충족할 수 있는 결과

서론

데이터별 EDA

행정동 필터링

최적 입지 선정

결과 분석

결론

1. 의의

- 현존하는 공동육아나눔터의 문제점을 파악해, 새로운 기준 제시
- 각 자치구 및 행정동의 특성을 고려하여, 논리적으로 설득력 있는 공동육아나눔터 입지를 제시

2. 한계

- 실제로 공동육아나눔터를 신설하기 위한, 행정동 내의 최적의 건물 혹은 부지를 구체적으로 제시하지 못함

감사합니다 😊