



서울시 공공자전거 따릉이 EDA



## [분석 개요] 분석 기준

### 시간적 변수

- 시간대별 이용건수
- 시간대별 평균이용시간
- 평일주말 별 이용빈도
- 평일주말 별 평균이용시간
- 월별 이용빈도
- 월별 평균이용시간

### 지리적 변수

- 자치구별 이용건수
- 자치구별 평균이용시간&거리
- 자치구별 신규 대여소 설치 기준
- 자치구별 이용자 연령대 분포
- 자치구별 이용권 유형 분포

Plus) 대여소 특성 지도 시각화

### 이용자 특성 변수

- 이용자 연령대 분포
- 연령대-시간대 별 이용건수
- ‘슬로우 유저’ 이용 특성

### 기상 변수

- 여름/겨울 온도에 따른 이용건수
- 여름/겨울 온도에 따른 평균이용시간
- 강수량에 따른 이용건수
- 강수량에 따른 평균이용시간

각 기준에 따라 살펴본 값은 대표적 값은 ‘이용건수’, ‘이용시간’ 임

- ➔ “어떤 조건에서 많이(적게) 이용하는가?”
- ➔ “어떤 조건에서 어떻게 이용하는가?”

## [분석 개요] 전처리 및 참고사항

### 1. 이상치 제거

- 속력(이용거리/이용시간) 시속 30km 초과 데이터포인트 제거
- 이동거리가 0인 데이터 포인트 제거

### 2. 해석상 편의를 위한 처리

- 시간대를 기준으로 분석(시각화)할 때 0시~23시가 아닌 '오전4시대~정오~오전3시대' 순으로 정렬
- 이른 새벽 → 심야 순으로 이용건수의 변화 흐름을 더 잘 파악하기 위함

### 3. 성별을 제외

- 이용정보 테이블에서 '성별' 속성의 결측치가 절반 가량
- 결측 값을 제외하면 전체 데이터에 대한 대표성이 떨어진다고 보아 '성별'을 기준으로 한 분석을 수행하지 않았음



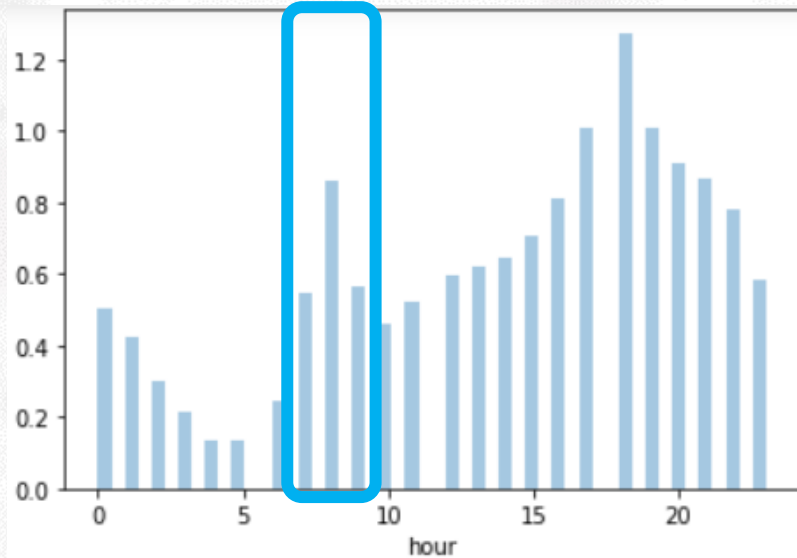
# 01. 시간적 변수



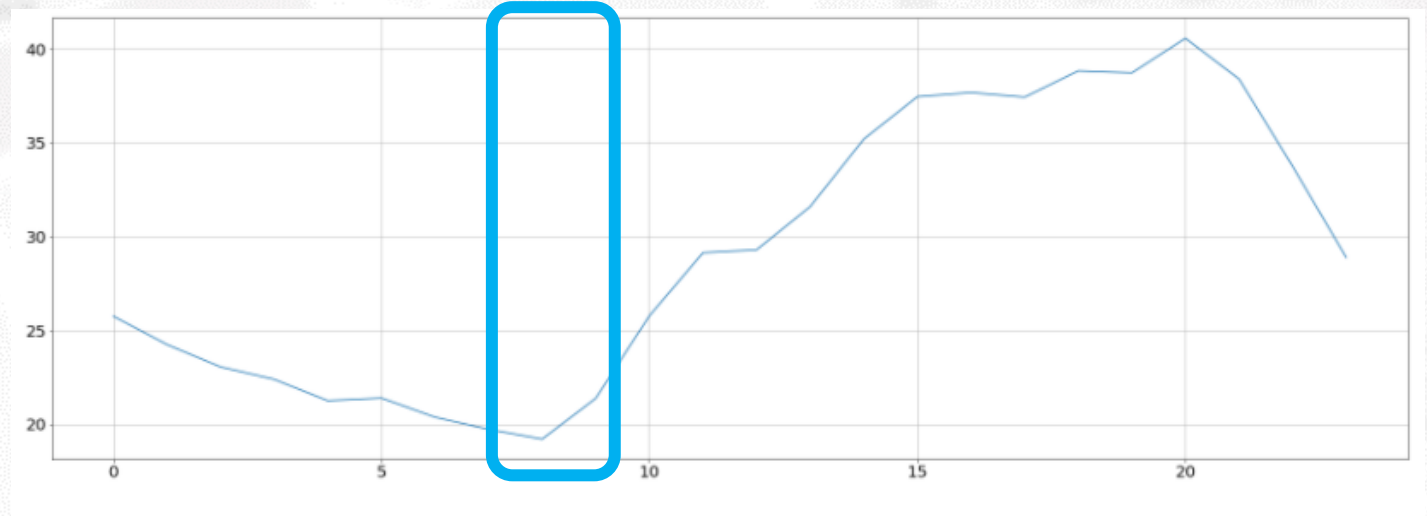
# 시간별 이용 특성

Enjoy your stylish business and campus life with BIZCAM

〈시간별 이용 빈도〉



〈시간별 평균 이용 시간〉



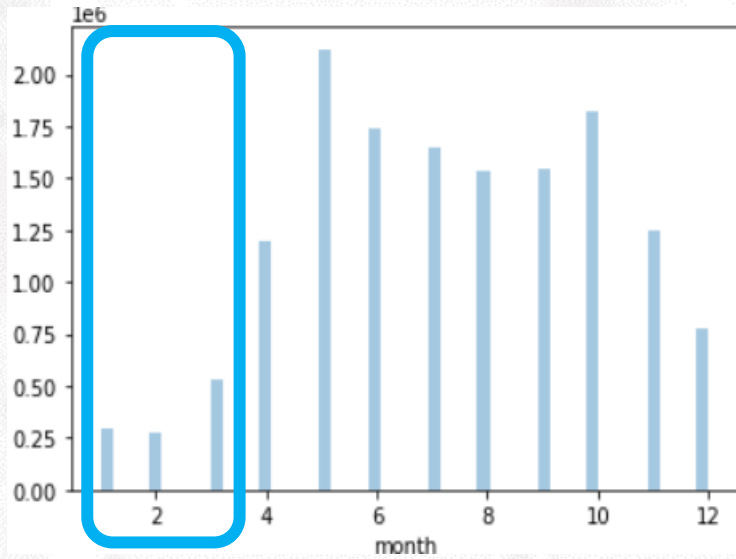
- 출근 시간에 따릉이 이용 빈도 많아지지만, 평균 이용 시간은 적음



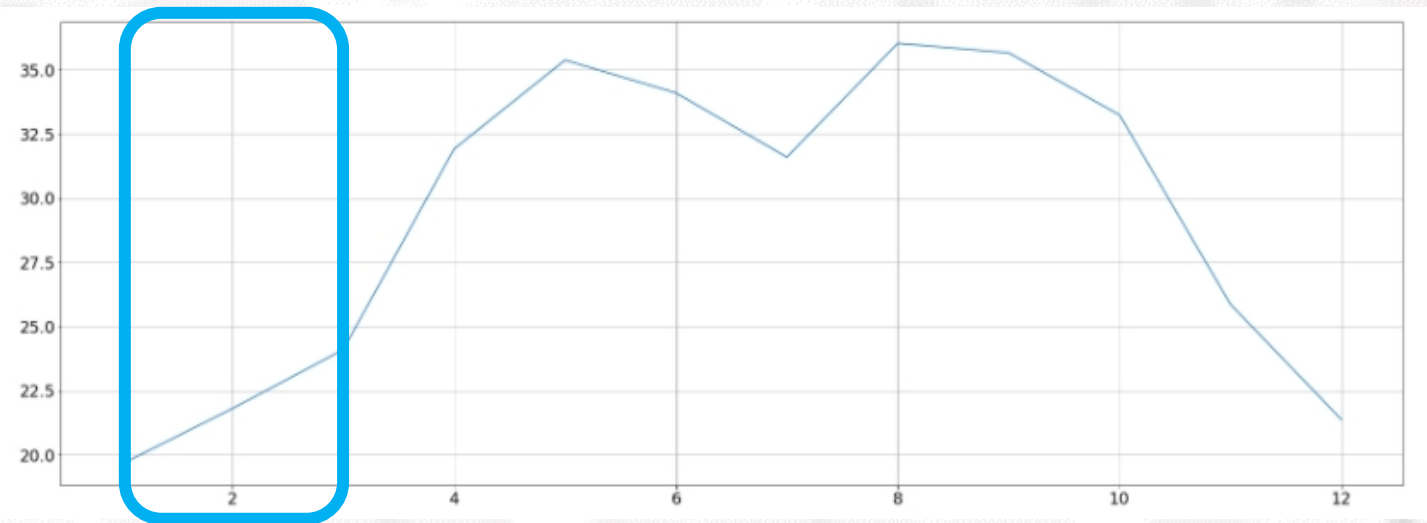
## 월별 이용 특성

Enjoy your stylish business and campus life with BIZCAM

〈월별 이용 빈도〉



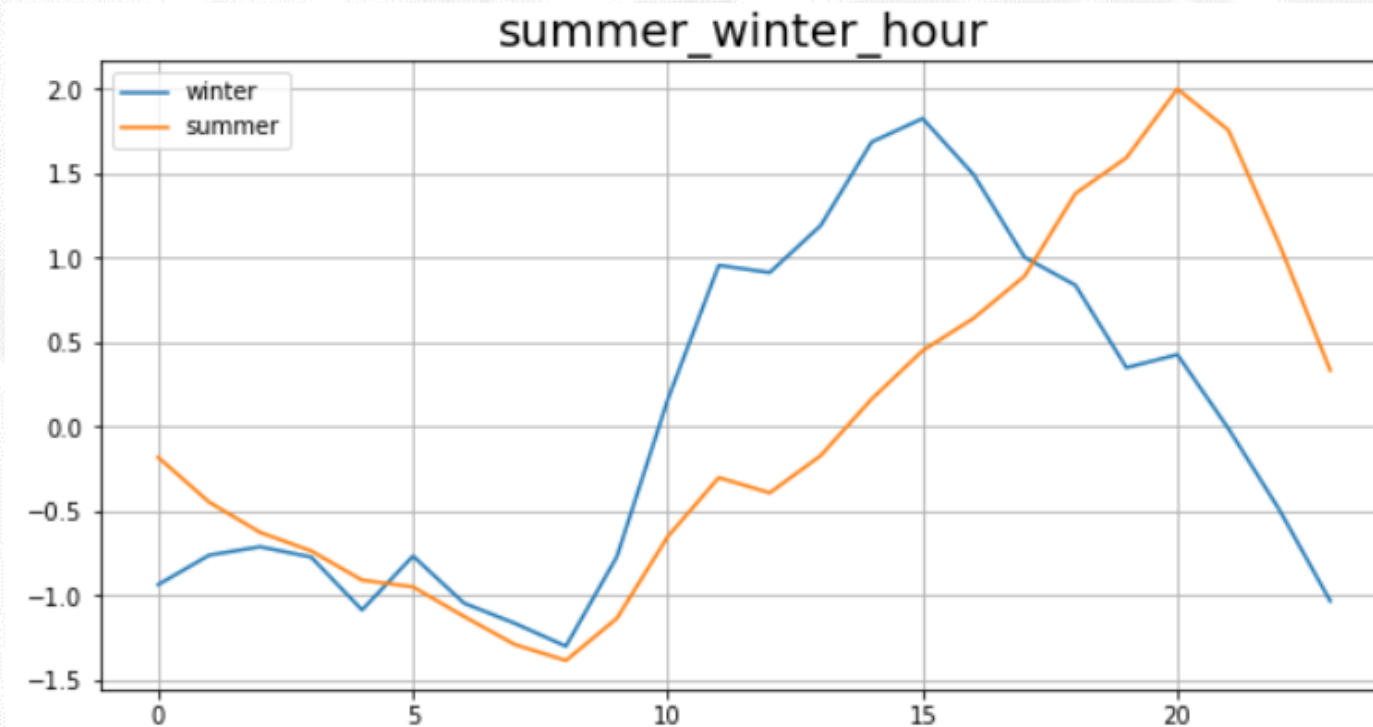
〈월별 평균 이용 시간〉



- 겨울에 특히 이용 빈도와 평균 이용 시간이 적음

## 계절 및 시간별 이용 특성

〈계절/시간대 별 이용 시간〉



- 겨울 대비 여름에, 저녁시간 따릉이 이용 비율이 더 높음



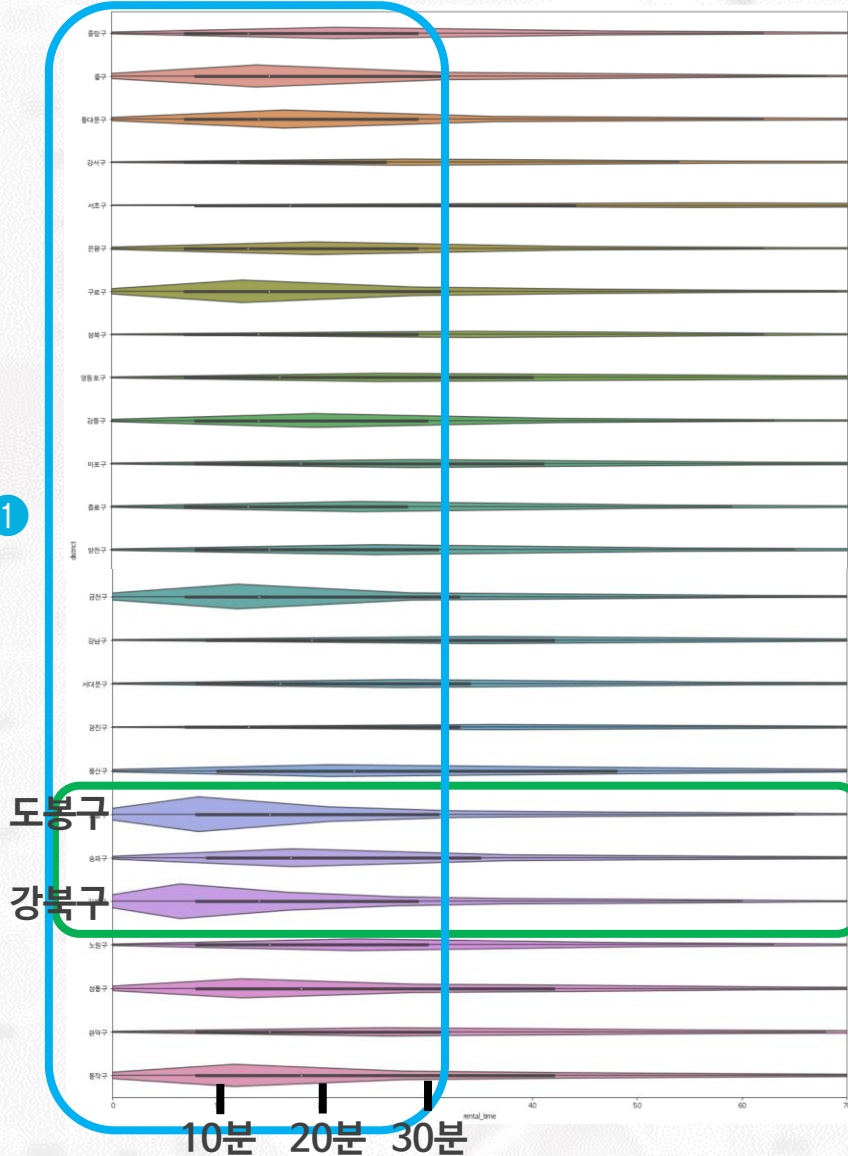
## 02. 지리적 변수





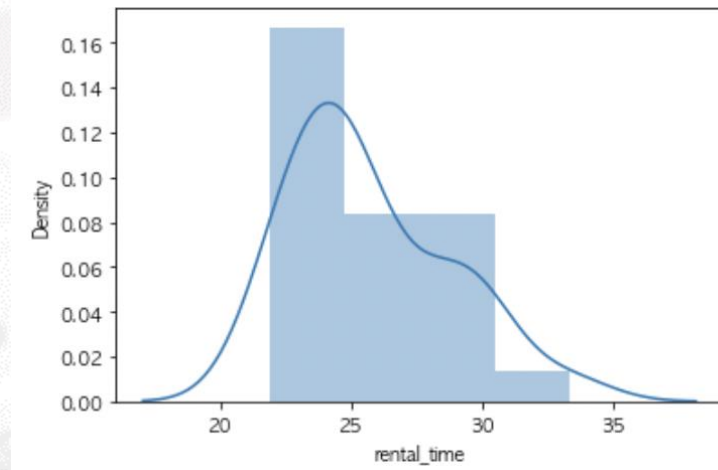
# 자치구별 이용 시간 분포의 차이

1



2

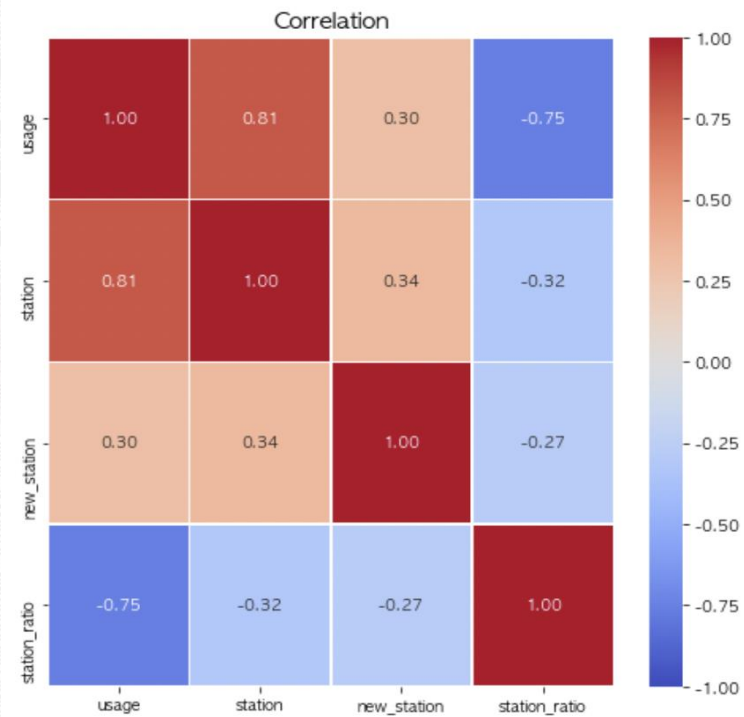
- 자치구별 1회 대여 시 평균 이용 시간은 24~25분



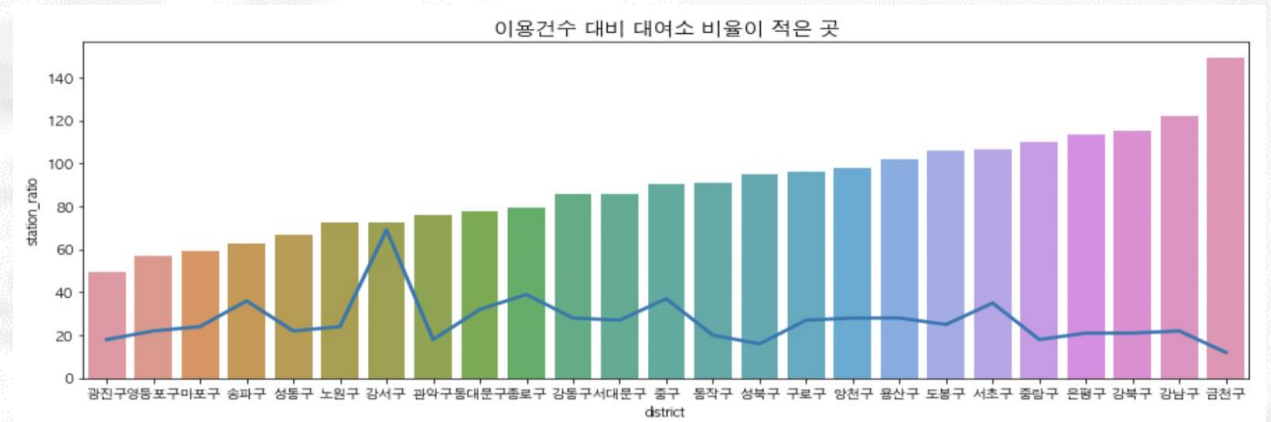
- 1회 대여 시 이용 시간의 분포에 대해서는 구별로 차이를 보임

- ① 자치구별 이용자들의 대부분은 1회 대여 시 30분 미만으로 이용함
- ② 도봉구, 강북구는 1회 대여 시 10분 미만으로 짧게 타는 이용자의 비율이 높음
- ③ ①, ②를 제외한 자치구에서는 1회 대여 시 이용시간에 대한 이용자의 분포가 비교적 일정함

## 자치구별 신규 대여소 설치 기준



- 이용 건수와 기존 대여소(19년 기준) 수는 비교적 강한 상관관계를 보임

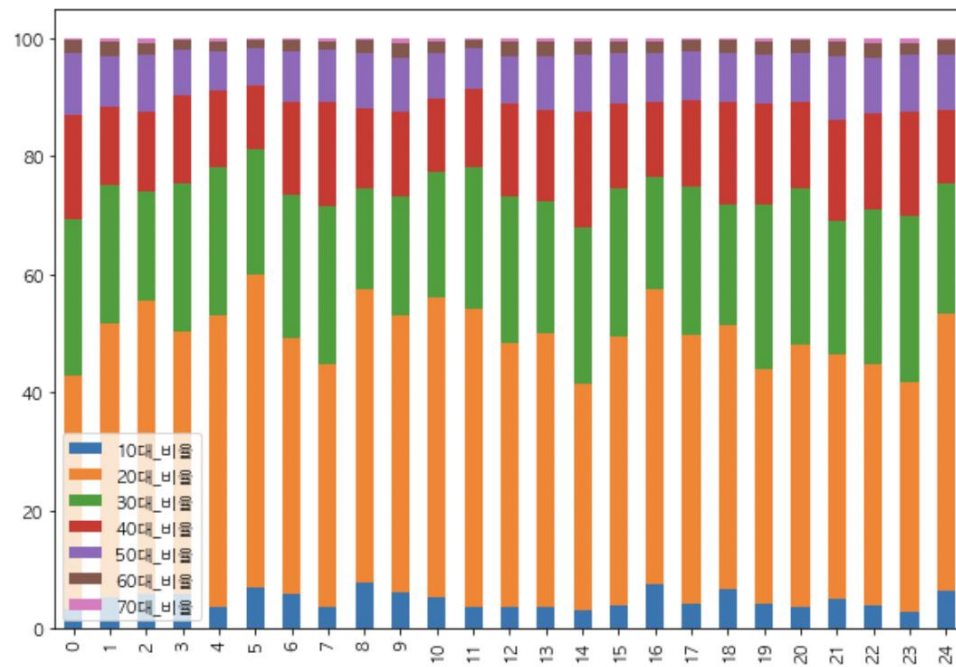


- 기존 대여소 수와 이용 건수는 서로 유의미한 영향을 끼치는 것으로 나타남
- 이를 근거로 “이용 건수에 비해 대여소 수가 적은 지역에 대여소를 추가 설치하는 경향이 있을 것” 이라고 예측
- 막대 그래프는 이용 건수 대비 대여소 비율, 꺾은선 그래프는 신규 설치 대여소 수
- 시각화 결과 신규 대여소 설치에 대한 유의미한 관계성을 찾기 어려움



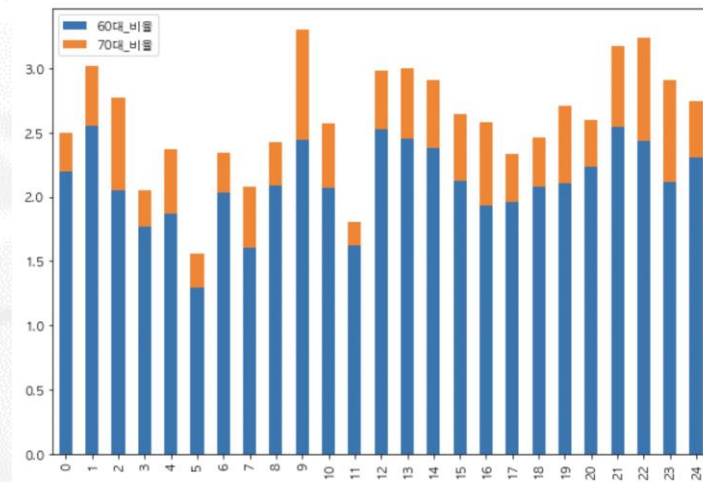
# 자치구/연령대별 이용자 비율

〈자치구별 이용자 연령대 비율〉

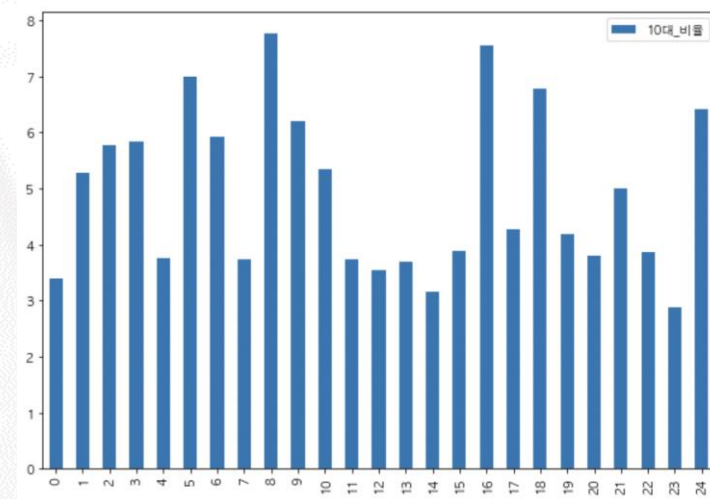


- 청년, 중년층의 비율은 자치구마다 비슷하게 나타남
- **도봉구, 종로구, 은평구**는 **노령층(6, 70대)** 이용률이 타 지역보다 높은 편이며, 광진구, 동작구는 타 지역보다 적은 편
- **노원구, 성북구, 광진구**는 **청소년층(10대)** 이용률이 타 지역보다 높은 편이며, 중구, 서초구, 강남구는 타 지역보다 적은 편

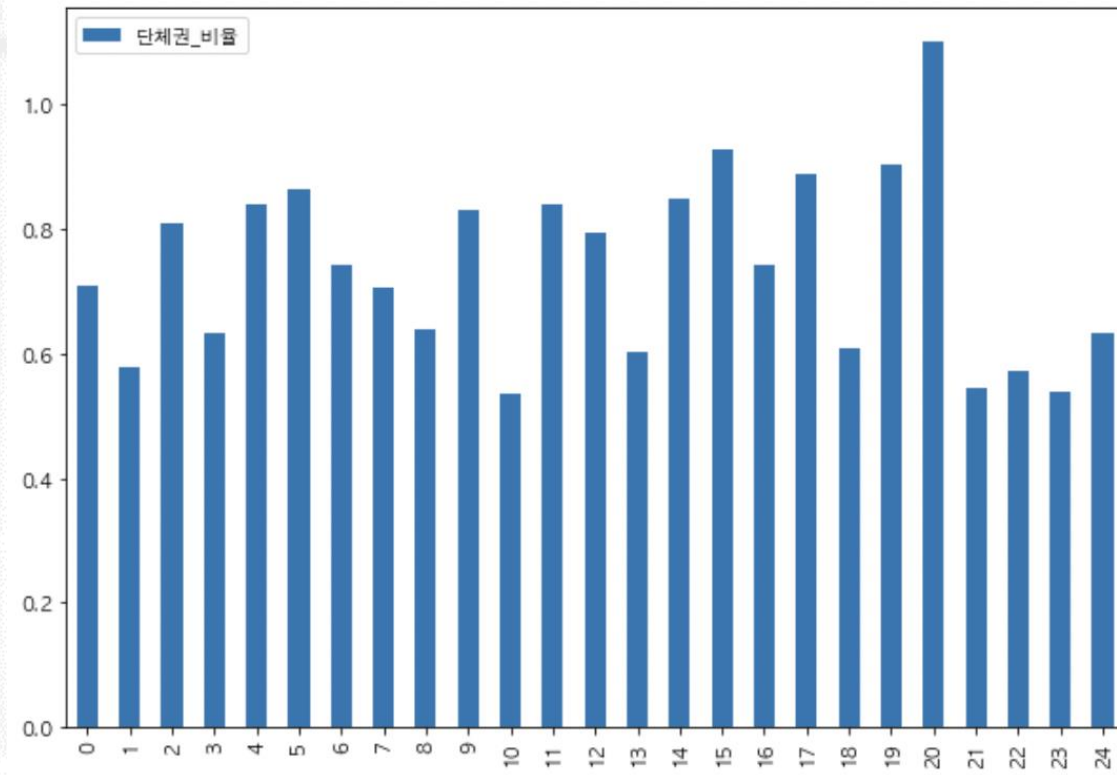
〈자치구별 노령 이용자 비율(60,70대)〉



〈자치구별 청소년 이용자 비율(10대)〉



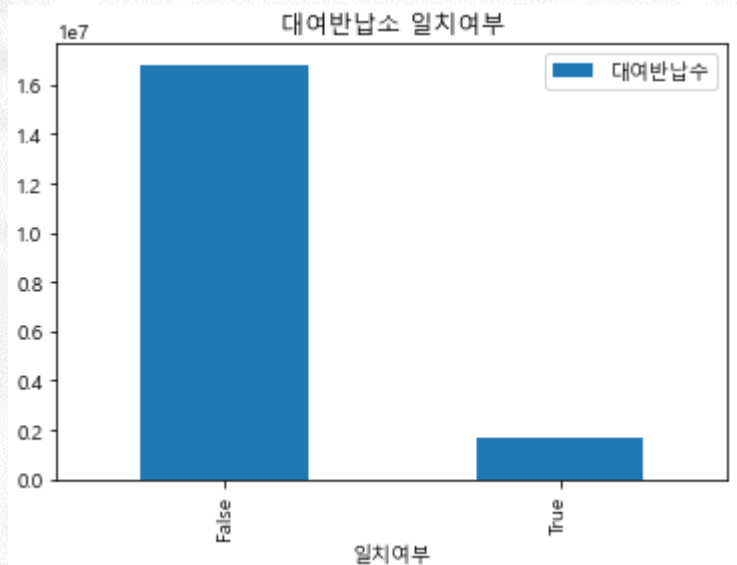
## 자치구별 이용권 유형



- 이용권 내 정기권, 일일권, 단체권 세가지 유형 존재
- 이 중, 단체권 이용 비율을 자치구별로 나타냄
- 자치구 중, 단체권 비율이 높은 구는 용산구, 성동구, 영등포구이며 단체권 비율이 가장 낮은 구는 동대문구, 중구, 은평구임
- 단체권은 일상적 용도 보다, 함께 즐기는 레저 용도가 강함  
용산구, 성동구, 영등포구 모두 한강공원이나 서울숲 등이 위치



## 대여 후 동일한 곳에 반납하는가?



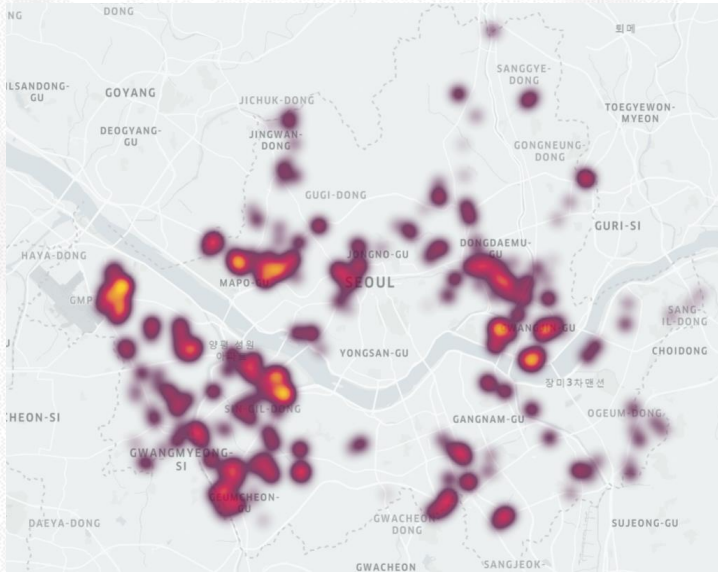
- 같은 곳에서 대여 반납을 하는 것보다 다른 곳에서 대여 반납을 하는 경우가 10.18배 더 많음
- 평균 이용 시간은 같은 곳에서 대여 반납이 이루어지는 곳은 51분, 다른 곳에서 이루어지는 곳은 23분으로, 같은 곳에서 대여 반납 할 경우 평균 이용 시간이 두배 가량 김

	rental_loc_name	이용횟수	rental_time
470	독성유원지역 1번출구 앞	27872	57.883180
1106	여의나루역 1번출구 앞	27388	66.495071
819	서울숲 관리사무소	13456	58.816662
160	고속터미널역 8-1번, 8-2번 출구 사이	13025	75.245528
522	마포구민체육센터 앞	12407	58.230918
664	봉림교 교통섬	10342	55.995455
318	노들역 1번출구	8621	62.616402
1178	옥수역 3번출구	8490	56.817668
364	당산육갑문	8100	61.698889
6	CJ 드림시티	8025	71.581059

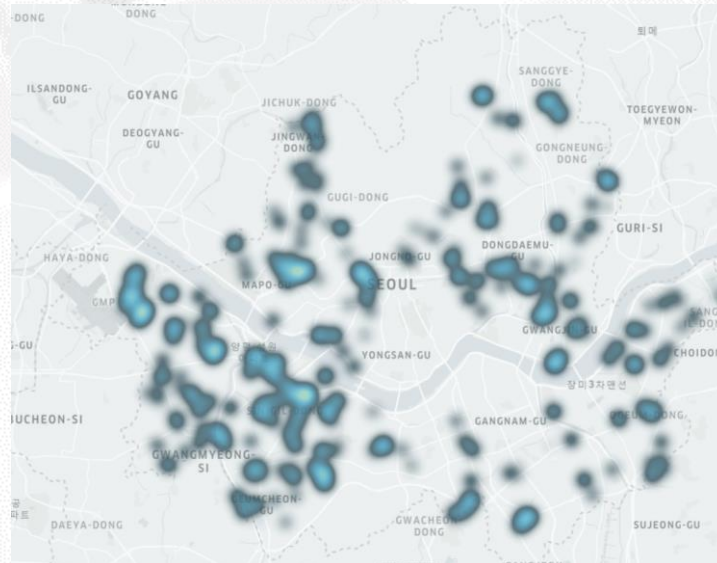
- 같은 대여소에서 대여 반납이 이루어지는 곳은 대부분 공원, 한강 주변으로 나타남

# 대여소별 출근/퇴근/야간 시간대 대여비율

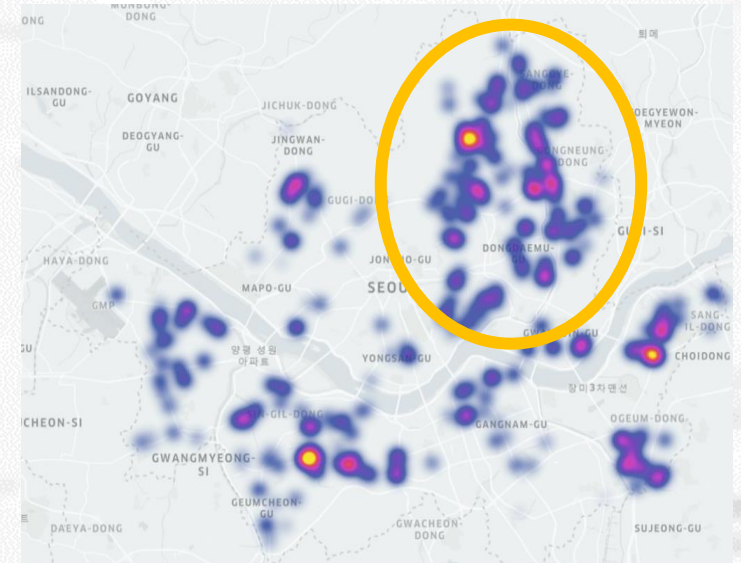
〈출근 시간대 대여 비율 상위 300〉



〈퇴근 시간대 대여 비율 상위 300〉



〈야간 시간대 대여 비율 상위 300〉



- 출근 시간대 대여 비율이 높은 지역과 퇴근 시간대 대여 비율이 높은 지역이 비슷하게 나타남
- 야간 시간대에는 서울 동북권쪽 대여소의 대여 비율이 높게 나타나며, 출퇴근 시간대와의 차이를 보임

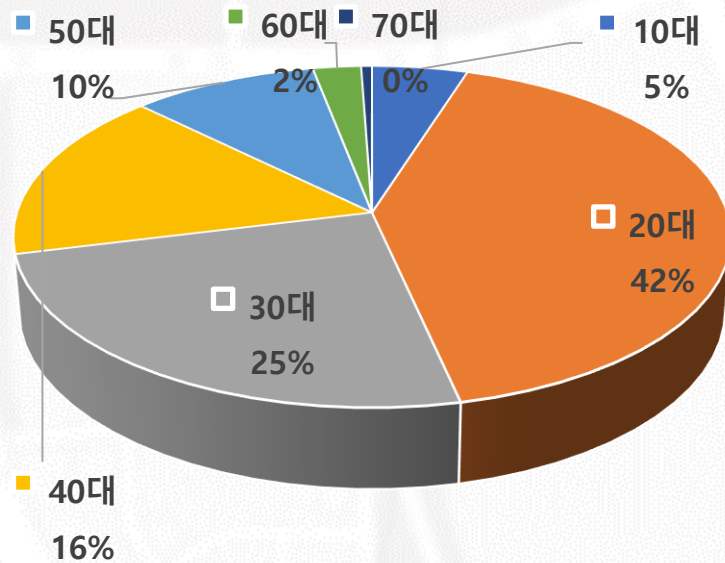


## 03. 이용자 특성 변수



## 연령대 별 이용량

- 이용량 순위: **20대** > 30대 > 40대 > 50대 > 10대 > 60대 > 70대
- 20~40대 이용량이 전체 이용량의 80% 이상 차지



- 고령층의 '디지털 소외' 문제 지적되어 왔음

뉴스 > 경제

### 노인을 위한 '따릉이'는 없다... "너무 어려워요"

기사입력 2019-08-04 19:30 | 최종수정 2019-08-04 20:19 | 0



자료: <https://www.mbn.co.kr/news/economy/3905195>

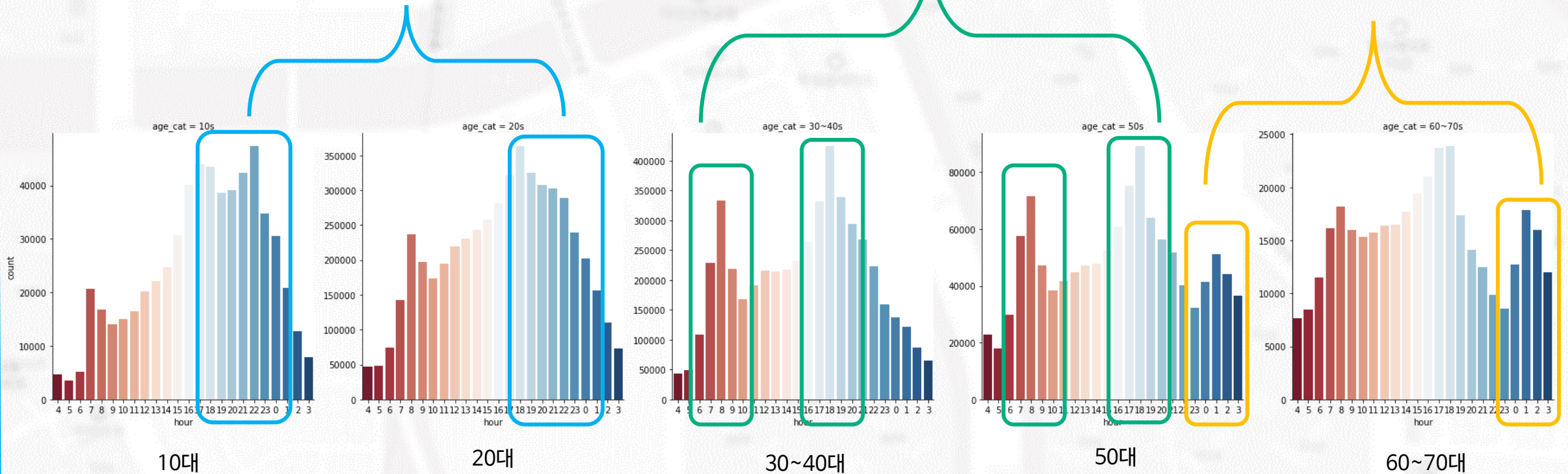


# 연령대-시간대별 이용 빈도

10~20대, “야간 이용 多”

30~50대, “오전 • 오후 피크가 뚜렷함”

50~70대, “심야 피크 존재”



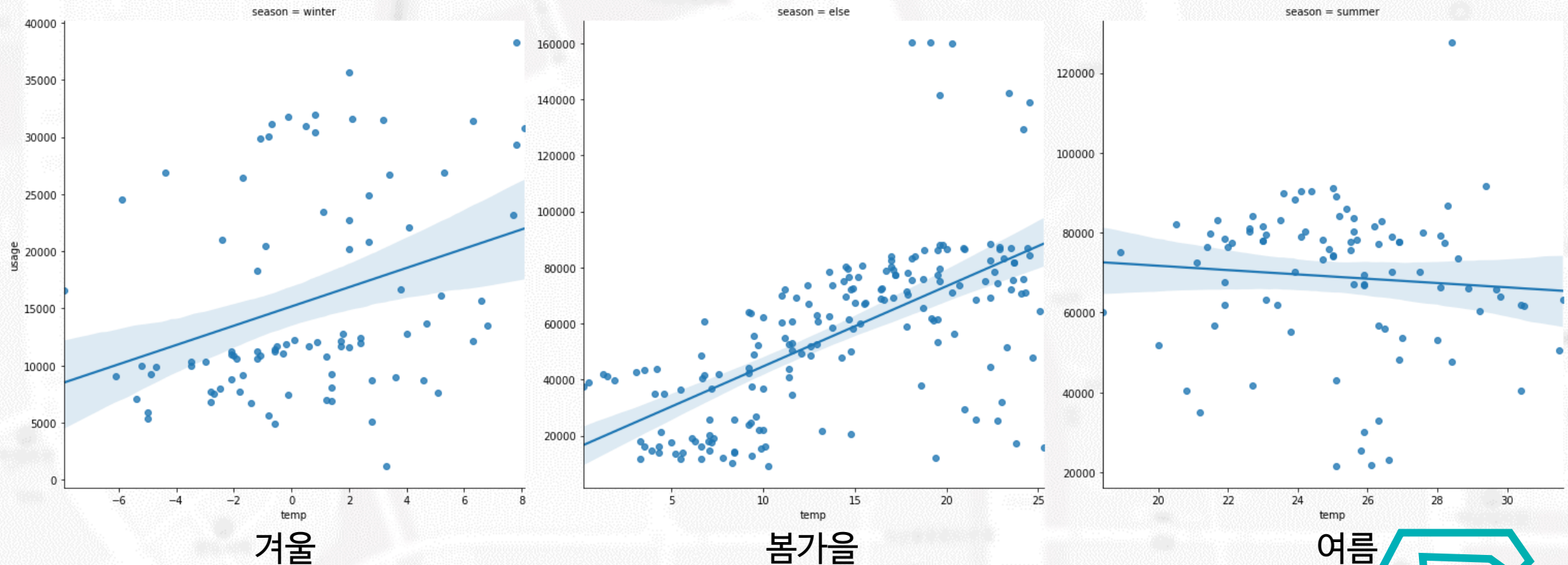
\*가로축: 4am ~ 정오 ~ 3am

## 04. 기상 변수



## 일평균 기온에 따른 일이용건수

- 겨울, 봄가을은 기온 상승에 따라 일이용건수 증가 → 따뜻하면 많이 탄다
- 여름은 기온 상승에 따라 일이용건수 감소 → 더우면 덜 탄다
- 회귀선 기울기의 절대값은 봄가을 > 겨울 > 여름 순 → 봄가을 이용이 온도에 더욱 민감하게 반응한다



# 느낀 점 및 향후 분석 방향





## 향후 분석 방향

1. 대여 이력 데이터를 하나의 네트워크로 보고 **네트워크 분석**(중심성 등) 분석을 시도해볼 예정
2. 선행연구조사를 통해 분석에 **유의미한 외부 변수**를 파악, 데이터를 수집하여 기존 데이터와 **연계 분석**을 시행