

• • • •  
Data Science Lab

# 우리의 삶에 필요한 도시환경은 무엇일까 ?

10기 박성원 11기 양주원 한은결 12기 김건우



Table of Contents

# 목차보기

도시환경 EDA



01

## 주제 선정 이유

도시환경이란 무엇이며, 왜 중요한가

02

## 데이터 소개

데이터의 출처 및 전처리 결과.

03

## 분석 결과

공원, 교육시설, 문화시설, 의료시설, 종합

04

## 결론 및 한계

최종 인사이트

# 주제 선정 이유

왜 도시환경에 집중해야 하는가



viva 브릿지경제 · 2일 전

## 구리시, 시민 만족도 높일 수 있는 도시 환경 사업 추진

구리시, 시민 만족도 높일 수 있는 도시 환경 사업 추진 구리시는 8월 1일 오전 10시 시청 본관 3층 종합상황실에서 2024년 8월 1주 차 정례 기자브리핑을 열었다. 이날 브리핑은 김천복 안전도시국장이 ▲여름철 자연 재...

시사포커스 · 2주 전

## 예천군, '출산 친화' 도시 환경 조성

예천 아이사랑안심케어센터 전경©예천군 경북 예천군이 출산 친화적인 도시환경 조성에 속도 내고 있다. 16일 예천군에 따르면 군은 '24시 시간제' 어린이집을 지난 1일부터 2개소 운영하고 있으며 다음 달 1일부터는...

강원일보 · 15면 1단 · 3일 전 · 네이버뉴스

## 영월군, 아이키우기 좋은 보육환경 도시 조성

영월군이 가족센터 개관에 이어 공공산후조리원 등을 완공하며 아이키우기 좋은 보육환경도시 조성에 나섰다. 영월군가족센터 개관식이 지난 31일 여월읍 우행나무3길 16-15에서 열렸다. 군은 출산과 보육이 지역에...



신문 · 6일 전

## 울산 도시환경 분과위원회 개최

울산시는 지난 26일 울산시청 국제회의실에서 이상식 위원장 주재로 도시환경분과 위원 및 관련부서 관계자들이 참석한 가운데 주민제안사업 심의를 위한 '주민참여예산 도시환경 분과위원회'를 가졌다. 울산시 제공





# 주제 선정 이유

왜 도시환경에 집중해야 하는가



뉴욕 센트럴파크 설계자

“지금 이 곳에 공원을 만들지 않는다면 100년 후에는  
이만한 크기의 **정신병원**이 필요할 것”


뉴욕의 센트럴파크, 런던의 하이드파크 등

세계 여러 주요 도시에는 그를 상징하는 공원이 존재하지만, 우리나라의 경우  
공원을 비롯한 여러 도시환경이 조성되어 있지 않은 편에 속한다.

# 주제 선정 이유

왜 도시환경에 집중해야 하는가




 한겨레

## 서울 집값 비싼 동네에 병원 70%·의사 2배 더 많다

서울 고가주택 밀집지역의 인구 대비 의료기관 수가 저가주택 지역보다 70% 이상 많은 것으로 나타났다. 공공병원을 세울 때 저가주택 지역을 중심...

2024. 1. 8.



 한국조경신문

## [집중점검] 도시공원일몰제 대응 '광역시' 예산 빈부격차 심각

[Landscape Times 지재호 기자] 도시공원일몰제 시행이 앞으로 550여일 남겨두고 있는 시점에서 지자체들은 그야말로 사면초가에 빠져 있는 듯 하다.

2018. 12. 21.



도시환경의 중요성은 높아져가지만,

그에 따라 지역 간 도시환경 격차는 더 커지고 있다

-> 어떠한 도시환경이 중요하며, 격차 해소를 위해 어떤 정책이 필요한지 검토할 시점

# 데이터 소개

녹지환경	의료 시설	교육 환경	문화 시설
1) 국토 이용 현황 by KOSIS	1) 의료기관별 상세정보 서비스 by 보건의료빅데이터개방시스템	1) 학원교습소정보 by 나이스 교육정보 개방포털	1) 지역문화재단 by 문화빅데이터플랫폼
2) 어린이 공원 현황 by KOSIS		2) 고등교육기관 데이터 by KOSIS	2) 박물관 by 문화빅데이터플랫폼
3) 전국도시공원정보 표준데이터 by 공공데이터포털		3) 유.초.중등 교육기관 by 교육통계서비스	3) 미술관 by 문화빅데이터플랫폼
			4) 공공도서관 by 문화빅데이터플랫폼

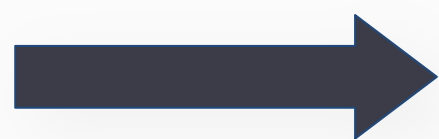
# 전처리 과정

## 1. 시도, 행정구역 기준으로 정리

ex) 의료 데이터

	요양기관명	종별코드명	설립구분코드명	개설일자	시군구 명칭	개설연도	응급실병상수	주소	longitude	latitude
0	강원대학교병원	종합병원	공립	19820708	강원도 춘천시	1982	30	강원도 춘천시 백령로 156, (효자동)	127.744926	37.875088
1	강동종로약국	약국	개인	20130506	강원도 강릉시	2013	0	강원특별자치도 강릉시 강동면 와천로 470, (강동면)	128.954460	37.727142
2	강동보건지소	보건지소	공립	19830329	강원도 강릉시	1983	0	강원특별자치도 강릉시 강동면 울곡로 2208, (강동면)	128.954222	37.728088
3	임곡보건진료소	보건진료소	공립	19821001	강원도 강릉시	1982	0	강원특별자치도 강릉시 강동면 임곡로 738, (강동면)	128.973049	37.688266
4	정동보건진료소	보건진료소	공립	19840201	강원도 강릉시	1984	0	강원특별자치도 강릉시 강동면 헌화로 1054, (강동면)	129.034062	37.688236

68383x10 columns



df\_medical\_sortcode\_2023

종별코드명	시군구 명칭	병원	보건소	보건의료원	보건지소	보건진료소	상급종합	약국	요양병원	의원	정신병원	조산원	종합병원	치과병원	치과의원	한방병원	한의원
0	강원도 강릉시	1	1	0	7	9	1	93	2	121	2	0	3	1	65	1	60
1	강원도 고성군	0	1	0	3	9	0	9	0	6	0	0	0	0	5	0	6
2	강원도 동해시	1	1	0	1	0	0	39	3	49	0	0	2	0	27	0	19
3	강원도 삼척시	1	1	0	8	6	0	24	1	23	0	0	1	0	13	0	10



# 전처리 과정

## 2. 스케일링

ex) 고등교육기관 데이터

	연도	학교종류	시도	행정구	학교명	학교명(영문)	본분교	학교상태	설립
0	2023	대학교	강원	강릉시	강릉원주대학교	Gangneung-Wonju National University	본교(제1캠퍼스)	기존	국립
1	2023	대학교	강원	원주시	강릉원주대학교	Gangneung-Wonju National University	본교(제2캠퍼스)	기존	국립
2	2023	대학교	강원	춘천시	강원대학교	Kangwon National University	본교(제1캠퍼스)	기존	국립



```
# 연령대별 기관수
df1_age['청년인구당_기관수'] = df1_age['학교수']/df1_age['고등교육기관_연령대_수']

# 연령대별 고등교육기관의 수 스케일링
# MinMaxScaler 객체 생성
scaler2 = MinMaxScaler()

# 데이터를 0과 1 사이로 스케일링
df1_age['청년인구당_기관수_스케일링'] = scaler2.fit_transform(df1_age[['청년인구당_기관수']])

# 결과 출력
df1_age.head(3)
```

✓ 0.0s

	시도	행정구	총인구수	학교수	고등교육기관_연령대_수	초/중/고_연령대_수	인당_기관수	인당_기관수_스케일링
0	서울	종로구	139189.0	10.0	11195	11308	0.000072	1.000000
1	서울	중구	121475.0	3.0	8040	7650	0.000025	0.343747
2	서울	용산구	206061.0	2.0	13976	16306	0.000010	0.135095



# 전처리 과정

## 3. 새로운 feature 및 지표 산출

ex) 공원시설 종류 5개 합쳐 공원시설 개수로 계산, 공원 조성 일 평균 년도 계산

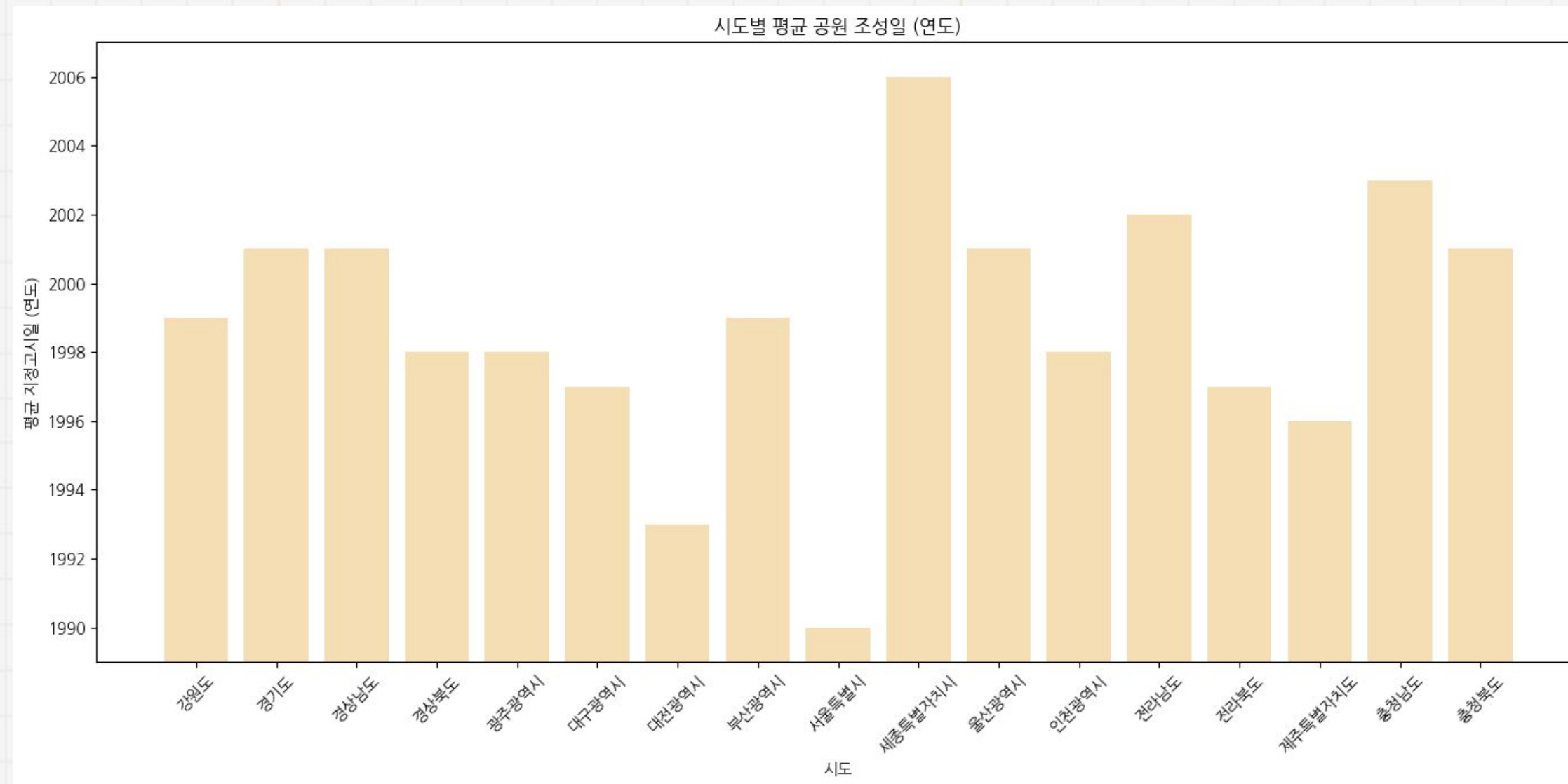
	공원명	공원구분	위도	경도	공원면적	공원보유시설(운동시설)	공원보유시설(유희시설)	공원보유시설(편익시설)	공원보유시설(교양시설)	공원보유시설(기타시설)	지정고시일	관리기관명	행정자치구역
0	봉화산2지구2(3호)	근린공원	37.353826	127.921297	33036.0	NaN	NaN	NaN	NaN	NaN	2009-11-20	강원도 원주시청	강원도 원주시
1	단구공원	근린공원	37.326433	127.953644	290813.0	농구장+배드민턴장	NaN	NaN	NaN	NaN	1967-04-19	강원도 원주시청	강원도 원주시
2	호국(태장2호)공원	근린공원	37.365319	127.952681	33539.0	NaN	NaN	NaN	NaN	NaN	1967-04-19	강원도 원주시청	강원도 원주시
3	단계조각(단계3호)공원	근린공원	37.340966	127.935690	22216.0	풋살경기장	NaN	NaN	청소년도서관+체육관	NaN	1974-11-26	강원도 원주시청	강원도 원주시
4	일산공원	근린공원	37.346851	127.937280	21182.0	NaN	NaN	NaN	NaN	NaN	1974-11-26	강원도 원주시청	강원도 원주시



	공원명	공원구분	위도	경도	공원면적	공원보유시설(운동시설)	공원보유시설(유희시설)	공원보유시설(편익시설)	공원보유시설(교양시설)	공원보유시설(기타시설)	지정고시일	관리기관명	행정자치구역	시도별	공원시설통합	공원시설통합num	평균년도
0	봉화산2지구2(3호)	근린공원	37.353826	127.921297	33036.0	NaN	NaN	NaN	NaN	NaN	2009-11-20	강원도 원주시청	강원도 원주시	강원도		0	1999
1	단구공원	근린공원	37.326433	127.953644	290813.0	농구장+배드민턴장	NaN	NaN	NaN	NaN	1967-04-19	강원도 원주시청	강원도 원주시	강원도	농구장+배드민턴장	2	1999
2	호국(태장2호)공원	근린공원	37.365319	127.952681	33539.0	NaN	NaN	NaN	NaN	NaN	1967-04-19	강원도 원주시청	강원도 원주시	강원도		0	1999
3	단계조각(단계3호)공원	근린공원	37.340966	127.935690	22216.0	풋살경기장	NaN	NaN	청소년도서관+체육관	NaN	1974-11-26	강원도 원주시청	강원도 원주시	강원도	풋살경기장+청소년도서관+체육관	3	1999
4	일산공원	근린공원	37.346851	127.937280	21182.0	NaN	NaN	NaN	NaN	NaN	1974-11-26	강원도 원주시청	강원도 원주시	강원도		0	1999

# 공원

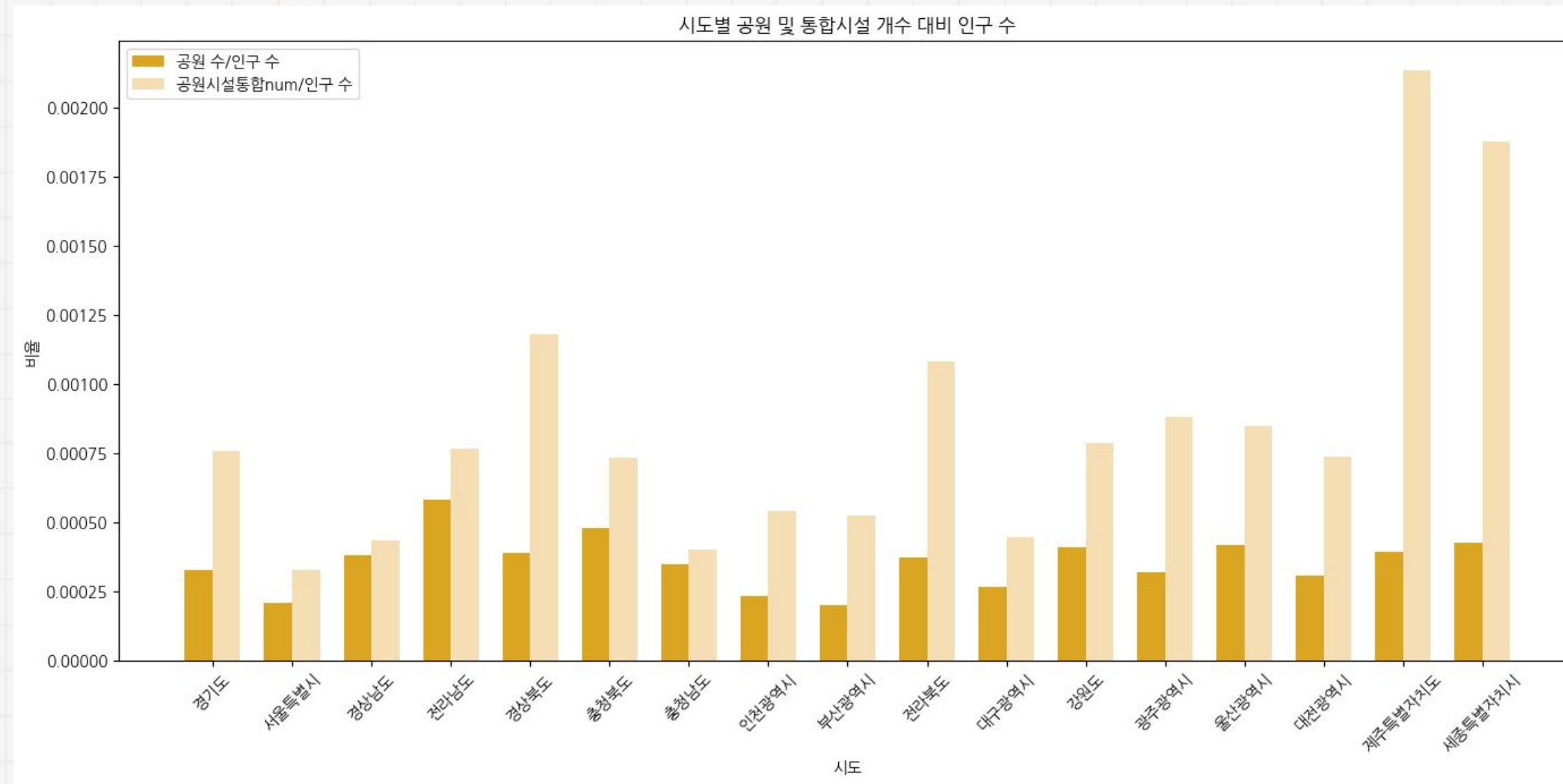
우리나라 공원 현황 시각화



1. 서울특별시의 공원 조성일이 다른 시도에 압도적으로 오래됨.
2. 세종특별자치시 등 비교적 최근에 조성된 도시의 공원 조성일이 최근이다.

# 공원

우리나라 공원 현황 시각화

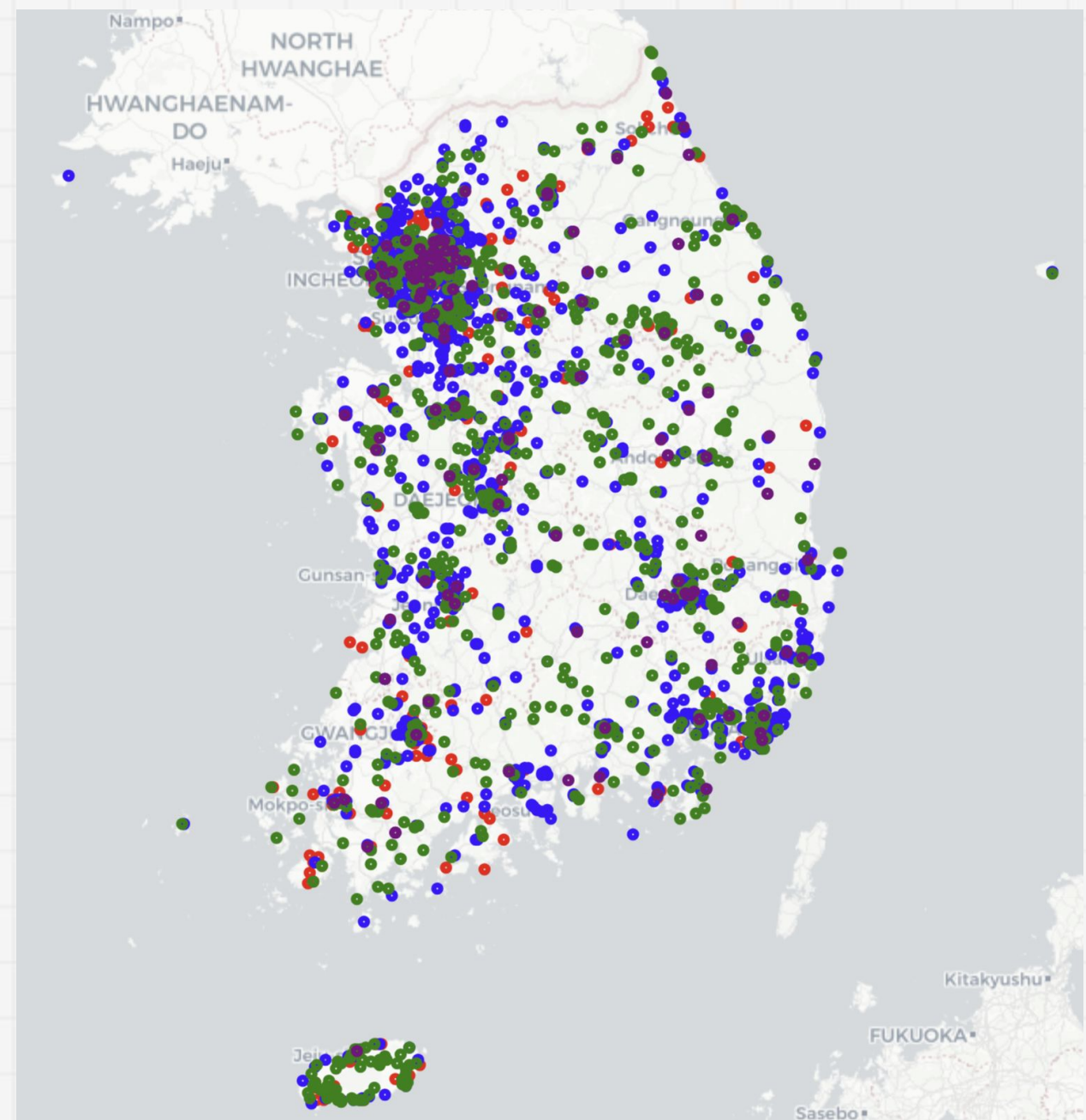


1. 경기도, 서울특별시의 절대적 공원 수는 많으나, 인구 수 대비 공원 수는 적다.
2. 공원 수 간 격차보다 공원 시설 격차가 극명하다.



# 문화시설

도서관 미술관, 박물관



● 박물관 ● 미술관 ● 도서관 ● 지역문화재단

---미술관---			
	count	인구	시설 대비 인구
울산	1	1106	1106.000000
인천	5	3009	601.800000
대구	4	2360	590.000000

---박물관---			
	count	인구	시설 대비 인구
대구	17	2360	138.823529
광주	13	1463	112.538462
울산	10	1106	110.600000

---도서관---			
	count	인구	시설 대비 인구
부산	51	3284	64.392157
대전	25	1474	58.960000
울산	20	1106	55.300000

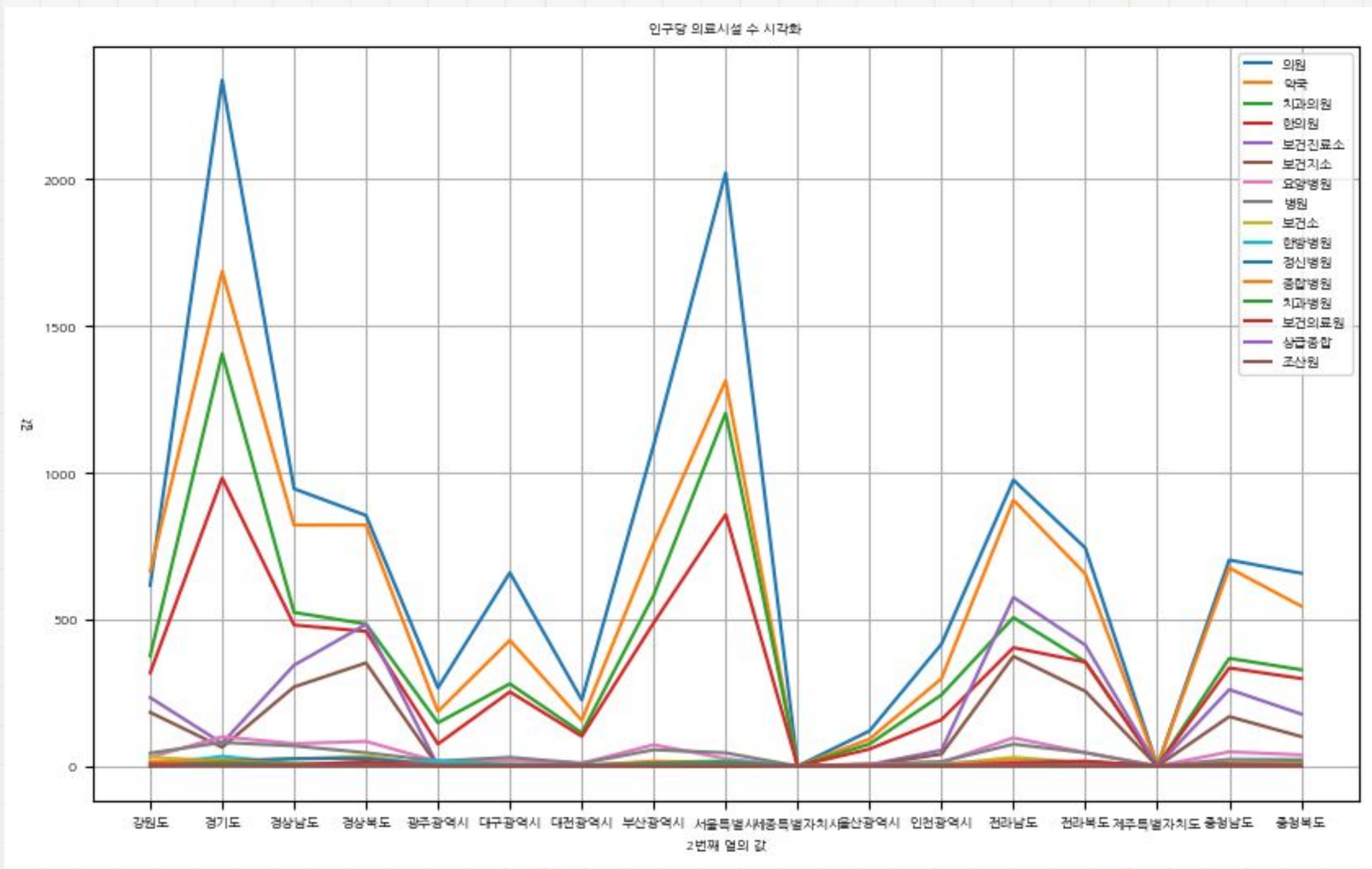
---지역문화재단---			
	count	인구	시설 대비 인구
광주	1	1463	1463.000000
부산	3	3284	1094.666667
대전	2	1474	737.000000

시설 대비 인구, 시설 1 개당 감당해야 하는 지역 내 인구를 뜻함.

문화시설들이 대부분 수도권에 집중되어 있어 상대적으로 남부 지역의 문화시설이 인구 대비 미비한 상태임을 확인

# 의료시설

우리나라 의료시설 시각화

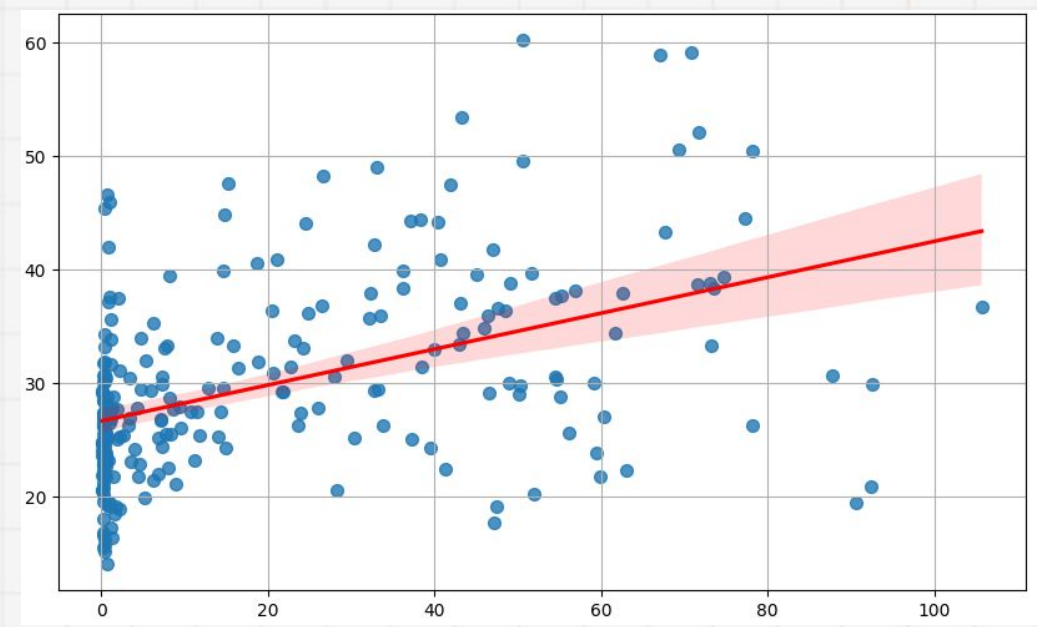


인구 10만명당 의료시설 수  
수도권 지역에 지나치게 몰려있다.

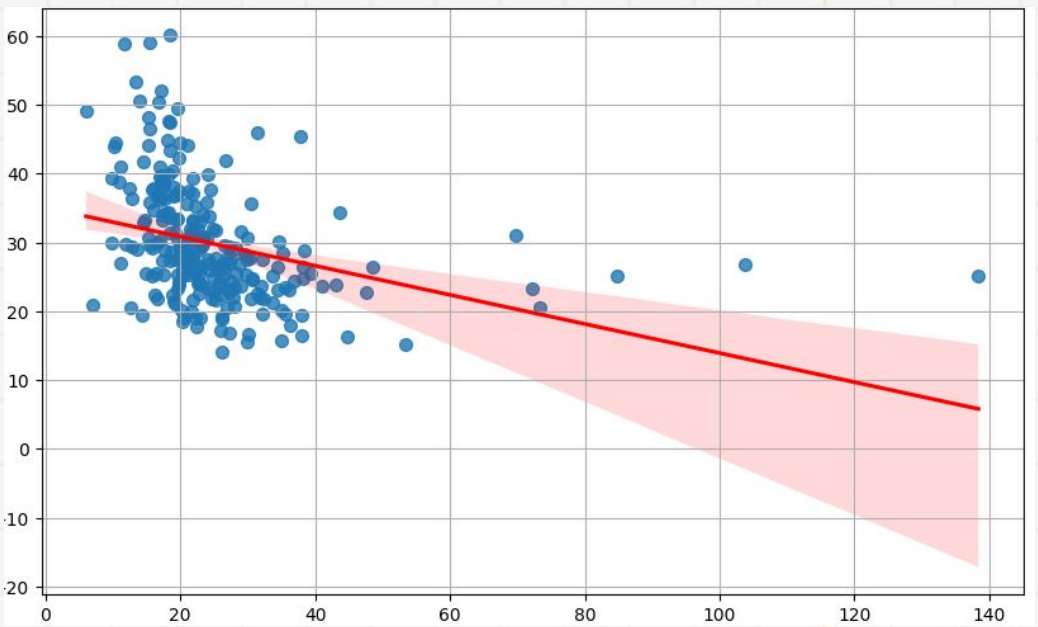
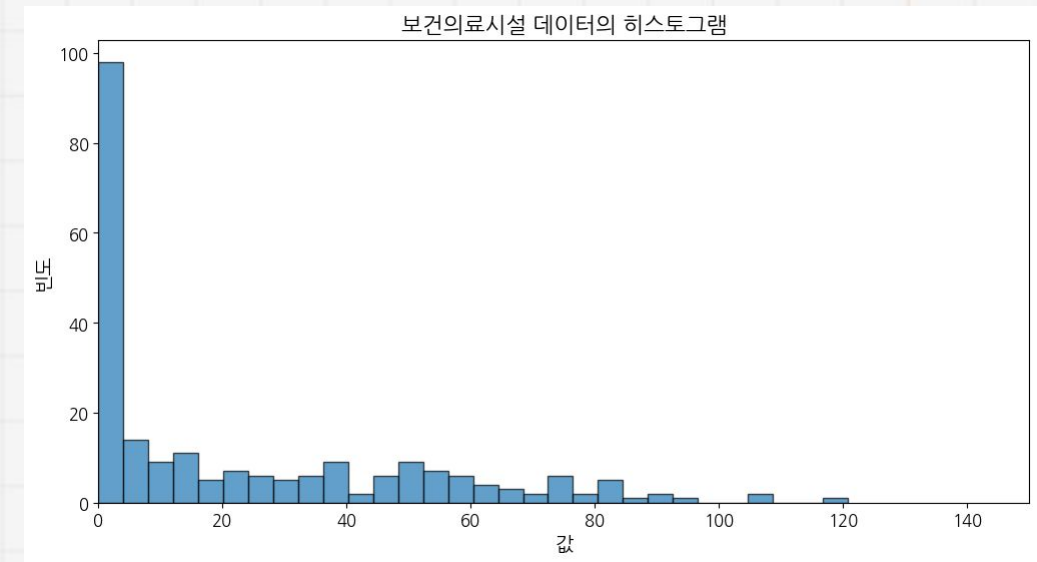


# 의료시설

우리나라 의료시설 시각화

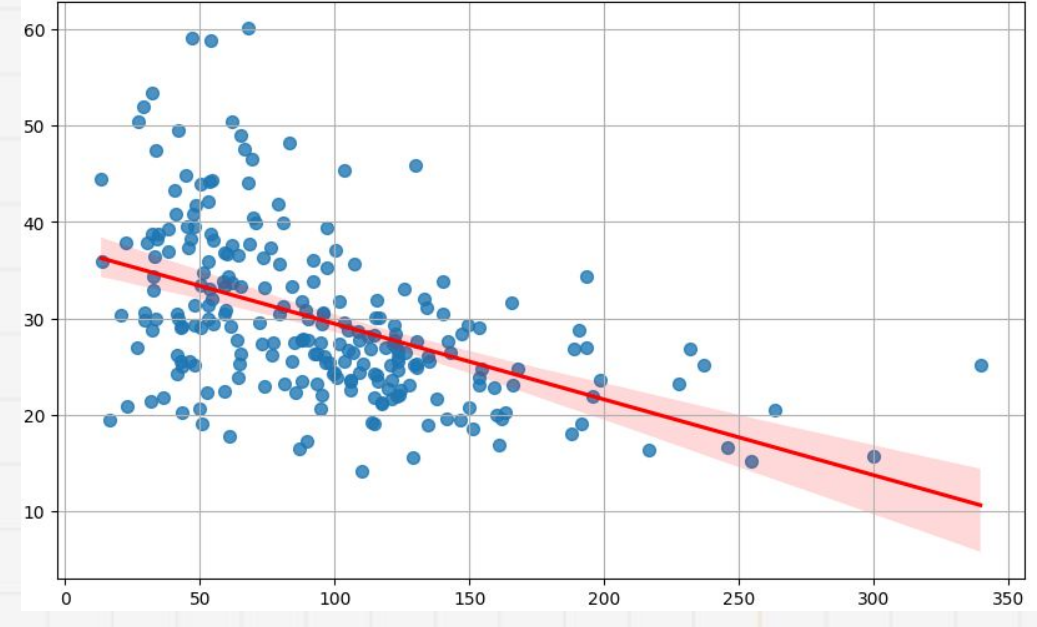


인구당 보건의료시설  
(보건소+보건의료원+보건지소+보건진료소)  
과 자살률 사이의 상관관계



인구당 치과의료시설  
(치과병원+치과의원)  
과 자살률 사이의 상관관계

2015년	0.4983	-0.5204	-0.5173
2016년	0.5600	-0.4915	-0.5315
2017년	0.4325	-0.3817	-0.4394
2018년	0.3999	-0.3402	-0.3871
2019년	0.3462	-0.3035	-0.3383
2020년	0.3163	-0.2837	-0.3408
2021년	0.4713	-0.3735	-0.4550
2022년	0.3097	-0.3099	-0.3619



노인 인구당 한방의료시설  
(한방병원+한의원)  
과 자살률 사이의 상관관계

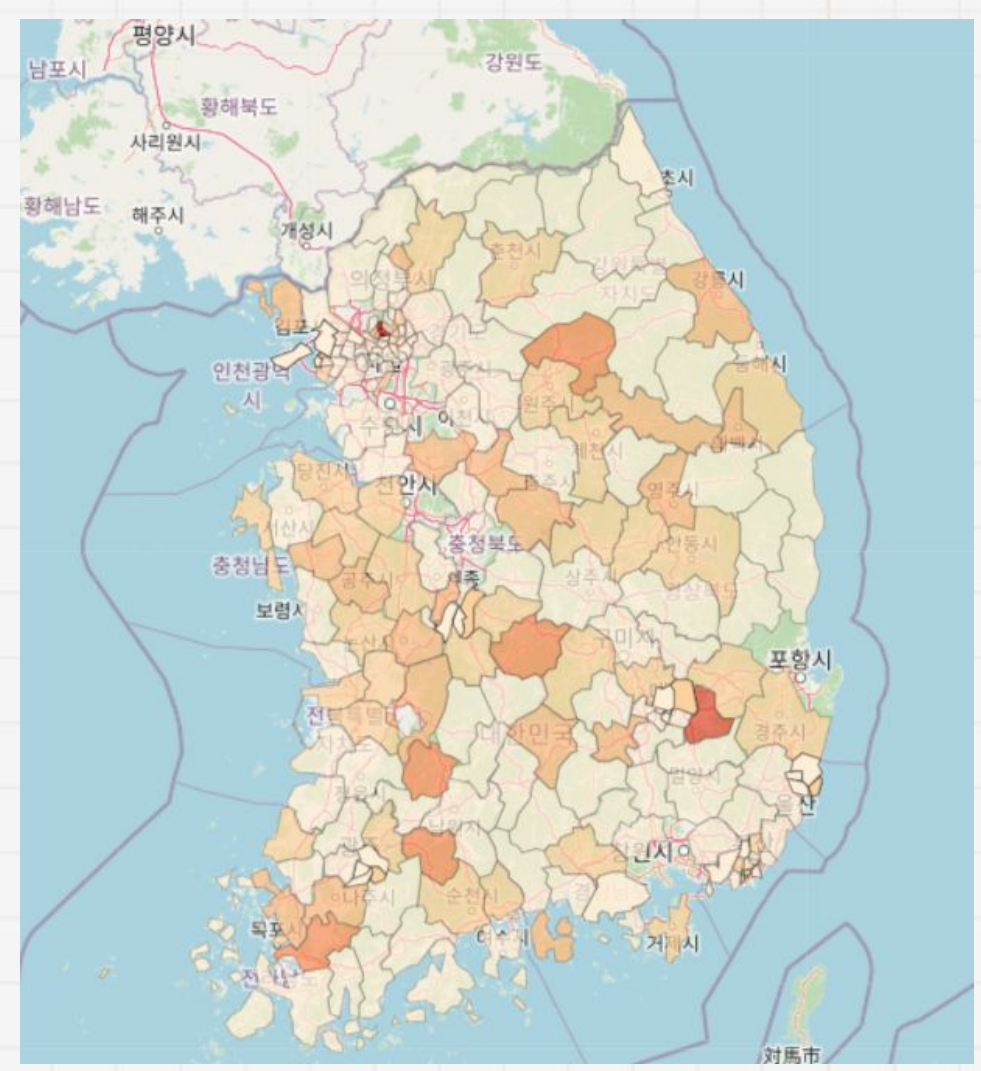
정규성을 만족하지 않아  
Spearman 계수를 사용함.



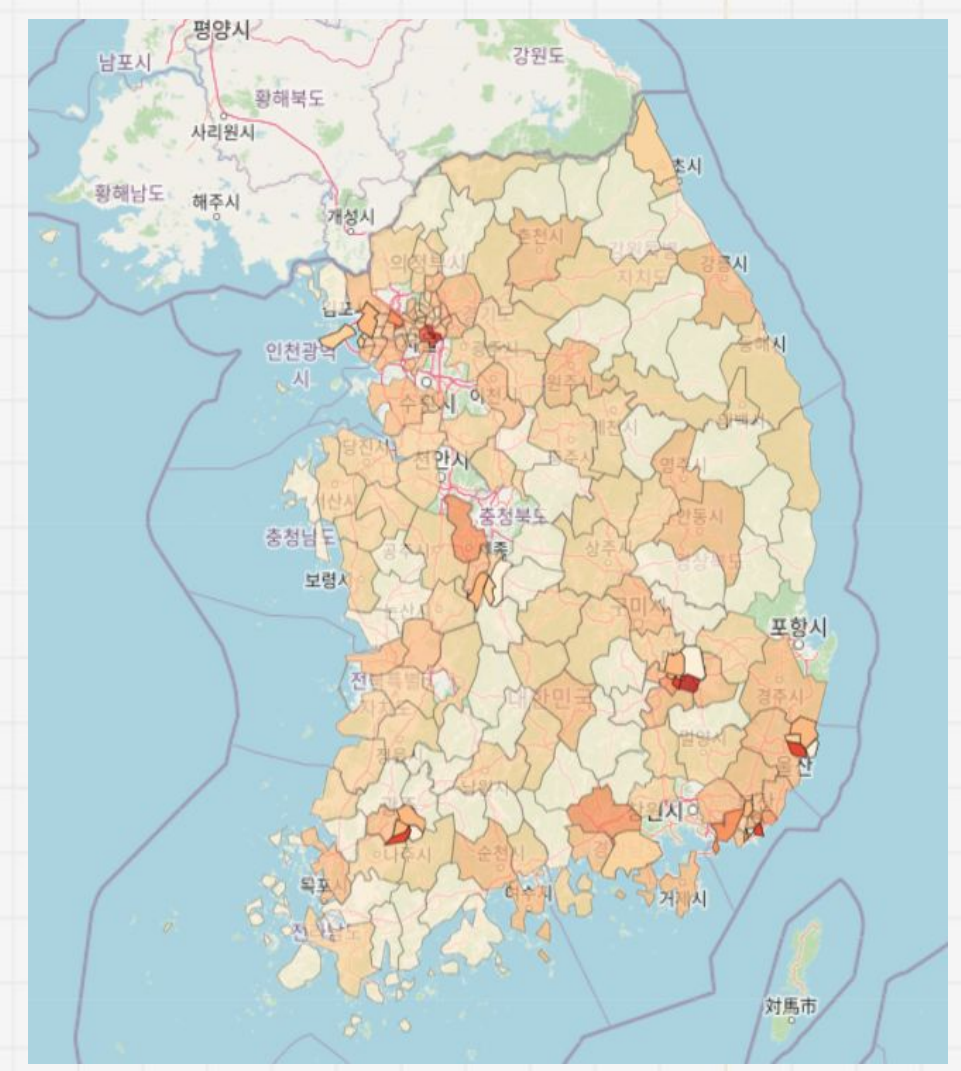
# 교육시설

학원 수, 고등교육기관, 유,초,중등 학교 분포 현황

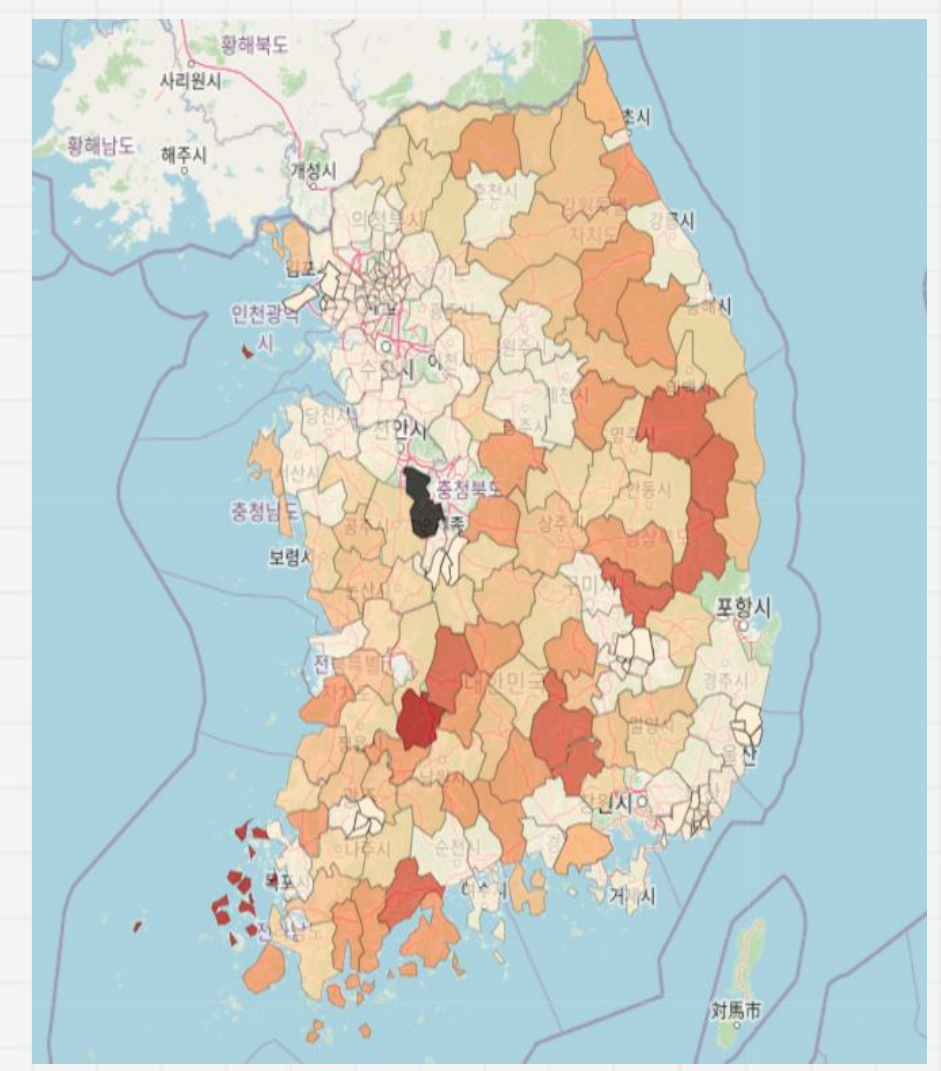
고등교육기관의 수



학원의 수



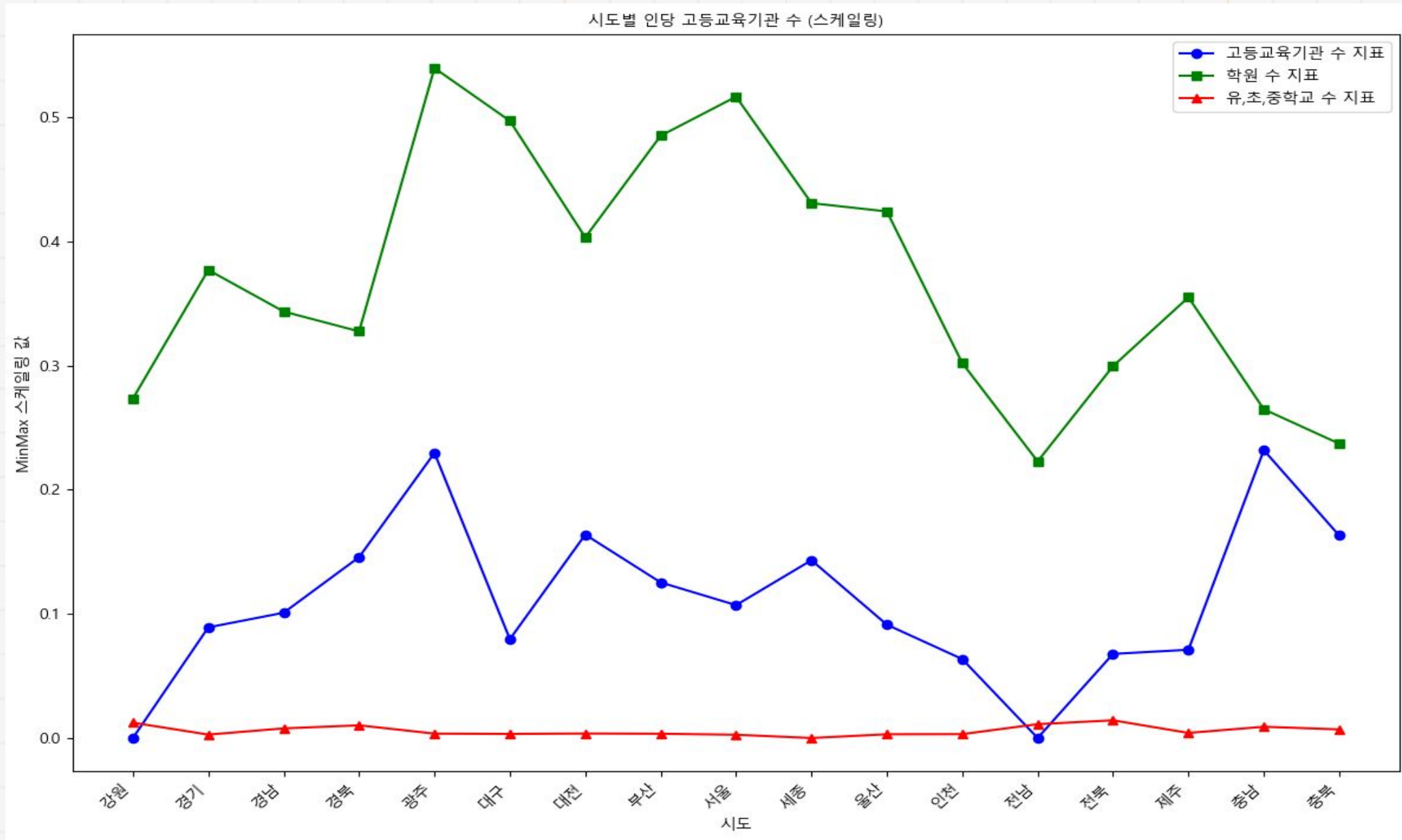
유/초/중등학교의 수



※ 세종시의 학교 수 데이터는 부재함

# 교육시설

## 시도별 지표 값 시각화



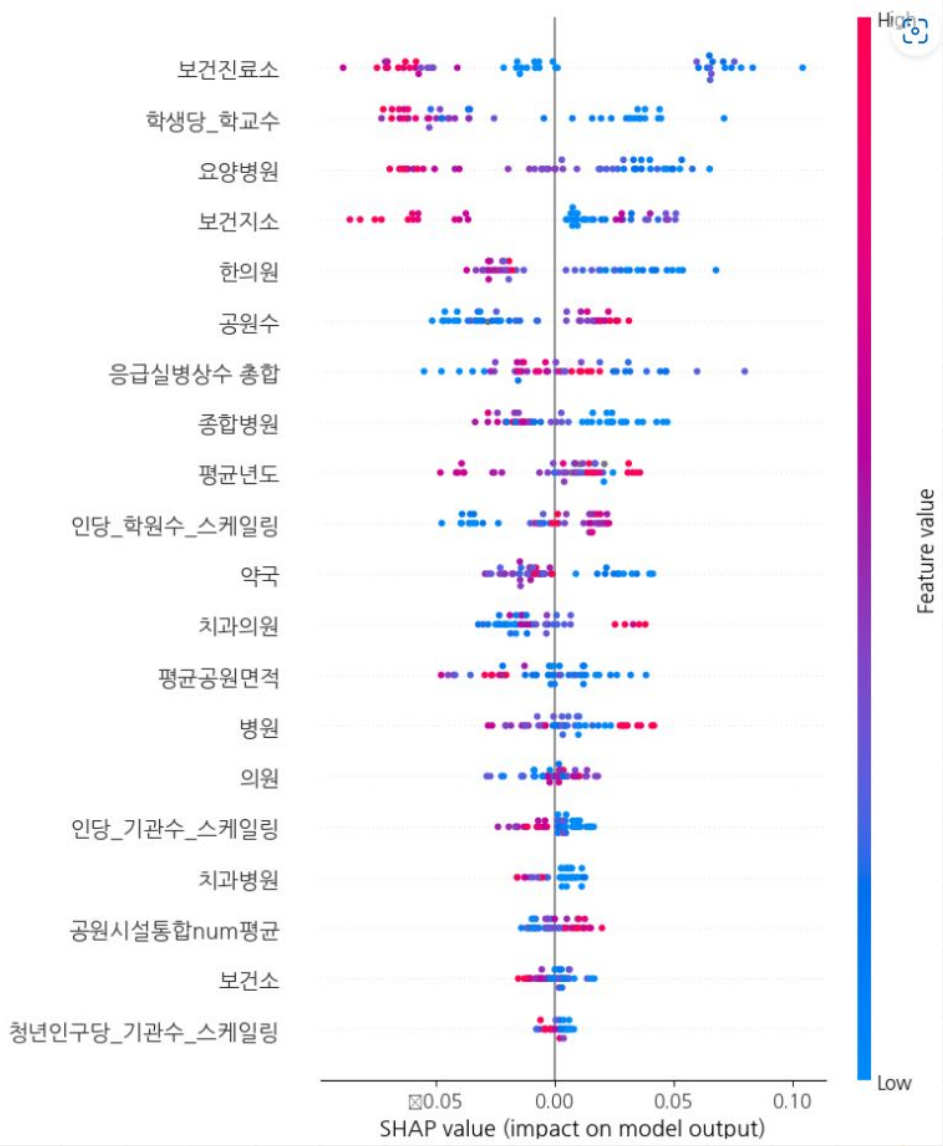
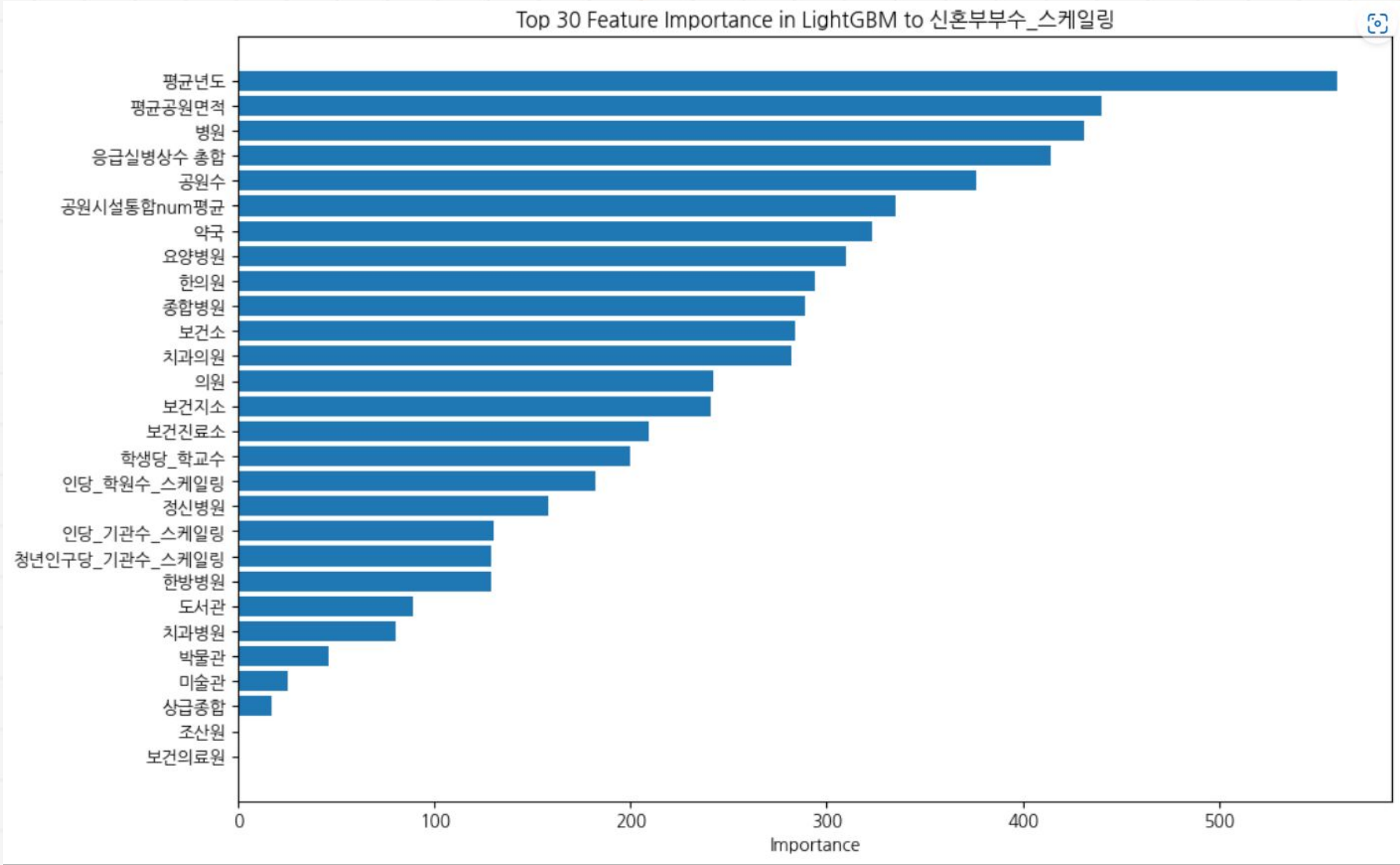
- 유/초/중학교는 고르게 분포되어 있음
- 학원 수 지표는 대도시 위주로 높은 경향을 보임
- 학원 수와 고등교육기관의 수는 지역 간 편차가 심함



# 도시환경 시설 종합

- 1. lightGBM 모델 이용하여 feature importance 계산
- 2. Shap Value 이용하여 종속 변수에 대한 영향력 계산

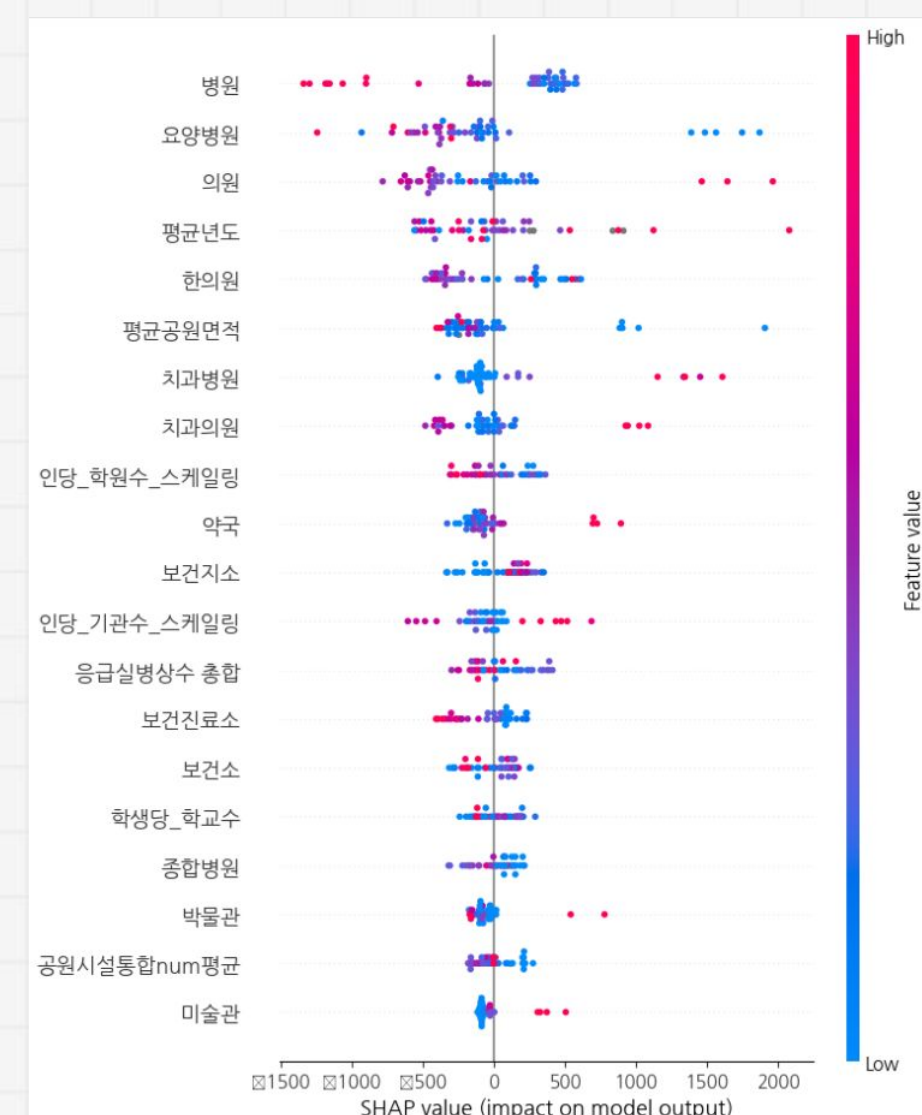
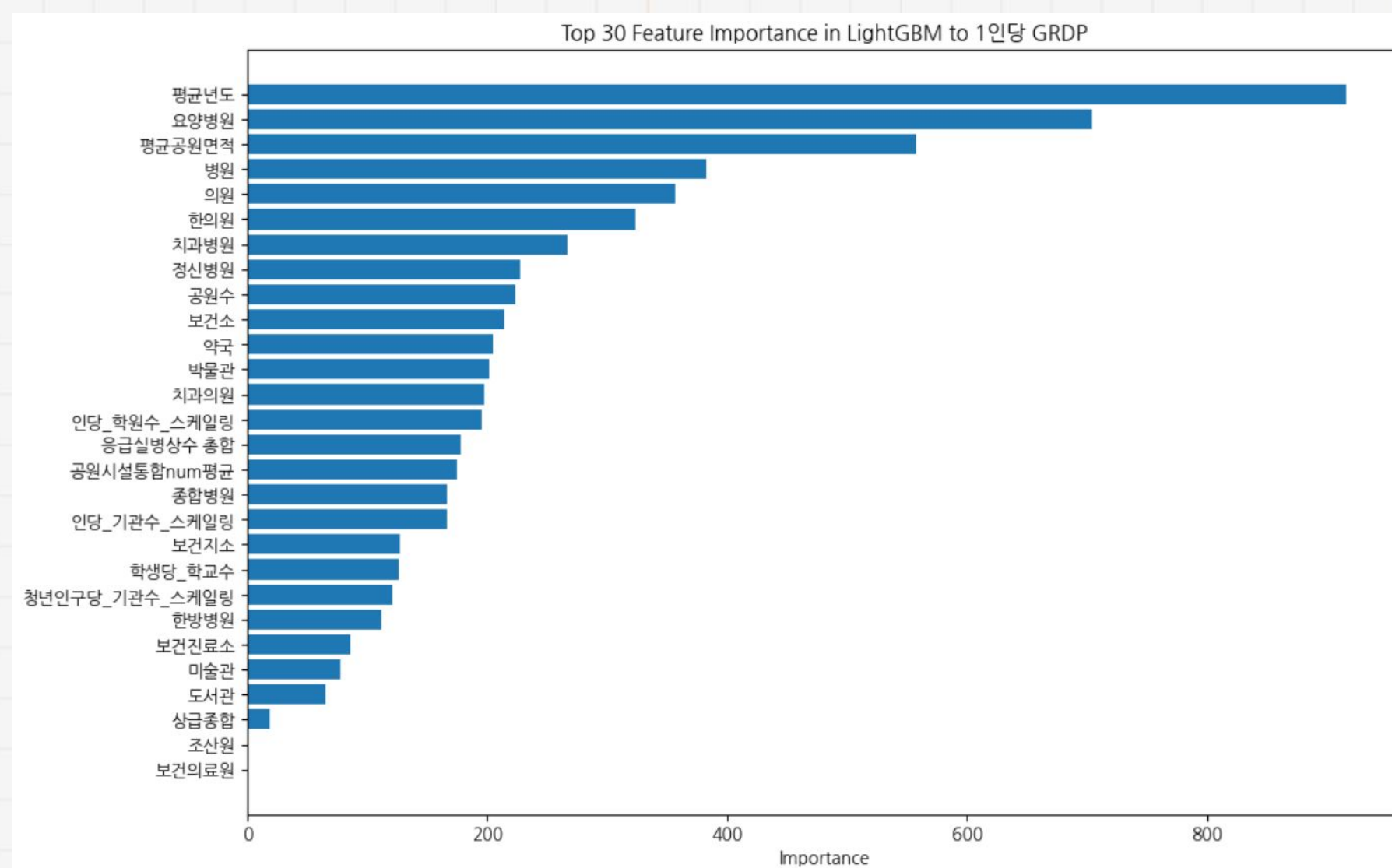
<신혼부부수 (인구 수로 나눈 후 스케일링)>





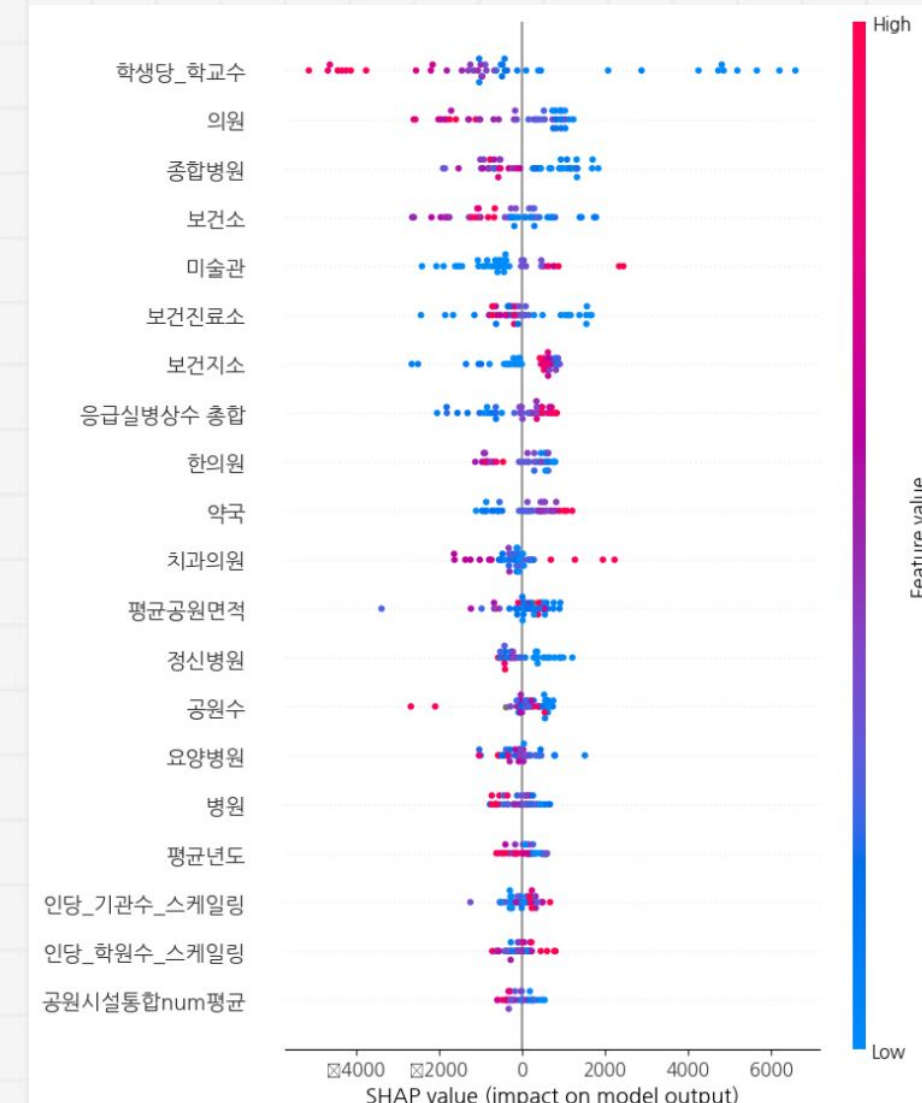
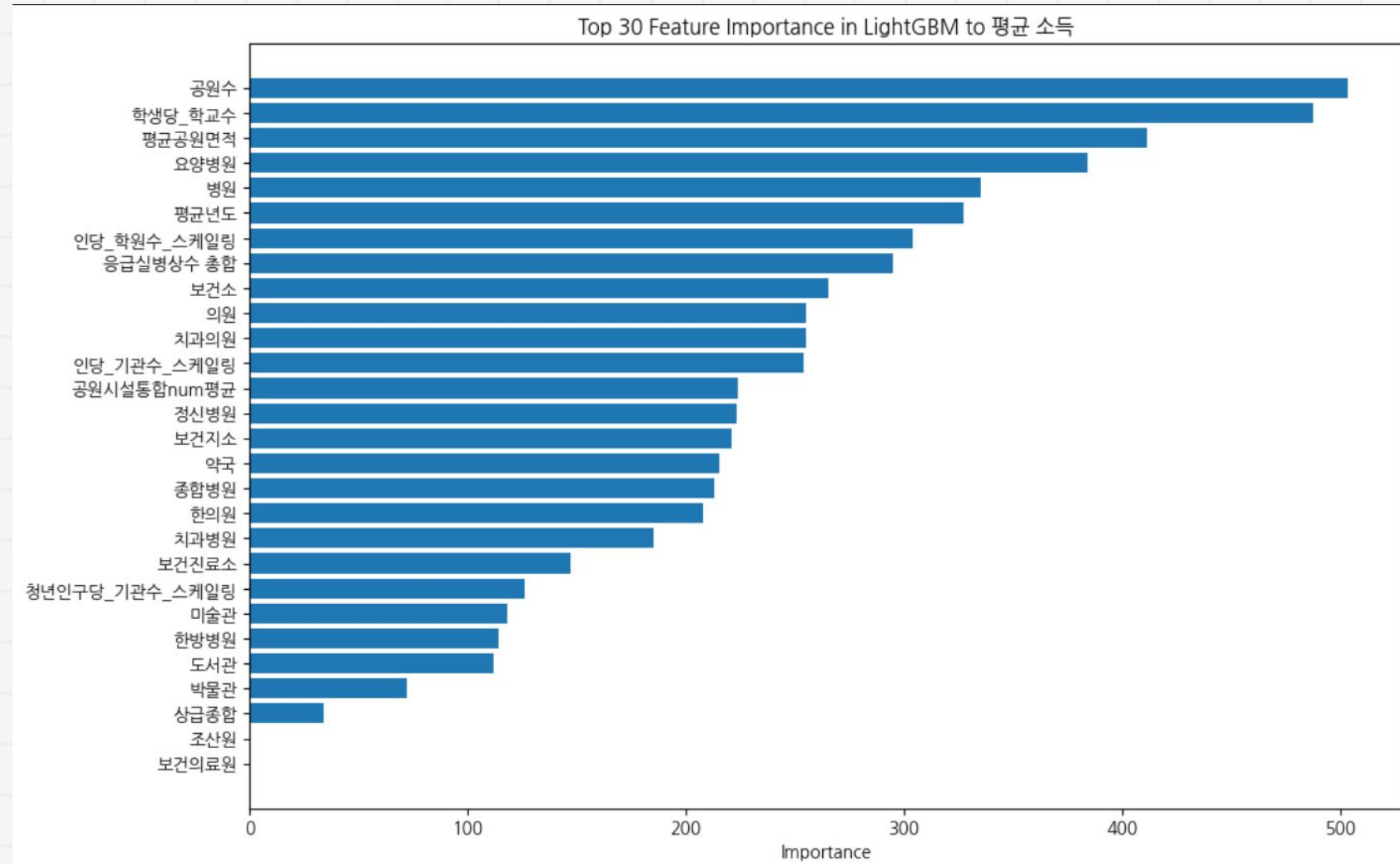
# 도시환경 시설 종합

<1인당 GRDP (지역내총생산)>



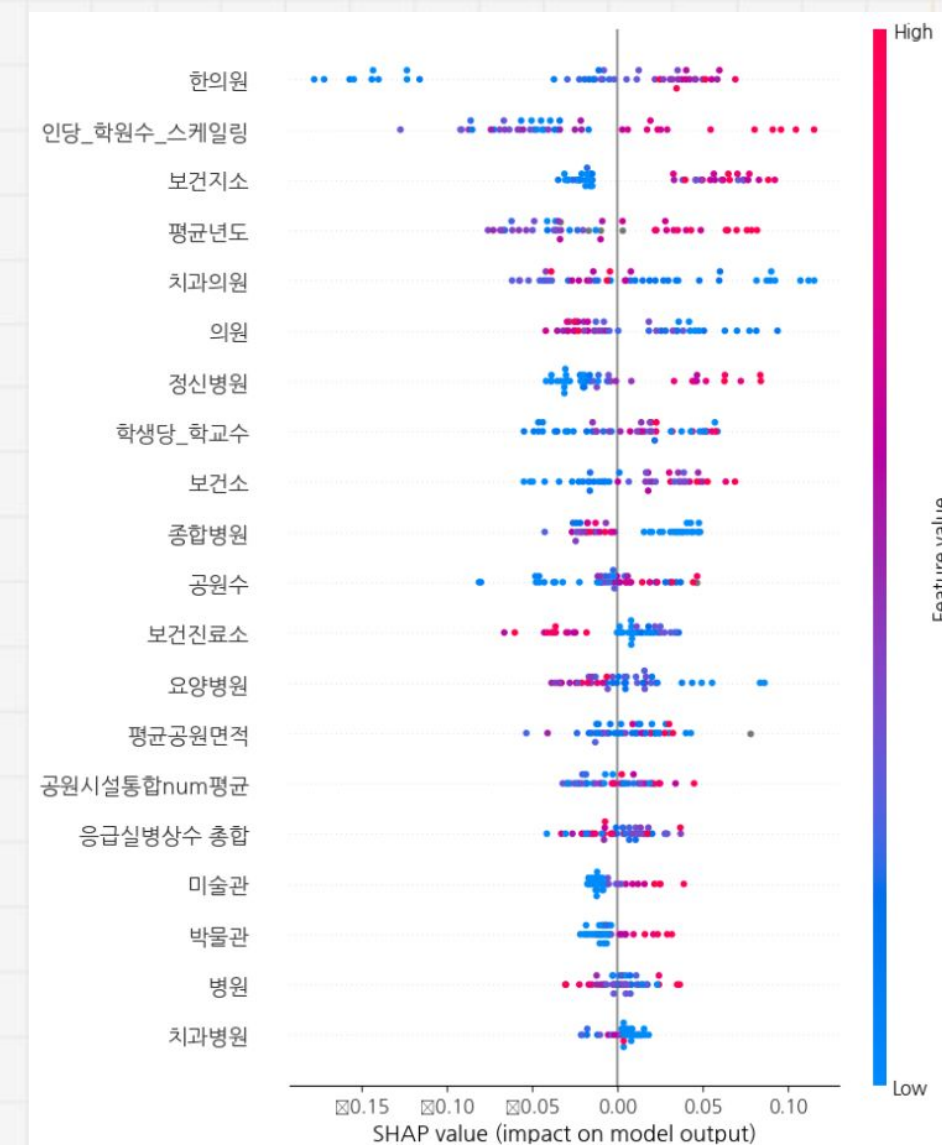
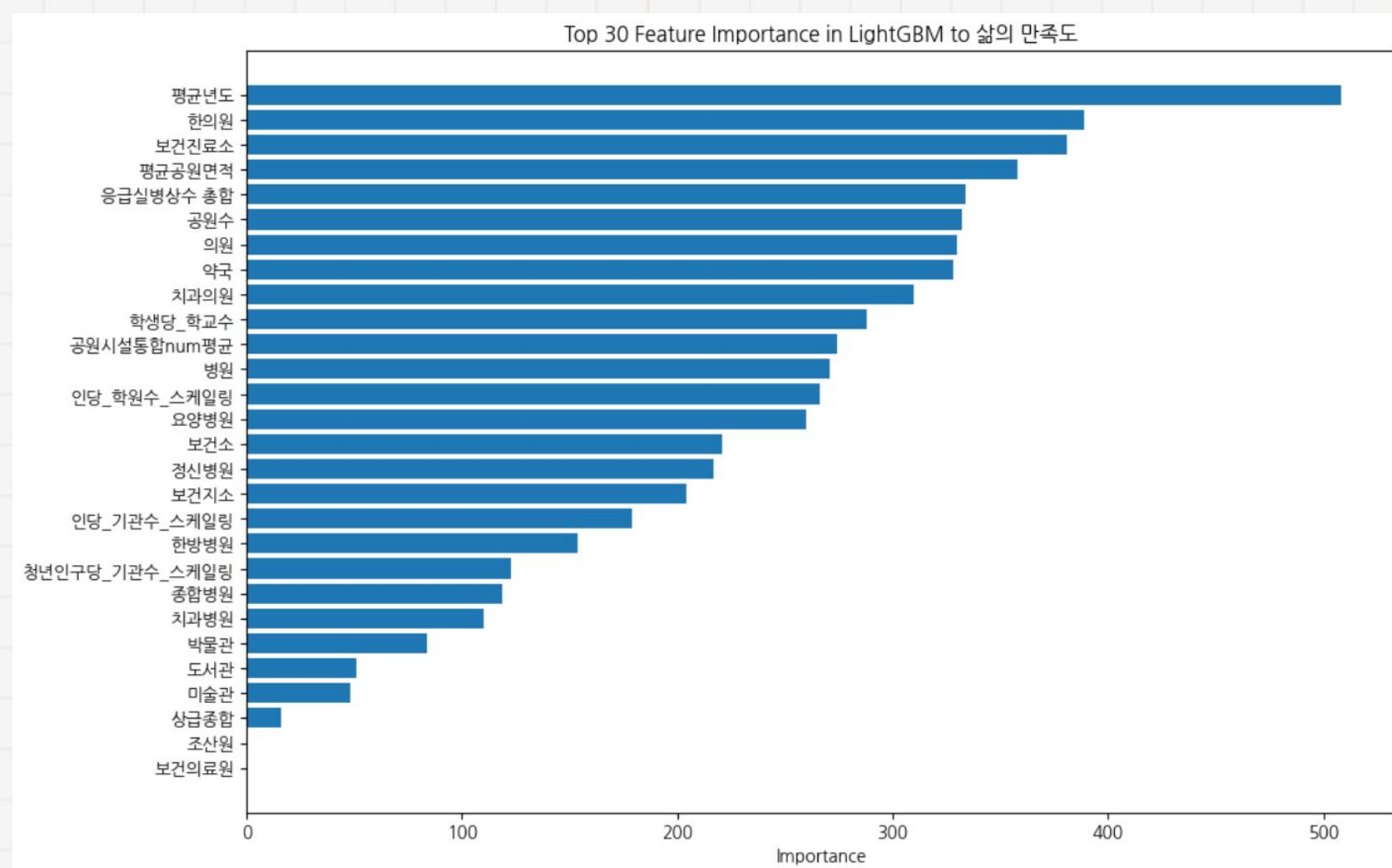
# 도시환경 시설 종합

## <평균소득>



# 도시환경 시설 종합

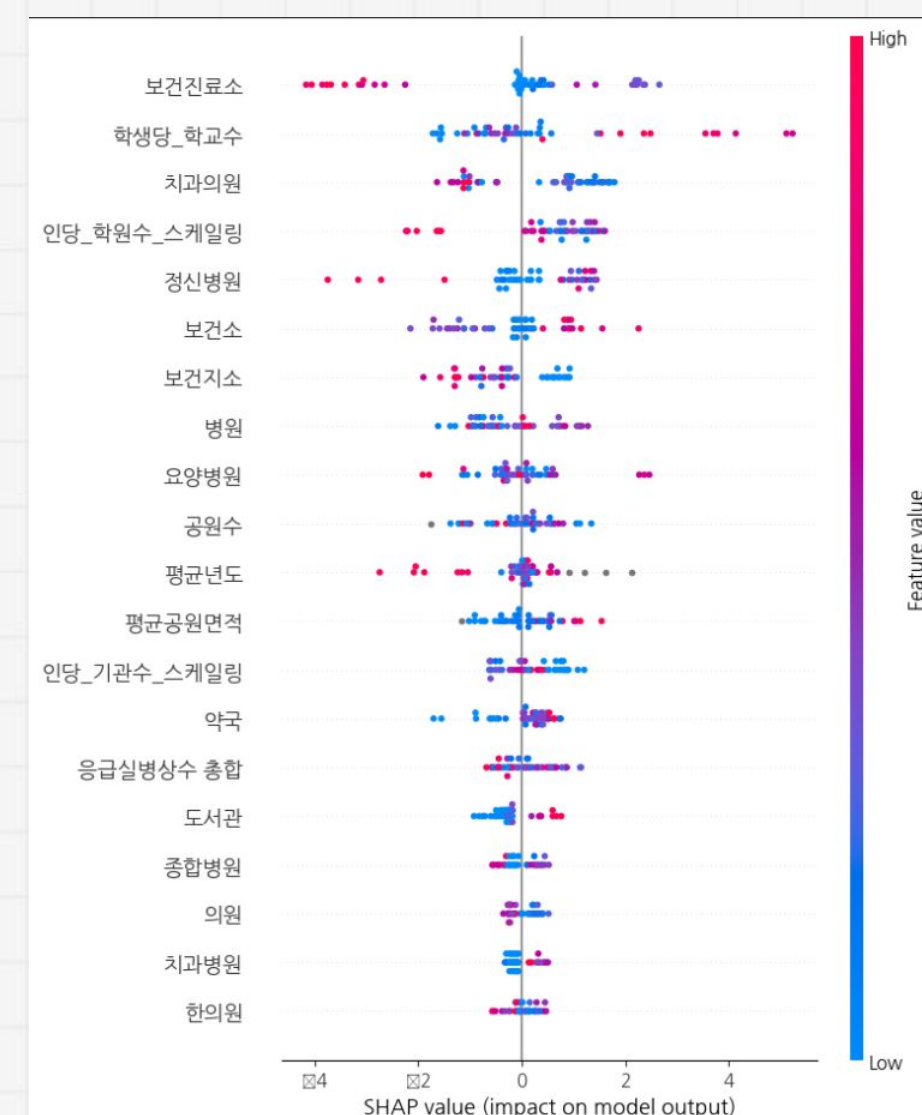
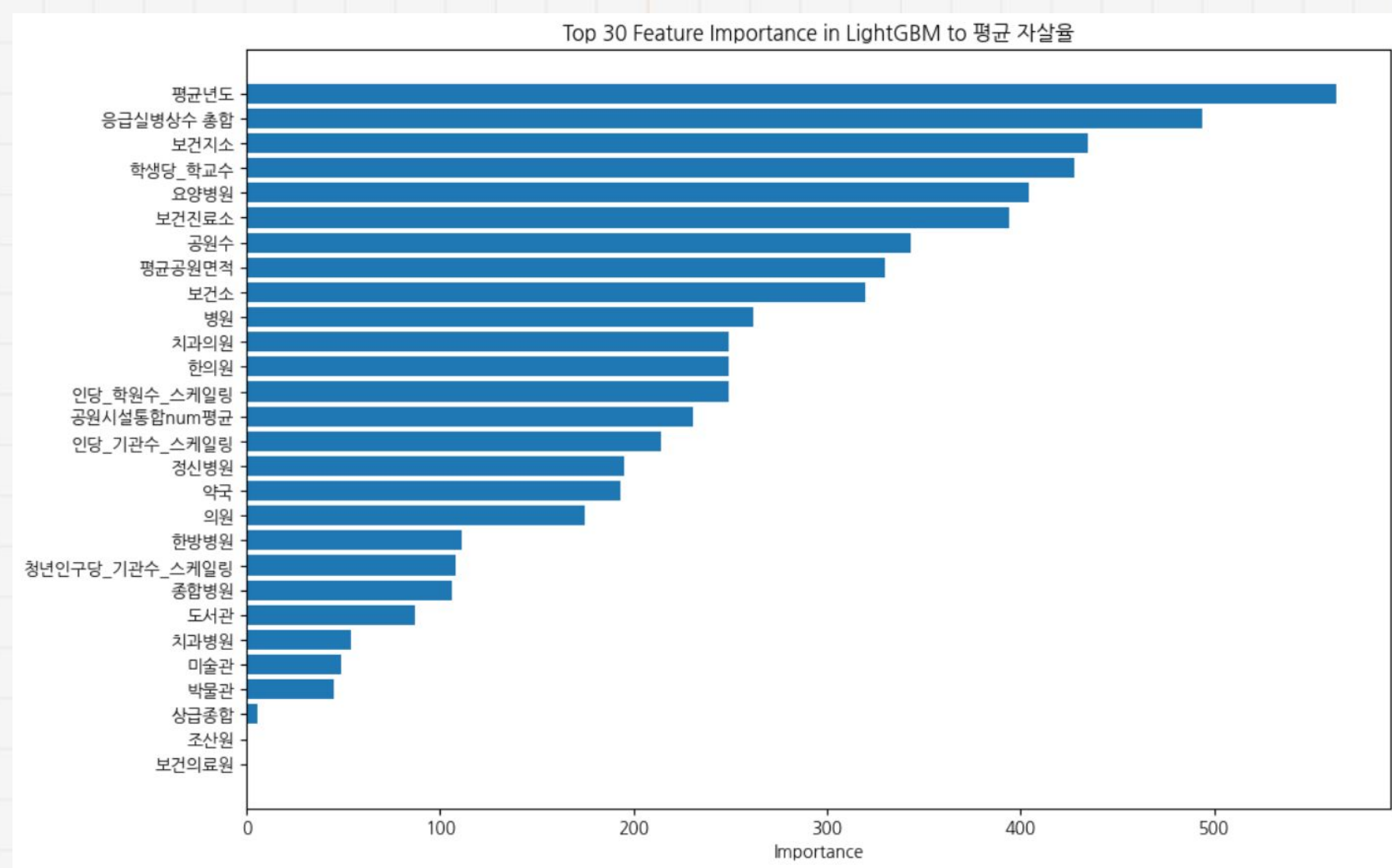
## <삶의 만족도>





# 도시환경 시설 종합

## <평균 자살율>

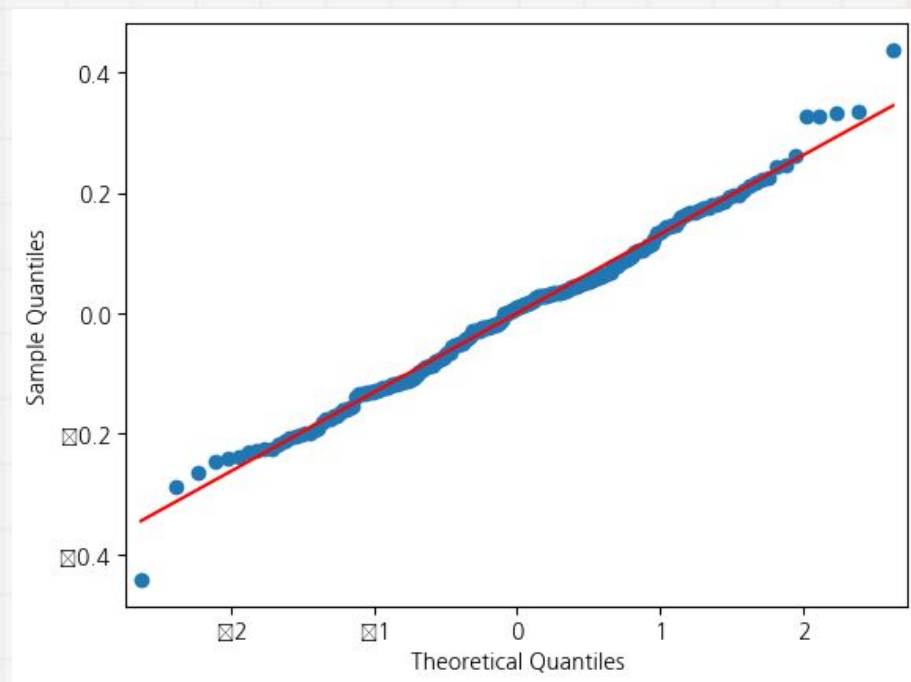


# 도시환경 시설 총합

## 3. 상위 5개 변수 선택하여 다중 회귀 분석 실시

	VIF Factor	features
0	68137.265999	const
1	1.012586	평균년도
2	1.101430	평균공원면적
3	1.430131	병원
4	1.440388	응급실병상수 총합
5	1.118809	공원수

종속 변수 5개 중  
 신혼부부수만 회귀 분석 가정 충족  
 +낮은 다중공선성



### 회귀분석 결과 정리

변수명	계수	표준 오차	t-값	p-값	95% 신뢰구간
Const	0.9627	2.295	0.420	0.675	[-3.559, 5.485]
평균년도	-0.0003	0.001	-0.294	0.769	[-0.003, 0.002]
평균공원면적	-0.0003	0.001	-0.294	0.769	[-0.003, 0.002]
병원	-2.675e-07	2.05e-07	-1.304	0.194	[-6.72e-07, 1.37e-07]
응급실병상수_총합	-0.0052	0.004	-1.278	0.203	[-0.013, 0.003]
공원수	0.0009	0.000	7.960	0.000	[0.001, 0.001]

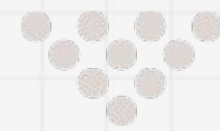
결과 : 공원수가 신혼부부수에  
 통계적으로 유의미한 양의 영향을 미친다.

# 인사이트

1. 공원 수과 신혼부부 수 사이에 유의미한 양의 영향력이 존재한다.
2. 삶의 만족도, 자살률, 신혼부부 수에 feature importance가 높은 변수는 **공원의 연도**이다.
3. 도서관 수, 박물관 수, 미술관 수는 도시환경 관련 종속변수에 영향을 미치지 않는다.
4. 공원과 달리, 의료시설과 문화시설은 수도권에 지나치게 집중되어 있다.



# 한계점



1. 데이터 연도 불일치 : 독립변수의 경우 2024년 데이터인 반면, 종속변수의 경우 2019/2022년
2. 정규성 만족 X : 거의 모든 컬럼이 정규성을 만족하지 못해 회귀 분석의 어려움
3. 결측치: 의료시설 데이터에서 값이 입력되지 않은 지역 다수
4. 불충분한 설명력: '도시환경'에 집중하게 위해 인구학적 특성을 거의 다루지 않아 설명력이 다소 부족해짐



# 감사합니다.

EDA 도시환경 | 10기 박성원 11기 양주원 한은결 12기 김건우

