

# 데이터 공학 최종발표

1901588 김동현  
1901599 이상민

- ❖ 팀 구성
- ❖ 팀원 역할 분담
- ❖ 분석 목적
- ❖ 사용 데이터
- ❖ 분석 결과
- ❖ 논의사항
- ❖ 소감 및 결론

## ❖ 김동현

- ✓ 전기사용량 데이터 정제 및 통합
- ✓ 연면적 및 층수에 따른 전기 사용량 분석 및 시각화
- ✓ 법정동별 전기사용량 분석 및 시각화
- ✓ 법정동 계절별 전기사용량 분석 및 시각화

## ❖ 이상민

- ✓ 건축물대장 데이터 정제 및 통합
- ✓ 주구조, 주지붕별 평균 전기 사용량 분석 및 시각화
- ✓ 건물 주용도별 평균 전기 사용량 분석 및 시각화
- ✓ 전기 데이터 건축물 데이터 상관관계 분석

## ❖ 분석 배경

- ✓ 건물의 에너지 사용량과 그로 인한 탄소배출량을 효율적으로 관리하기 위해, 건물 특성과 전기사용량 간의 관계를 이해하는 것이 중요
- ✓ 본 프로젝트는 건물축 대장 데이터와 서울시 전기사용량 데이터를 분석하여, 이를 통해 에너지 절약 및 탄소배출 감소를 위한 근거 자료를 마련하는 것을 목표
- ✓ 서울시 온실가스 배출량 산정 프로젝트에 활용할 수 있도록 데이터 분석을 하기 위함

## ❖ 분석 목적

- ✓ 건물 연면적이 크면 전기 사용량이 높을 것이다.
- ✓ 건물 건폐율이 크면 전기 사용량이 높을 것이다.
- ✓ 건물 용적률이 크면 전기 사용량이 높을 것이다.
- ✓ 건물 층수가 높으면 전기 사용량이 높을 것이다.
- ✓ 건물 노후도가 크면 전기 사용량이 낮을 것이다.
- ✓ 목구조가 철근콘크리트구조보다 전기 사용량이 낮을 것이다.
- ✓ 공업 시설이 다른 시설보다 전기 사용량이 높을 것이다.
- ✓ 세대수가 많으면 전기사용량이 높을 것이다.
- ✓ 여름과 겨울은 봄, 여름보다 전기 사용량이 높을 것이다.

## ❖ 서울시 건축물대장 데이터

✓ 건축 허브 - <https://www.hub.go.kr/portal/main.do>

✓ 데이터 규모

– 592,145(row) x 53(columns) 총 31,383,685 건

	PK	대장구분	시도	시군구	법정동	번	...	인근전기차대수	면제대수	특수공법여부	용도지역코드명정보	용도지구코드명정보	용
0	100212381	집합	서울특별시	종로구	청운동	1	...	0	0	0.0	NaN	NaN	NaN
1	100212382	집합	서울특별시	종로구	청운동	1	...	0	0	0.0	NaN	NaN	NaN
2	100212383	집합	서울특별시	종로구	청운동	1	...	0	0	0.0	NaN	NaN	NaN
3	100212384	집합	서울특별시	종로구	청운동	1	...	0	0	0.0	NaN	NaN	NaN
4	100212385	집합	서울특별시	종로구	청운동	1	...	0	0	0.0	NaN	NaN	NaN

[5 rows x 53 columns]

데이터 샘플 (head)

## ❖ 서울시 전기에너지 사용량 데이터

✓ 건축물 데이터 개방 - <https://open.eais.go.kr/main/main.do>

✓ 데이터 규모

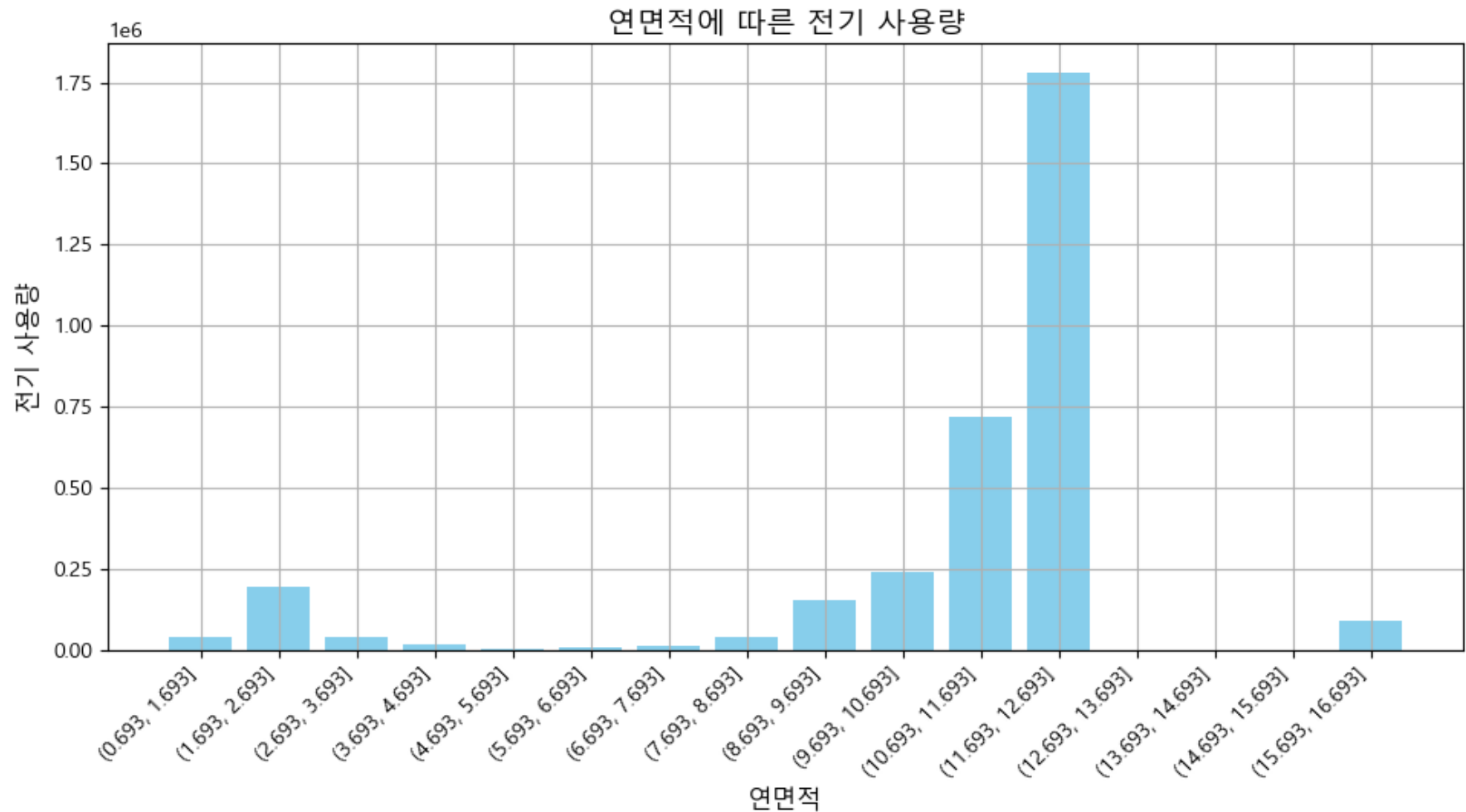
– 80,569(row) x 15(columns) 총 1,208,535 건

3 소본번	새주소부번	사용량(KWh)	순번	대지위치	시군구코드	법정동코드	대지구분코드	...
4			1	서울특별시 양천구 신정동 1000-3번지	11470	10100	0	1000 00
...			...	...	...	...	...	...
3732			3729	서울특별시 양천구 신월동 997-2번지	11470	10300	0	0997 00
3733			3730	서울특별시 양천구 신월동 998-12번지	11470	10300	0	0998 00
3734			3731	서울특별시 양천구 신월동 998-1번지	11470	10300	0	0998 00
3735			3732	서울특별시 양천구 신월동 998-3번지	11470	10300	0	0998 00
3736			3733	서울특별시 양천구 신월동 998-4번지	11470	10300	0	0998 00

[3737 rows x 15 columns]

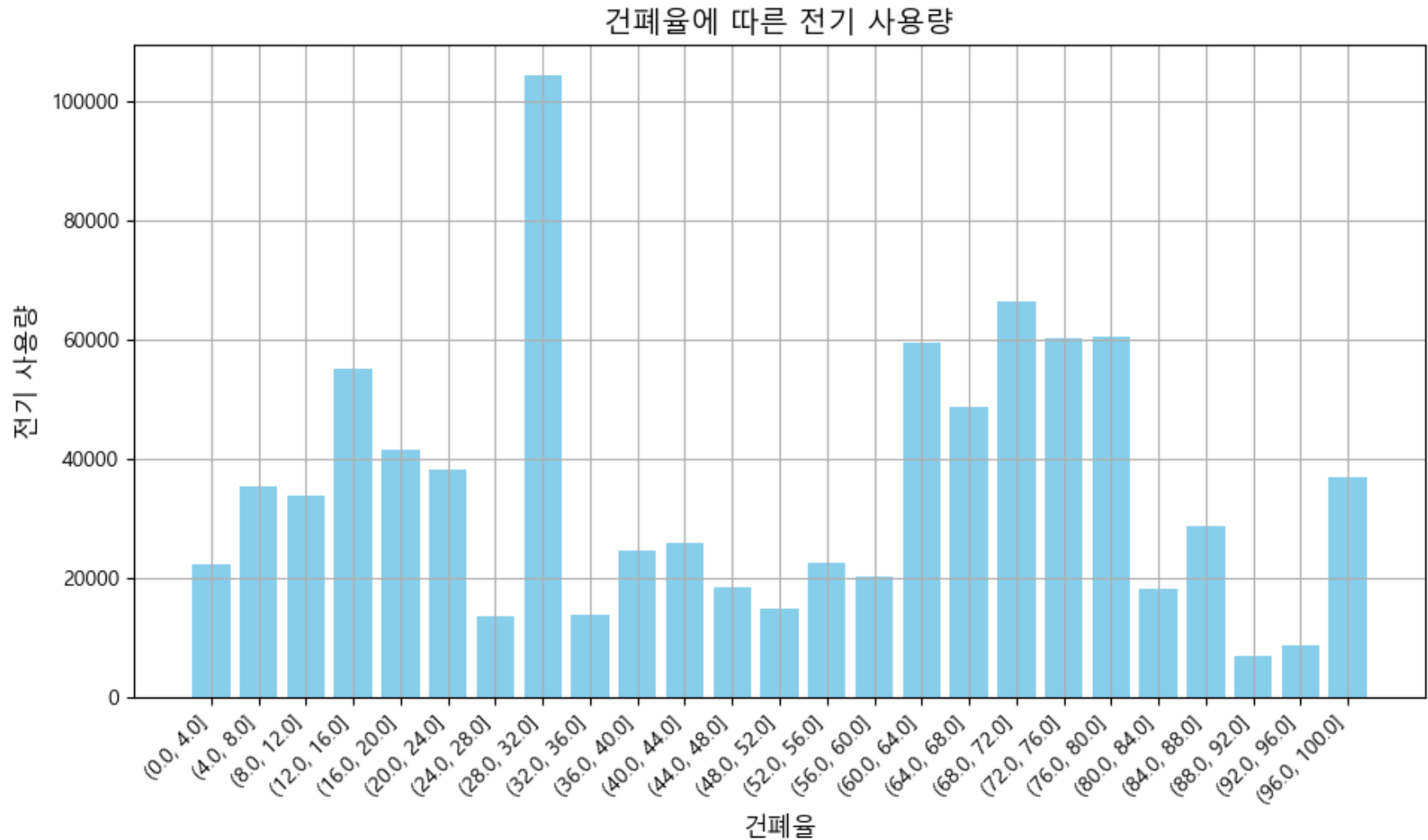
데이터 샘플(양천구)

❖ 건물 연면적이 크면 전기 사용량이 높을 것이다.



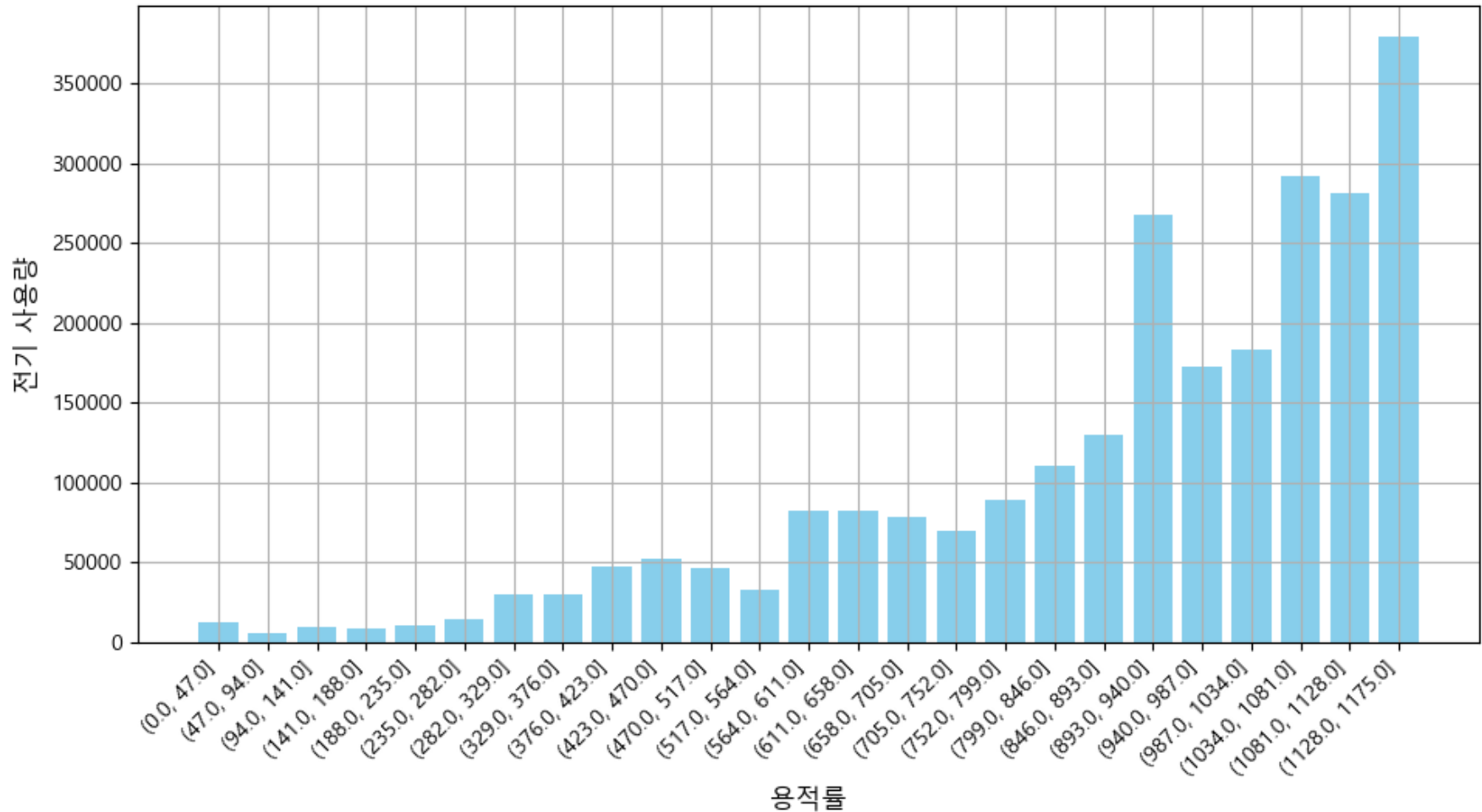


❖ 건물 건폐율이 크면 전기 사용량이 높을 것이다.

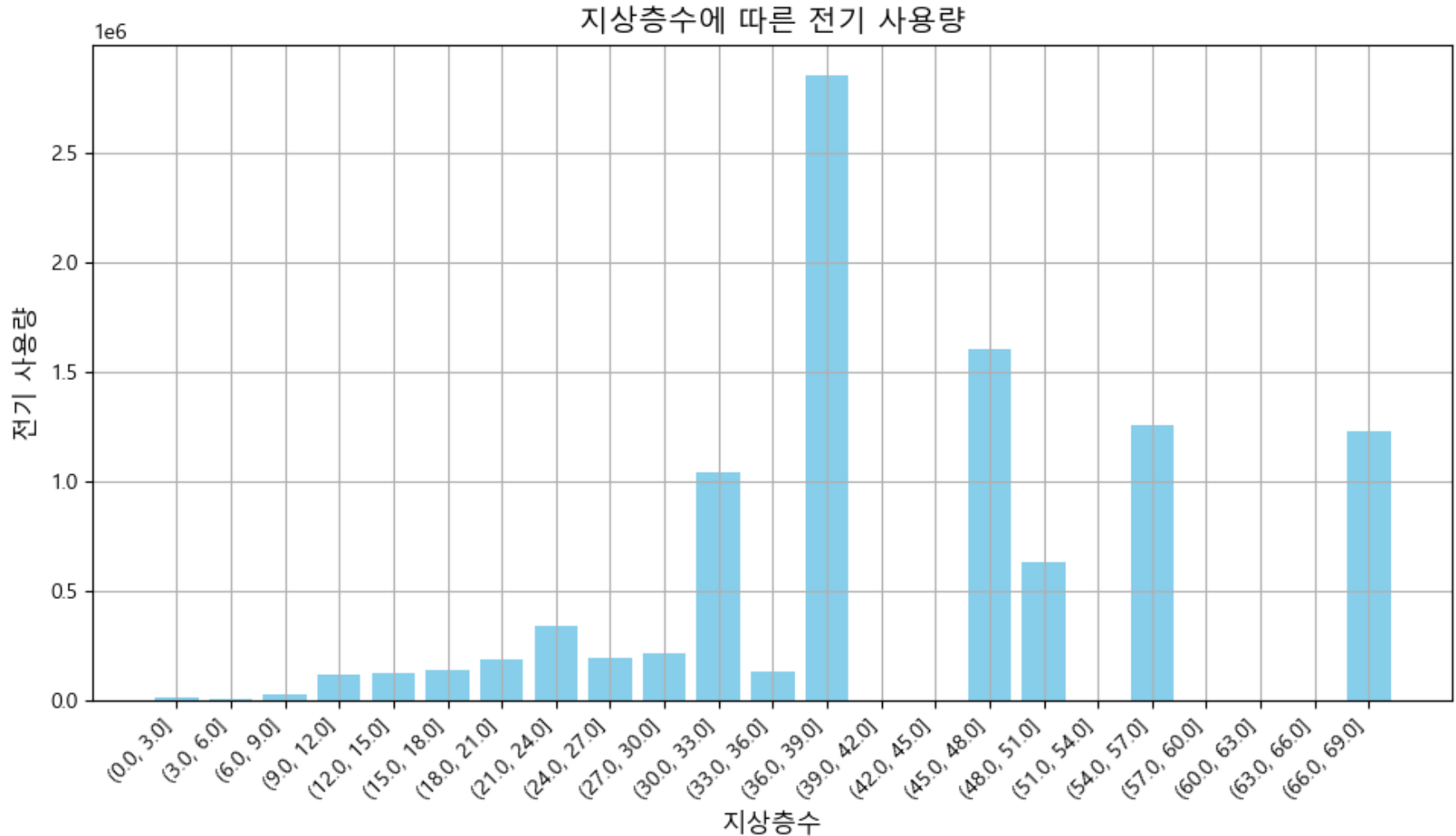


❖ 건물 용적률이 크면 전기 사용량이 높을 것이다.

용적률에 따른 전기 사용량

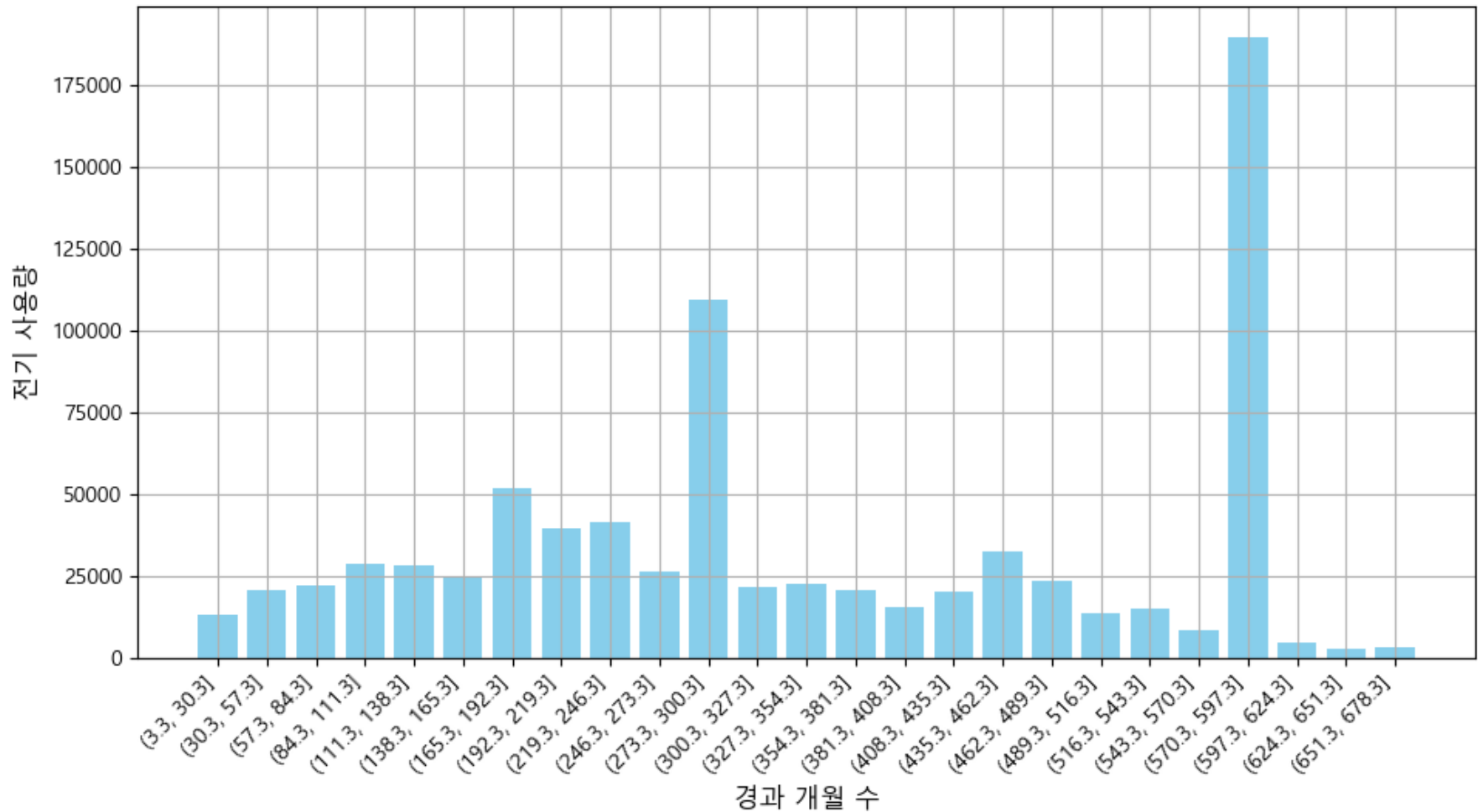


❖ 건물 층수가 높으면 전기 사용량이 높을 것이다.

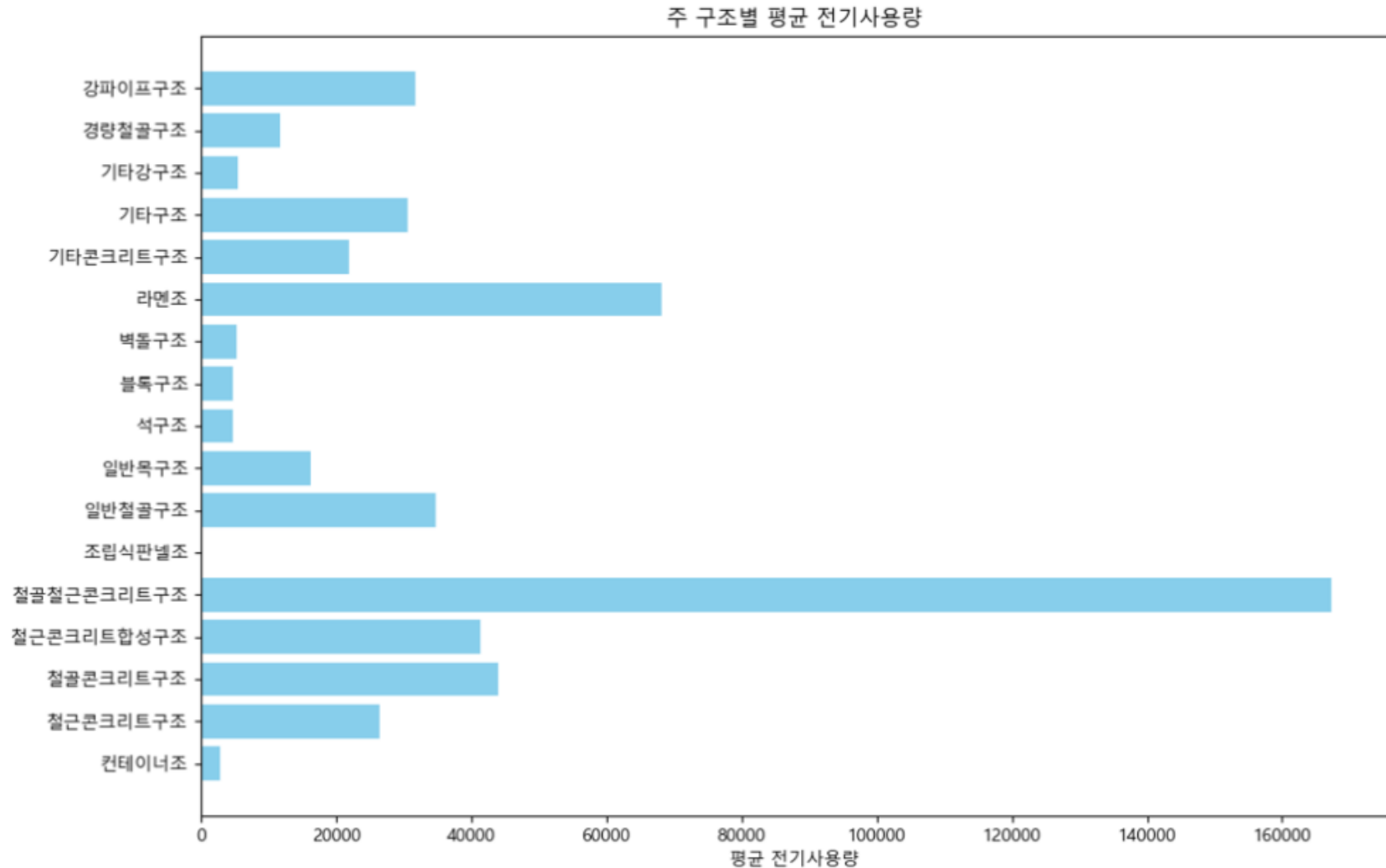


❖ 건물 노후도가 크면 전기 사용량이 낮을 것이다.

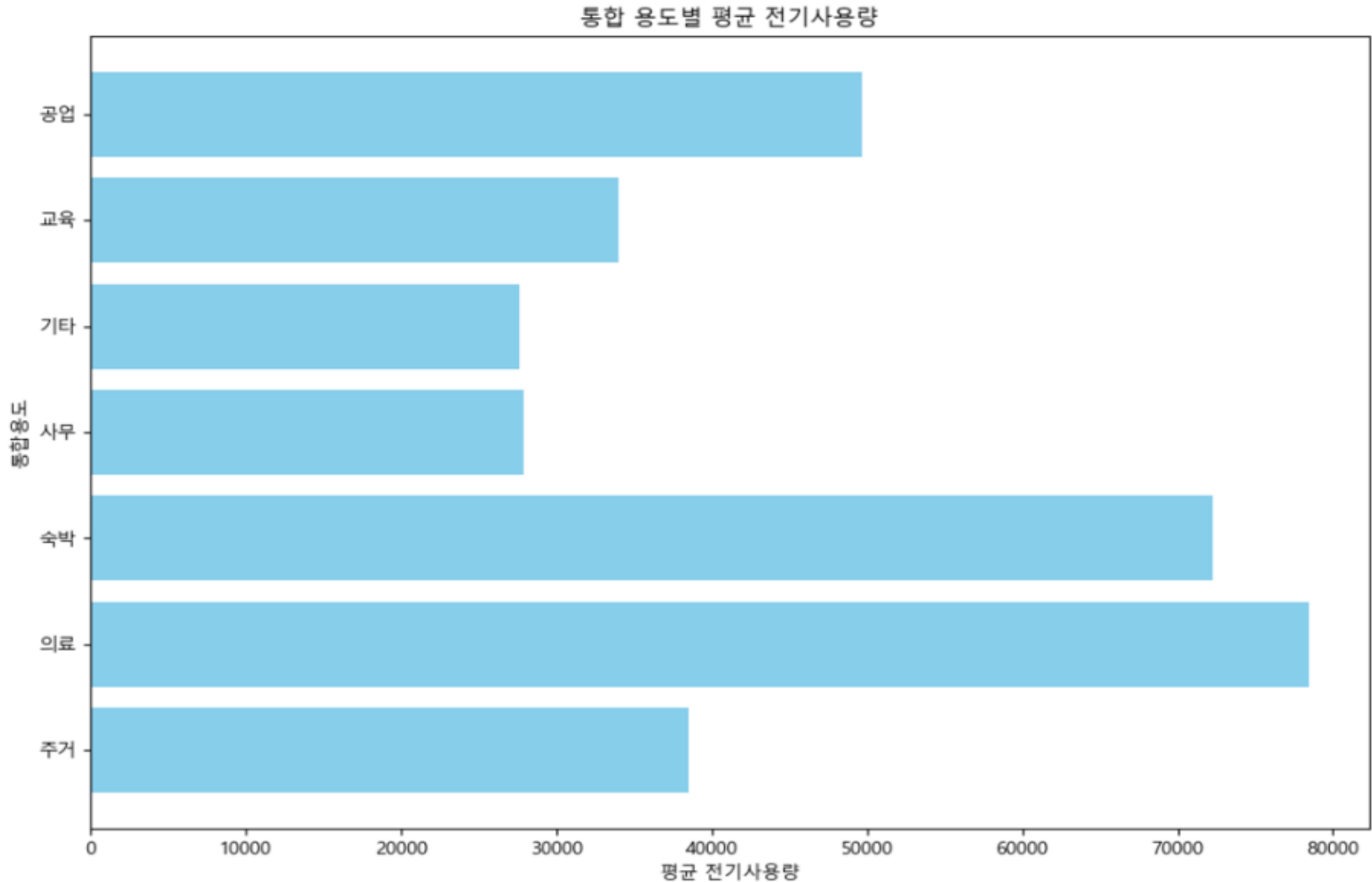
경과 개월 수에 따른 전기 사용량



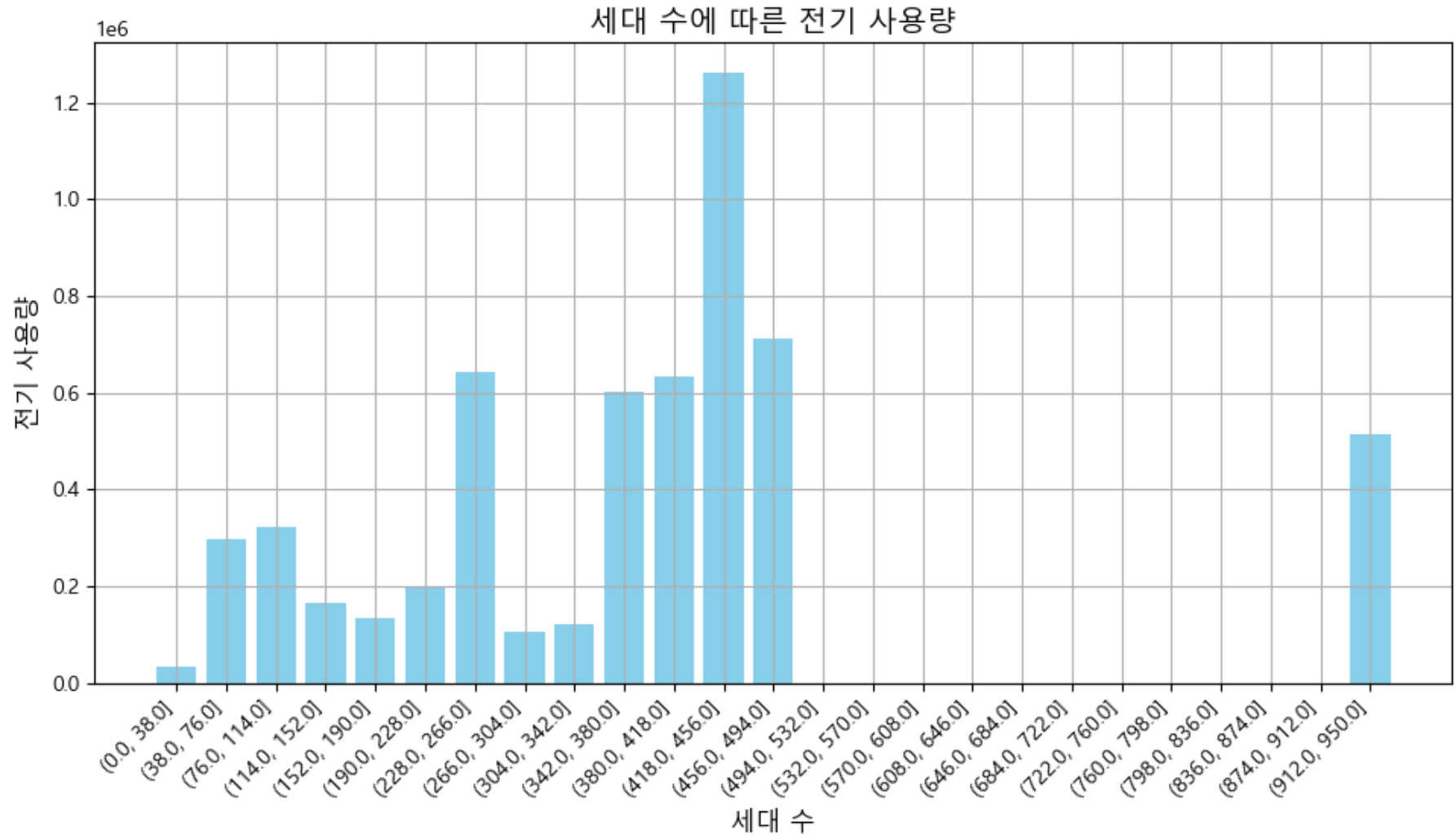
❖ 목구조가 철근콘크리트구조보다 전기 사용량이 낮을 것이다.



❖ 공업 시설이 다른 시설보다 전기 사용량이 높을 것이다.



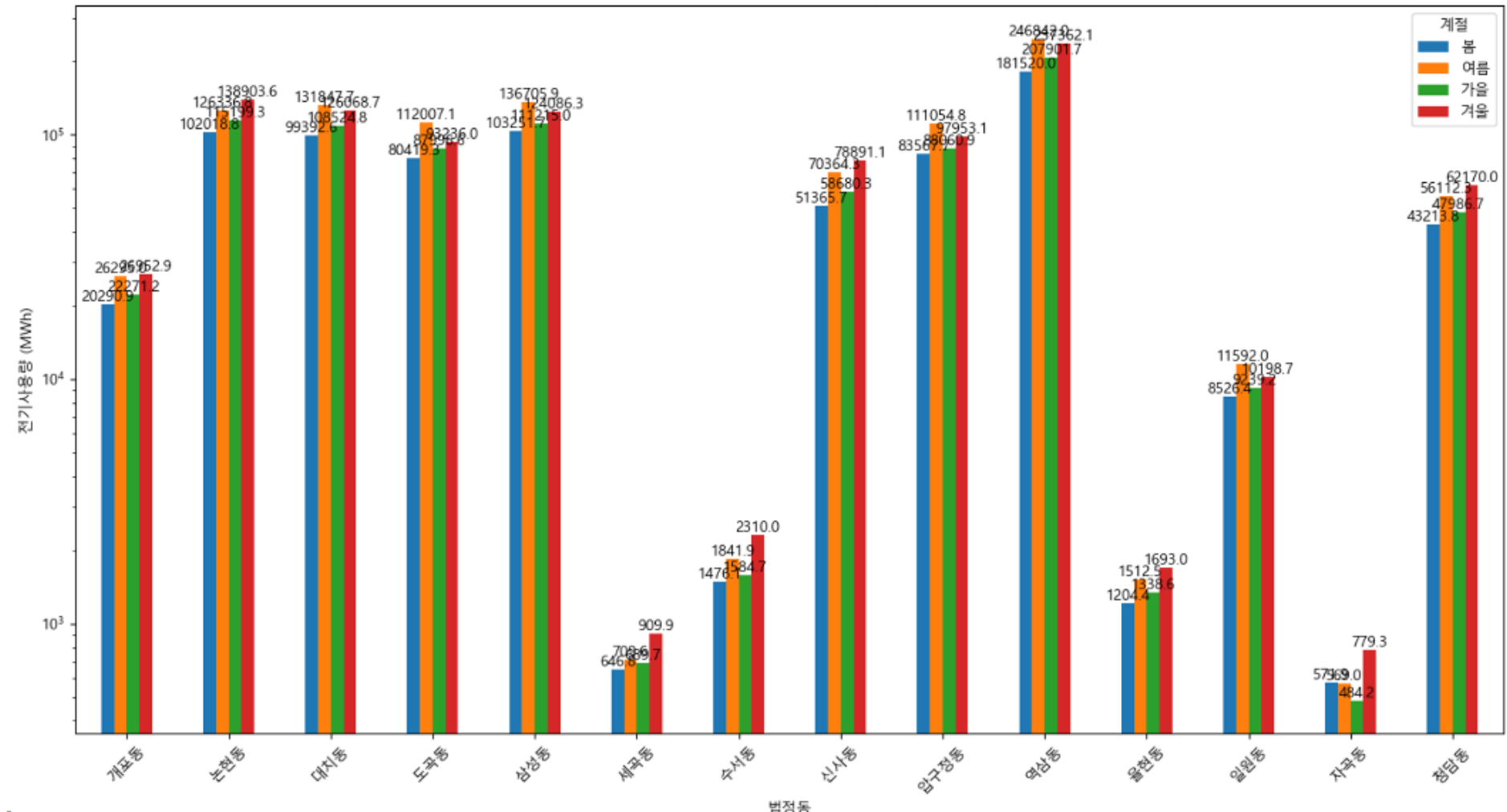
❖ 세대수가 많으면 전기사용량이 높을 것이다.



# 분석 결과

❖ 여름과 겨울은 봄, 여름보다 전기 사용량이 높을 것이다.

강남구 행정동별 계절별 전기사용량





## ❖ 건축물 대장 데이터

- ✓ 건축물 대장 데이터 주소형식(시도, 시군구, 법정동, 번, 지) -> 대지위치로 통합
- ✓ 건축 사용승인일 이용 건물 연식 컬럼 추가

## ❖ 전기 사용량 데이터

- ✓ 월에 2번 측정하는 경우 월별로 통합

## ❖ 건축물 대장 데이터와 전기 사용량 데이터 조인

- ✓ 건축물 대장 데이터와 전기 사용량 데이터 대지위치로 결정
- ✓ 200세대 미만 전기 사용량 개인정보보안으로 미제공 -> 데이터 조인 nan값이 생김
- ✓ 통합 후 전기사용량 nan값 -> 해당 행 삭제

## ❖ 소감

데이터를 보고 분석하고 전처리, 통합, 인사이트 도출까지 전 과정을 경험해보면서 유의미한 정제 데이터 하나 만드는 것이 매우 까다롭고 어렵다는 것을 느꼈고 얼마나 중요한 과정인지 알게 된 것 같다.

데이터를 수집하는 과정도 쉽지는 않았지만 데이터를 정제하는 과정이 매우 까다롭다고 느꼈습니다. 또 분석한 결과가 유의미한지 고민하는 것도 어렵다고 느꼈습니다.

## ❖ 결론

- ✓ 프로젝트를 통해 건물의 다양한 특성과 전기사용량 간의 관계를 분석하여 계절별, 월별, 주구조별, 주용도별, 법정동별 등 다양한 특징별 전기사용량 분석을 통해 에너지 절약 방안 및 전력 수급 계획을 마련할 수 있습니다. 더 많은 데이터와 정교한 분석을 통해 보다 구체적인 에너지 절약 방안을 도출할 수 있을 것입니다.