

26-1 도시환경 EDA 프로젝트

따릉이 운영 전략 제언

신지원, 이윤서, 이재원, 최현, 황수민

Contents

01

따릉이 서비스 개요

서비스 규모 및 비용 구조

02

따릉이 이용 행태 분석

날씨, 이용권, 연령, 시간대, 행정구역

03

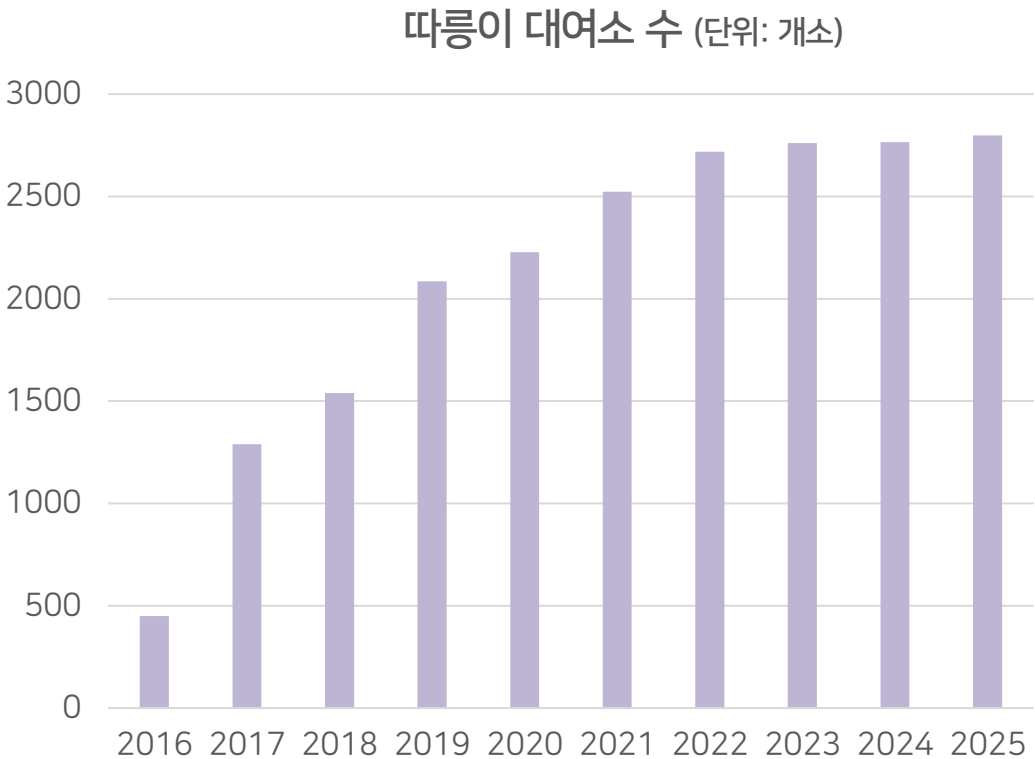
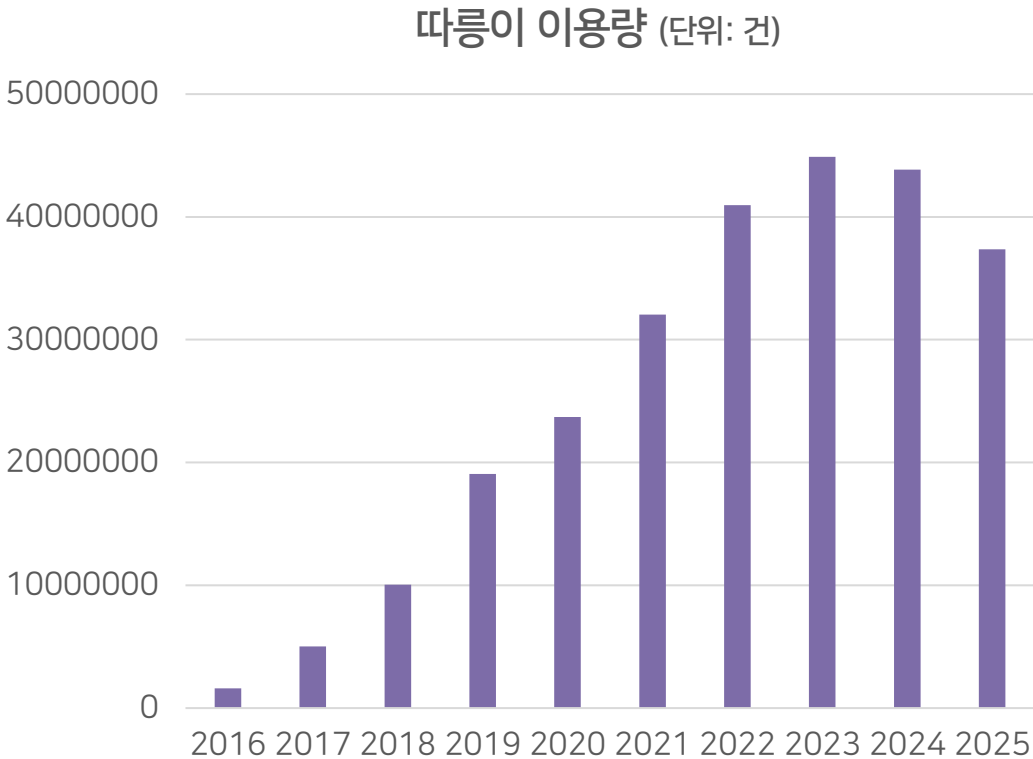
이용률 영향 요소 분석

다중선행회귀

따릉이 서비스 개요

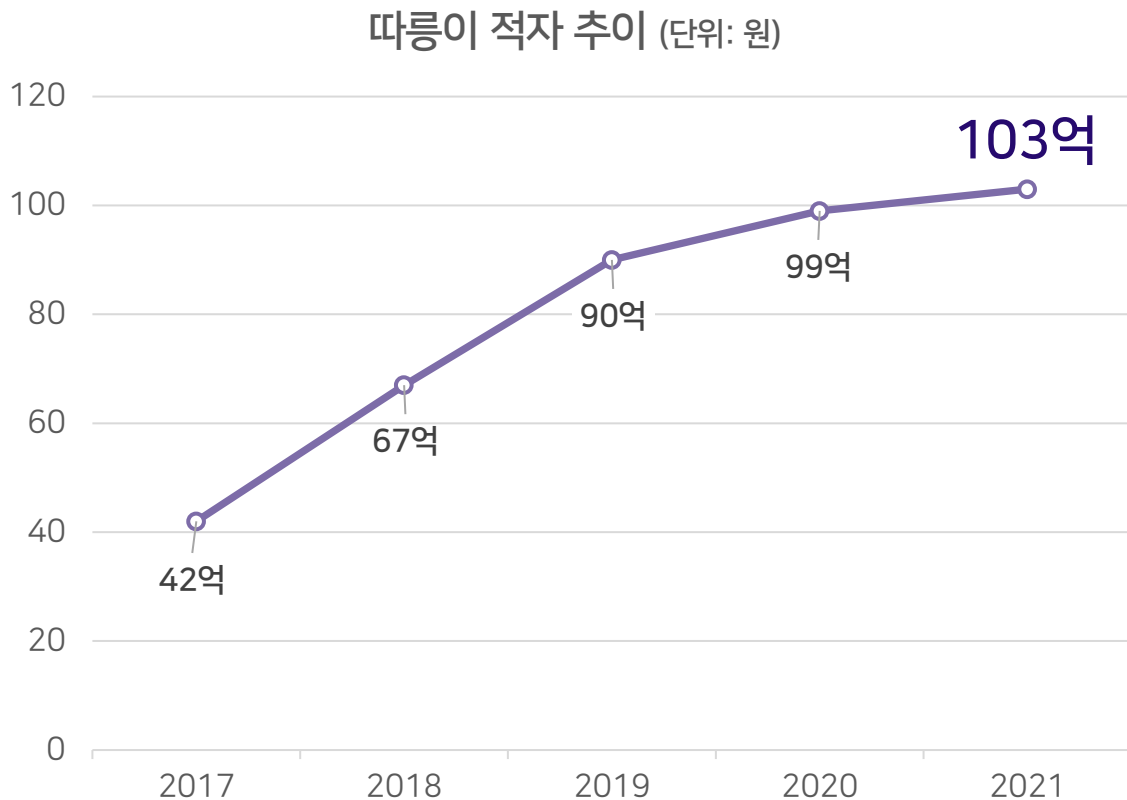
따릉이는 지난 10년간 꾸준히 서비스 확장을 이뤄 옴

연도별 따릉이 이용 통계



공공 교통 서비스 특성에 따른 구조적 적자가 매해 지속적으로 발생함

연도별 따릉이 적자 규모

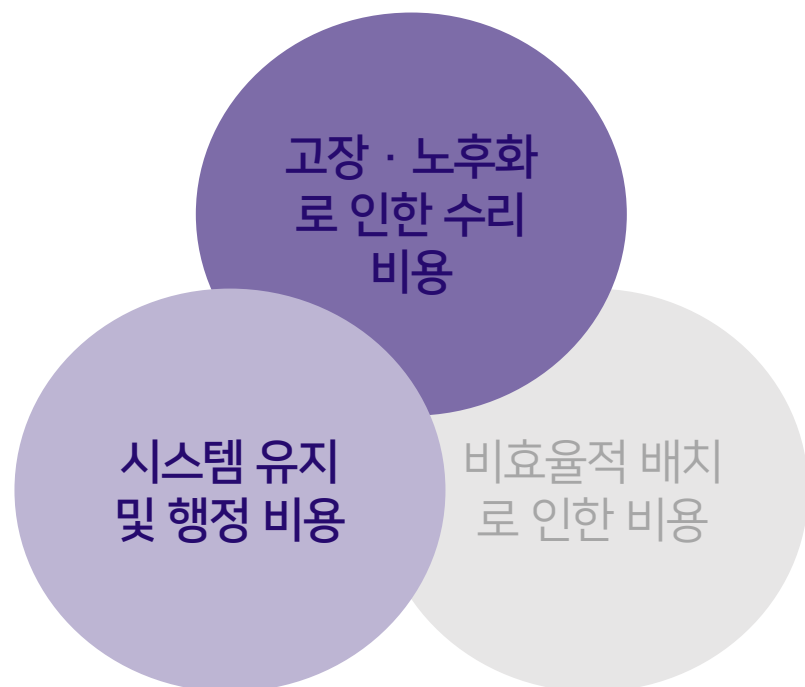


따릉이의 정책 목적은 수익 극대화가 아닌
사회적 편익 제공을 목적으로 운영되는 공공 교통 서비스

요금 정책과 접근성 유지로 인해
구조적인 적자 발생은 불가피

따릉이 적자 개선을 위해 비용 구조에 대한 **선택적 접근**이 필요함

따릉이 비용 구성 요소



고장 · 노후화로 인한 수리비용

자전거 및 단말기 구매, 대여소 구축비, 수선유지비(부품비), 보험료(종합/영조물), 차량 유지비

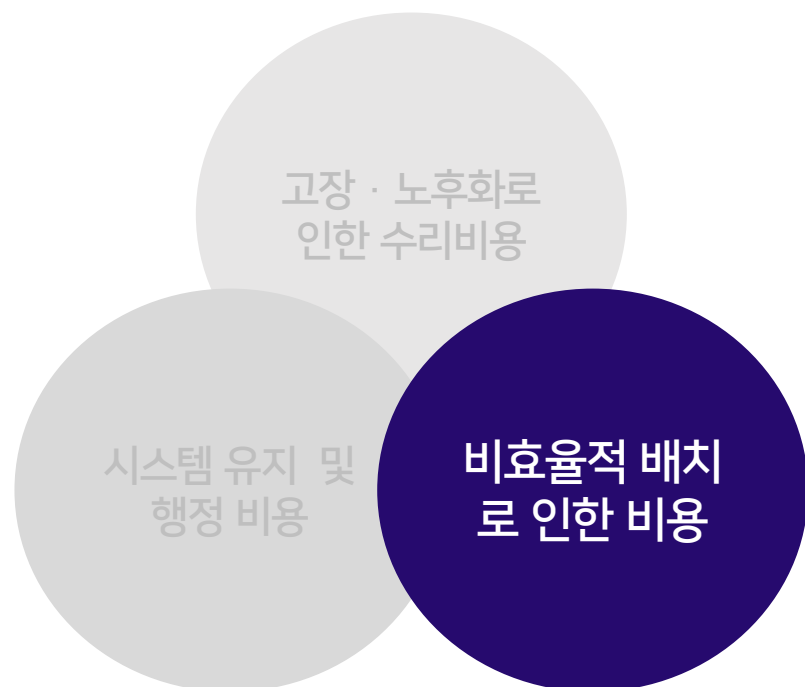
시스템 유지 및 행정 비용

대여소 및 단말기 통신비(LGU+, KT), 클라우드 이용료, 본인인증 및 문자 발송비

➔ 외생적인 요인이 크고 단기간에 개선하기 어려운 영역

따릉이 적자 개선을 위해 비용 구조에 대한 **선택적 접근**이 필요함

따릉이 비용 구성 요소



비효율적 배치로 인한 비용

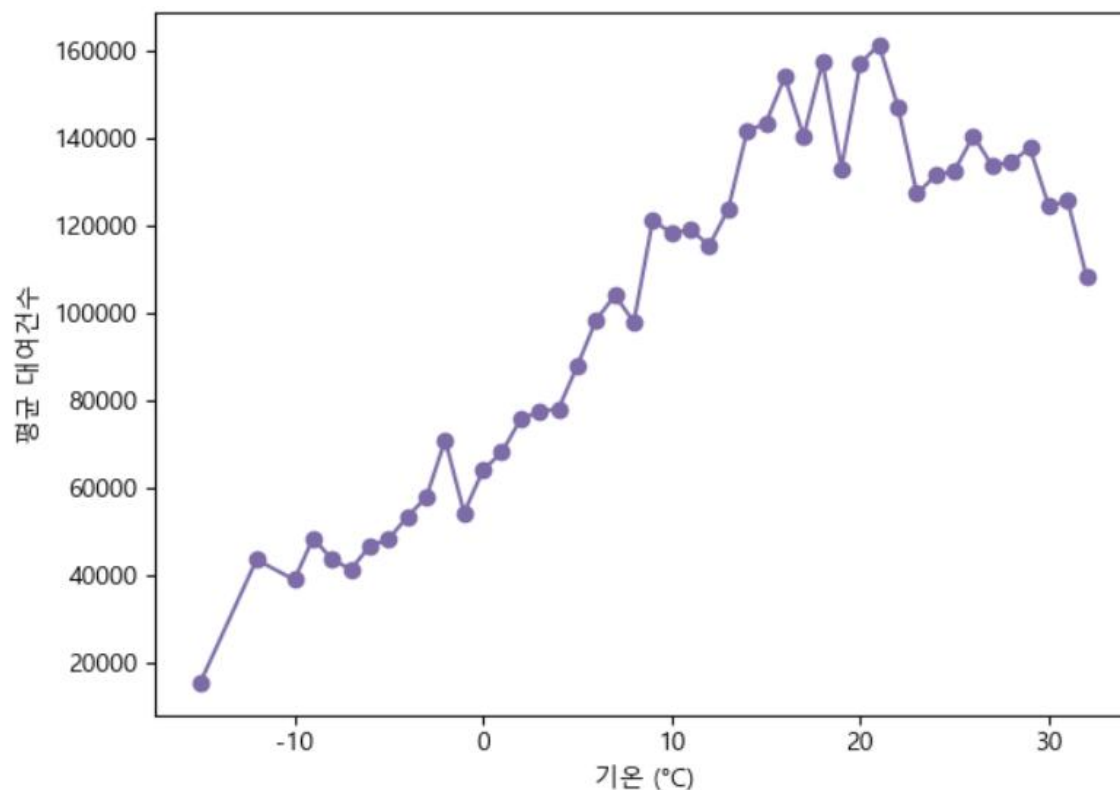
일반직, 공무원직(배송, 정비, 안내), 계약직 인건비 및 복리후생비

➔ 비용 구조 중 가장 지출이 많은 부분이며
운영 전략에 따라 조정 가능한 영역에 해당

따릉이 이용 행태

기온과 평균 대여 건수는 비선형적 관계를 가짐

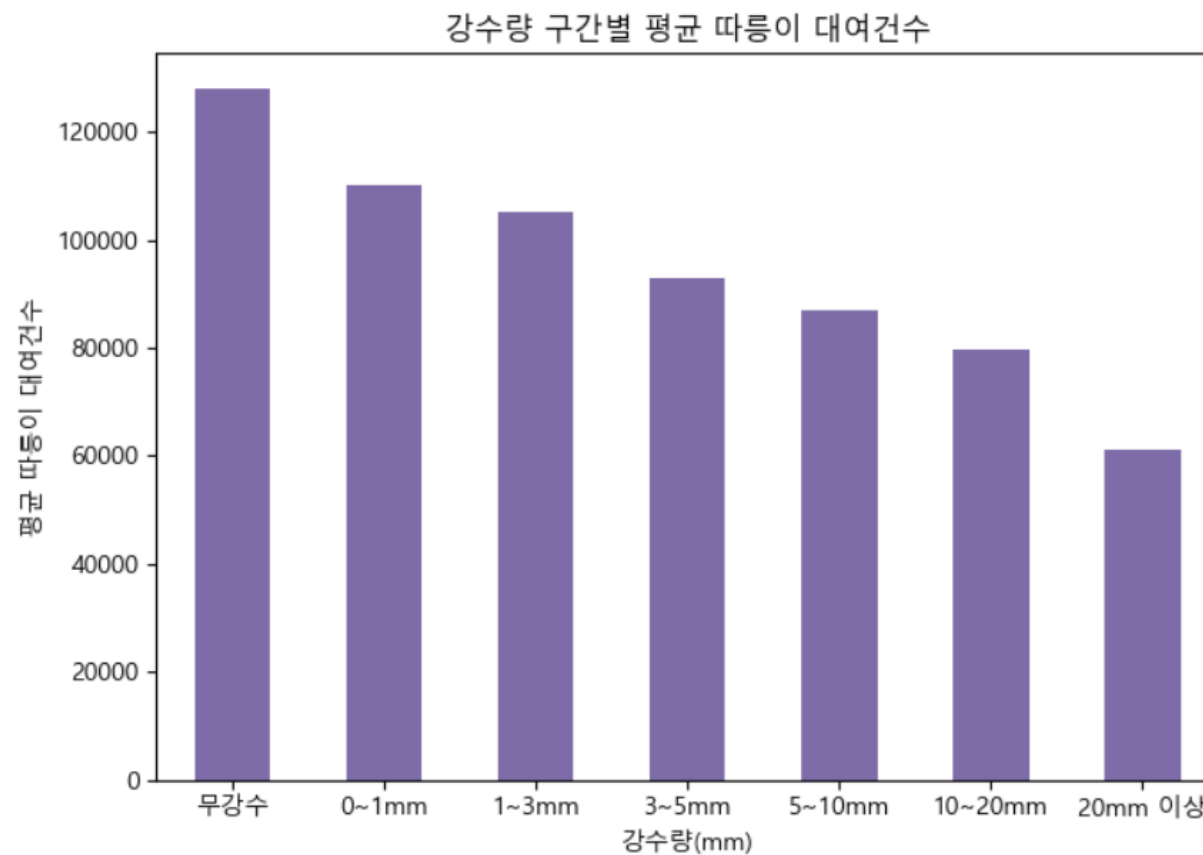
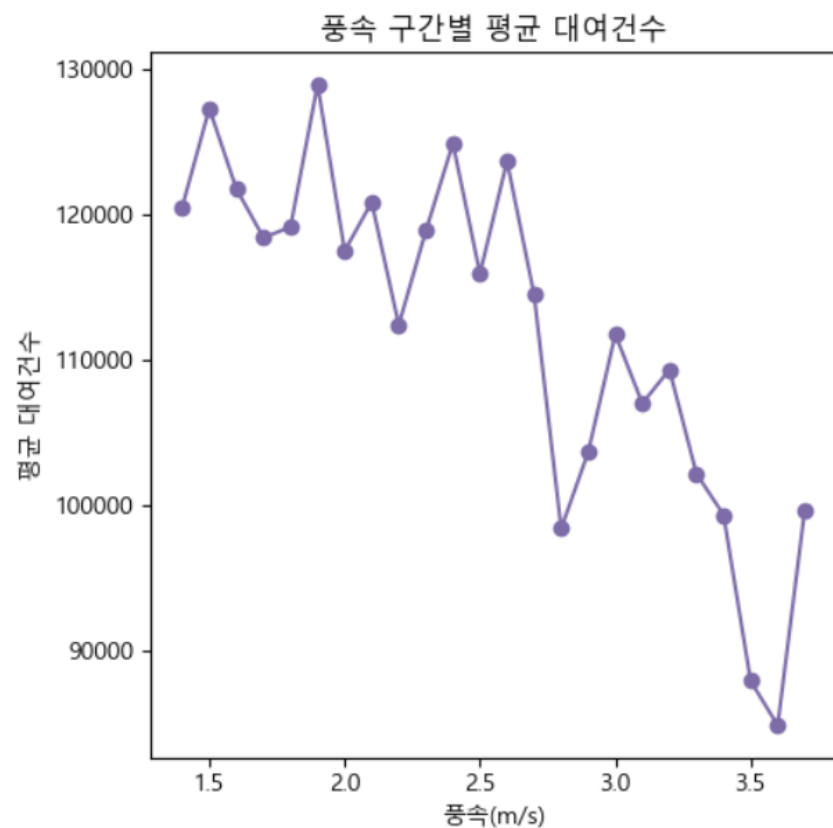
기온 구간별 평균 대여건수 분포



약 22°C 까지는 기온이 증가함에 따라
평균 대여건수도 증가함

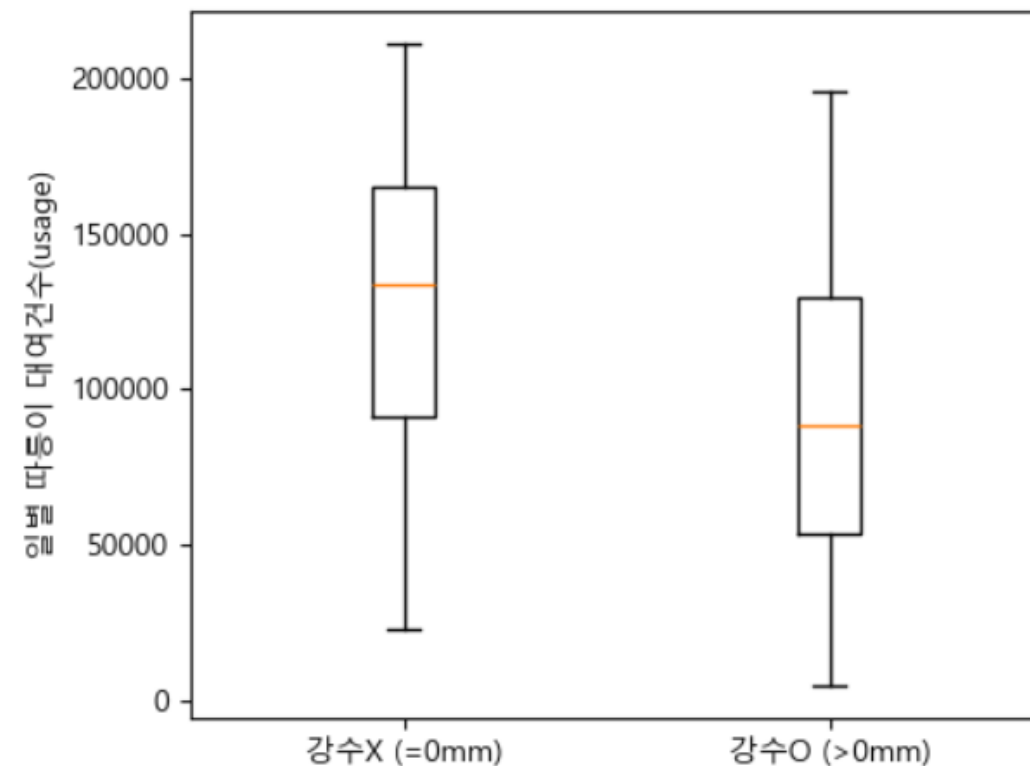
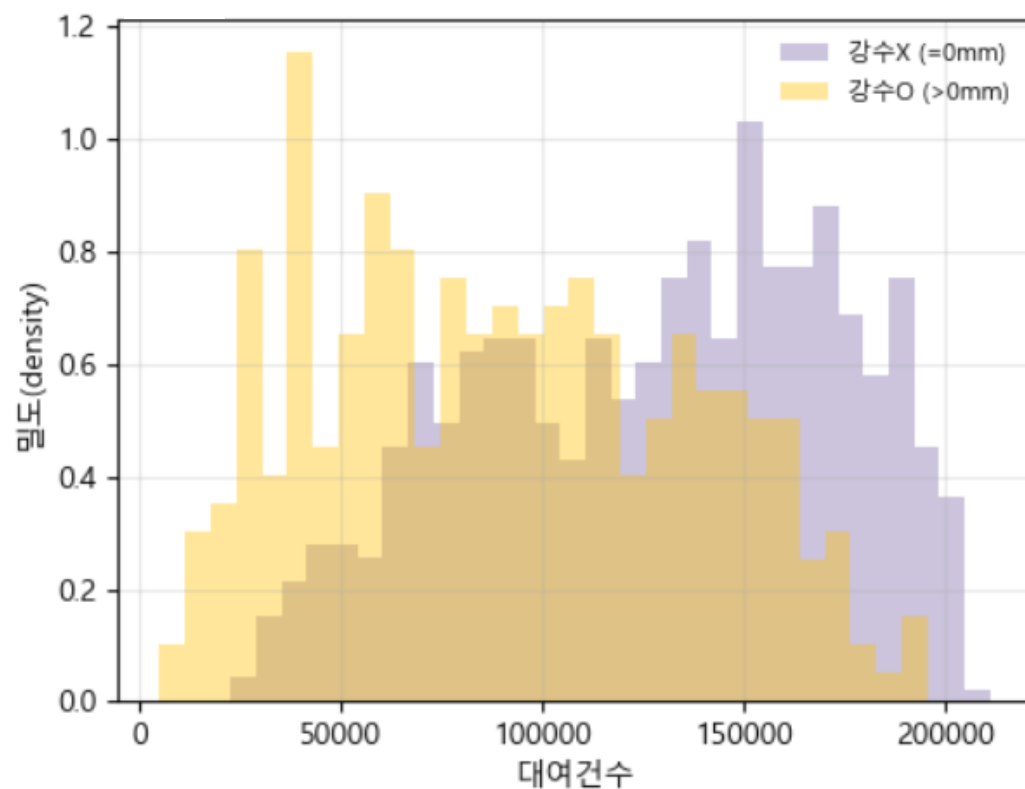
반면 기온이 더 높아지는 경우,
평균 대여건수가 감소하는 경향을 보임

풍속과 강수량은 평균 대여건수와 음의 상관관계



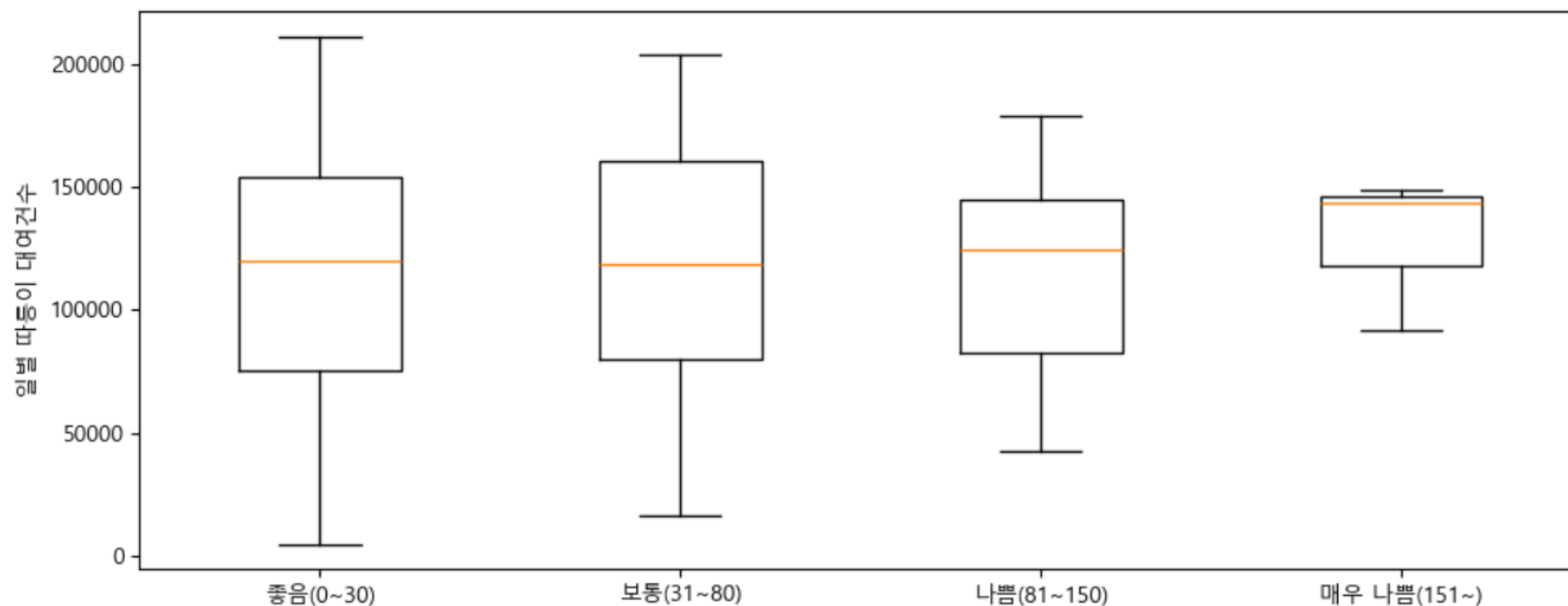
강수 여부에 따라 평균 대여건수는 **약 30%** 차이가 남

강수 여부에 따른 따릉이 대여건수 분포



미세먼지(PM10) 농도와 대여건수 사이에 유의미한 상관관계는 확인되지 않음

미세먼지(PM10) 농도에 따른 따릉이 대여건수 분포



따릉이 이용권은 4가지로 구분되지만 정기권 중심 고정 수요가 압도적임

따릉이 이용권 비중 3개년 평균

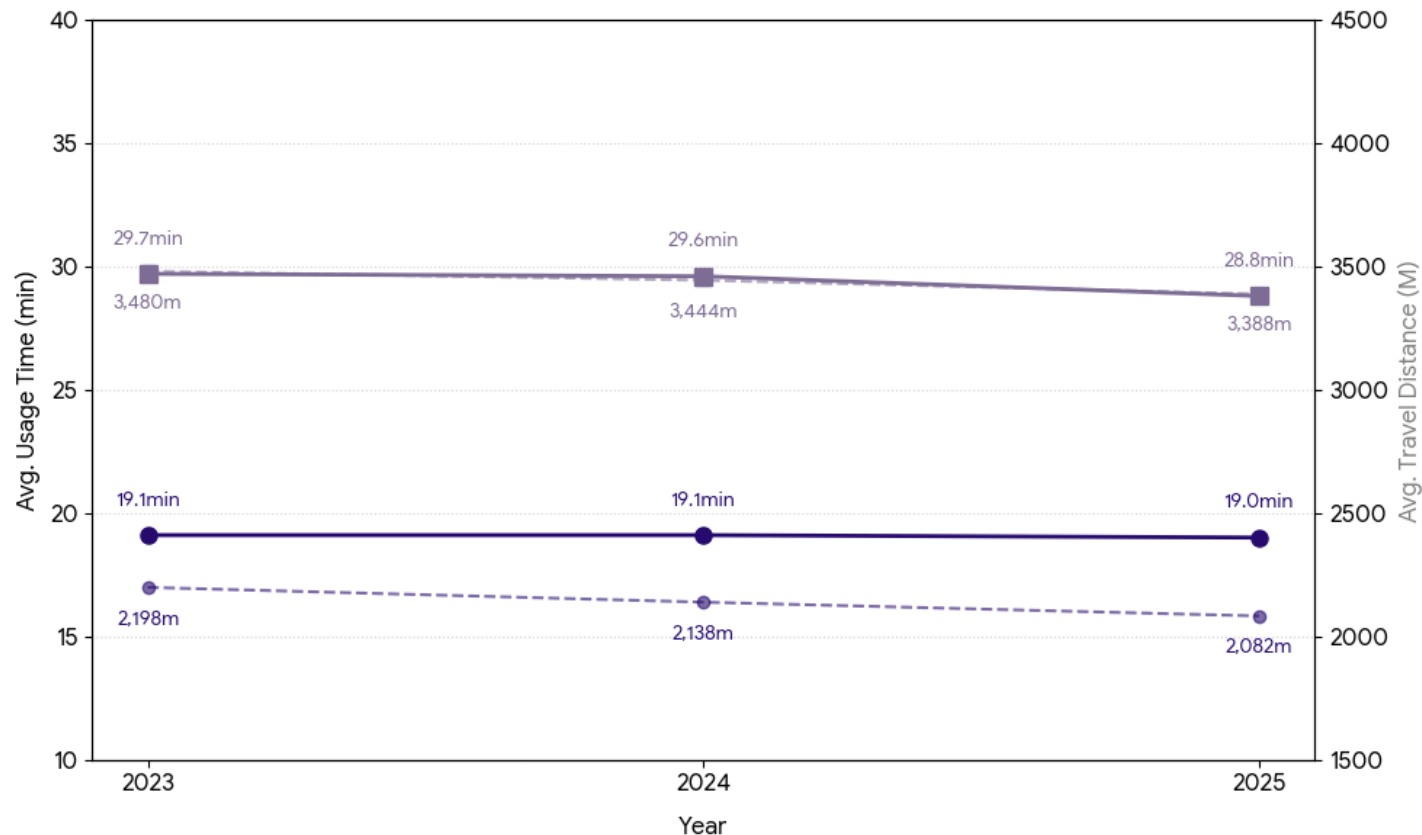
구분	이용 건수 (건)	비중 (%)
정기권	104,927,961	83.23%
일일권	20,054,924	15.91%
일일권(비회원)	1,028,791	0.82%
단체권	25,215	0.02%
가족권	24,768	0.02%
가족권(2시간)	9,224	0.01%
일일권(3시간)	4,696	0.00%
일일권(비회원 3시간)	667	0.00%
가족권(3시간)	218	0.00%
전체	126,076,464	100.00%

이용권은 연간 $\pm 0.3\%p$ 미만의 극소 변동만 존재

➔ 정기권과 일일권을 기준으로 한 분석 방향성 도출

정기권 이용자는 일일권에 비해 1.5배 단시간 & 단거리 이용 특징을 보임

이용권별 이용 시간 및 거리



일일권

평균 이용시간: 약 29.4분

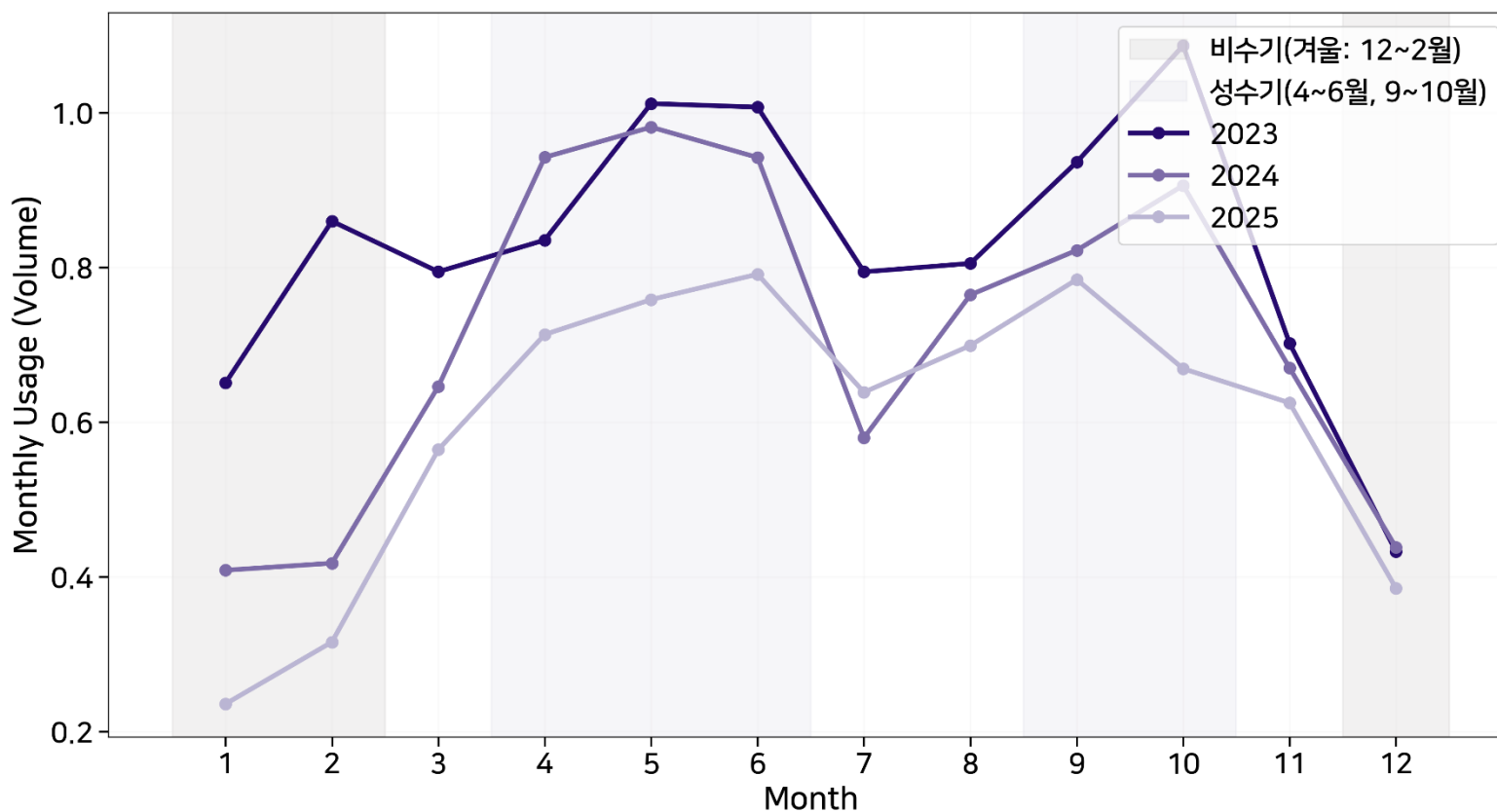
평균 이동거리: 약 3,437m

정기권

평균 이용시간: 약 19.1분

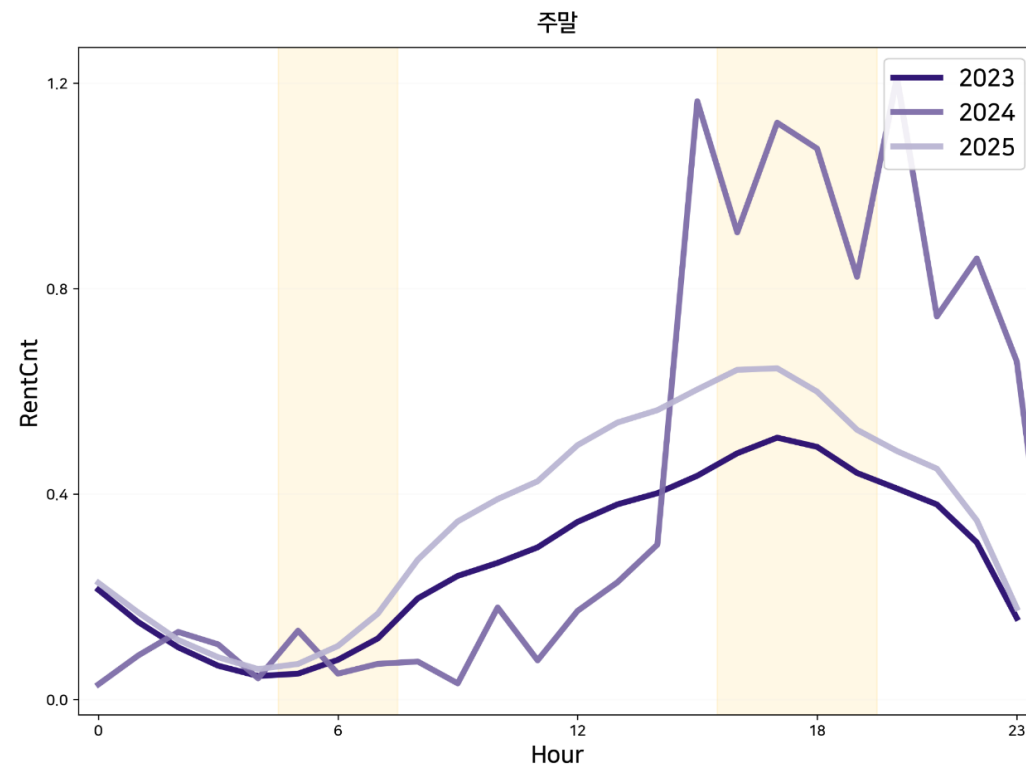
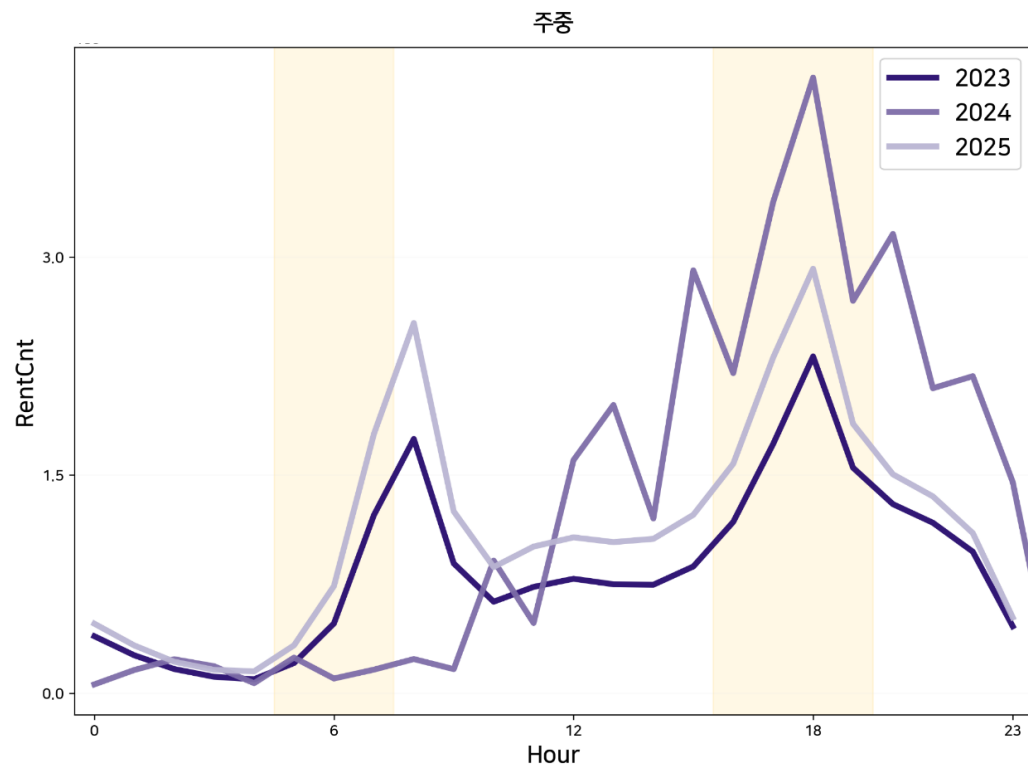
평균 이동거리: 약 2,139m

봄·가을(성수기)과 겨울(비수기)에 월별 이용량 계절성이 두드러짐



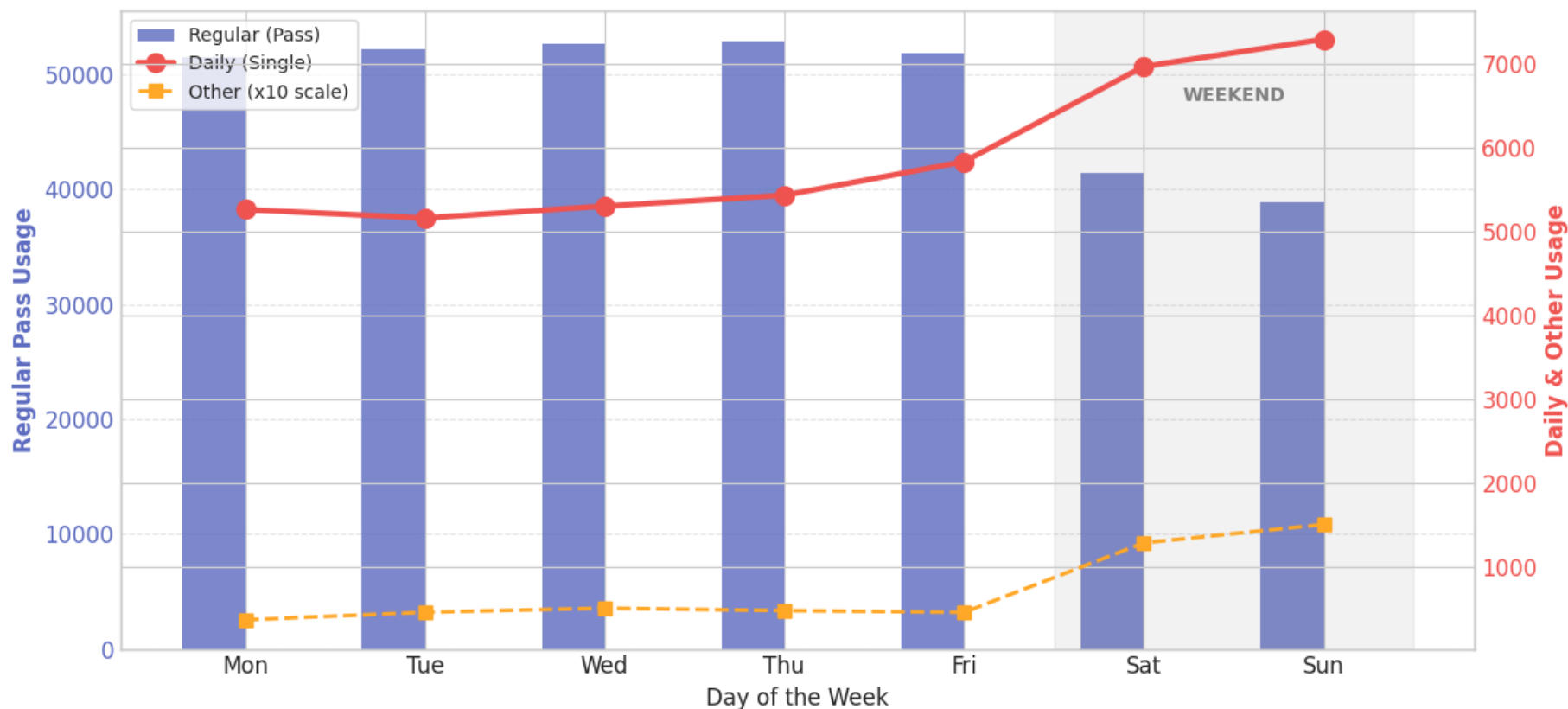
주중에는 주말 대비 출퇴근 구간에서 대여가 집중됨

시간대별 따릉이 대여건수 분포



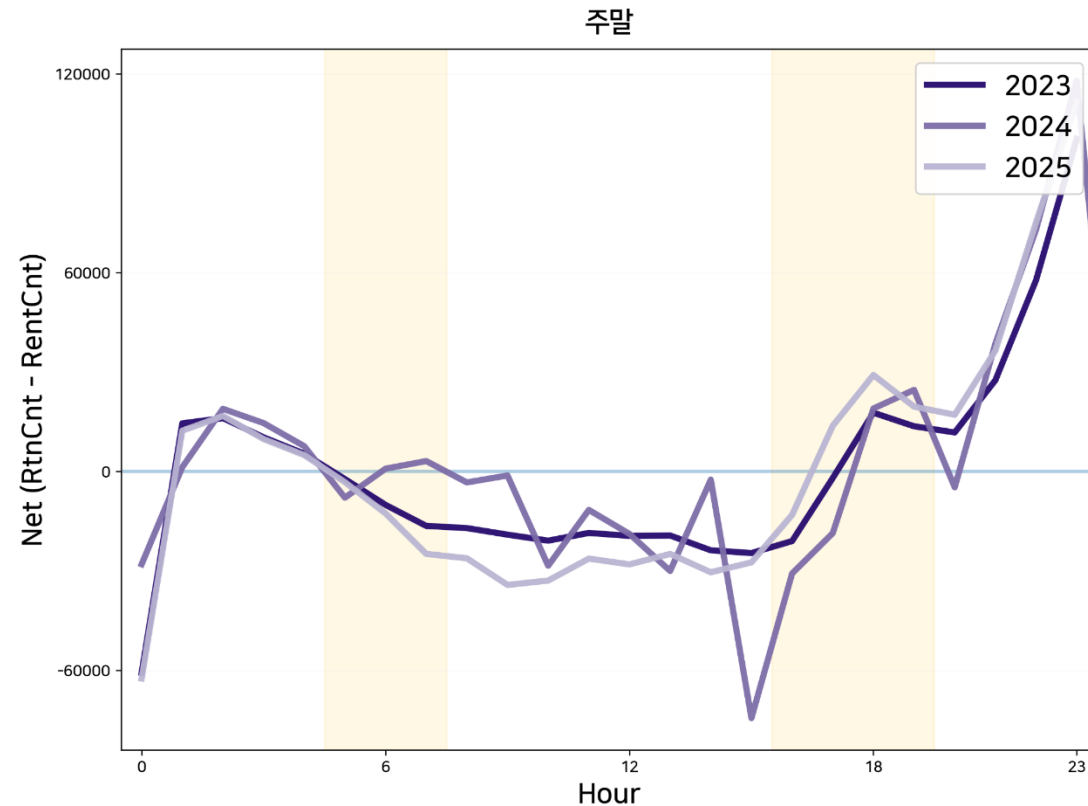
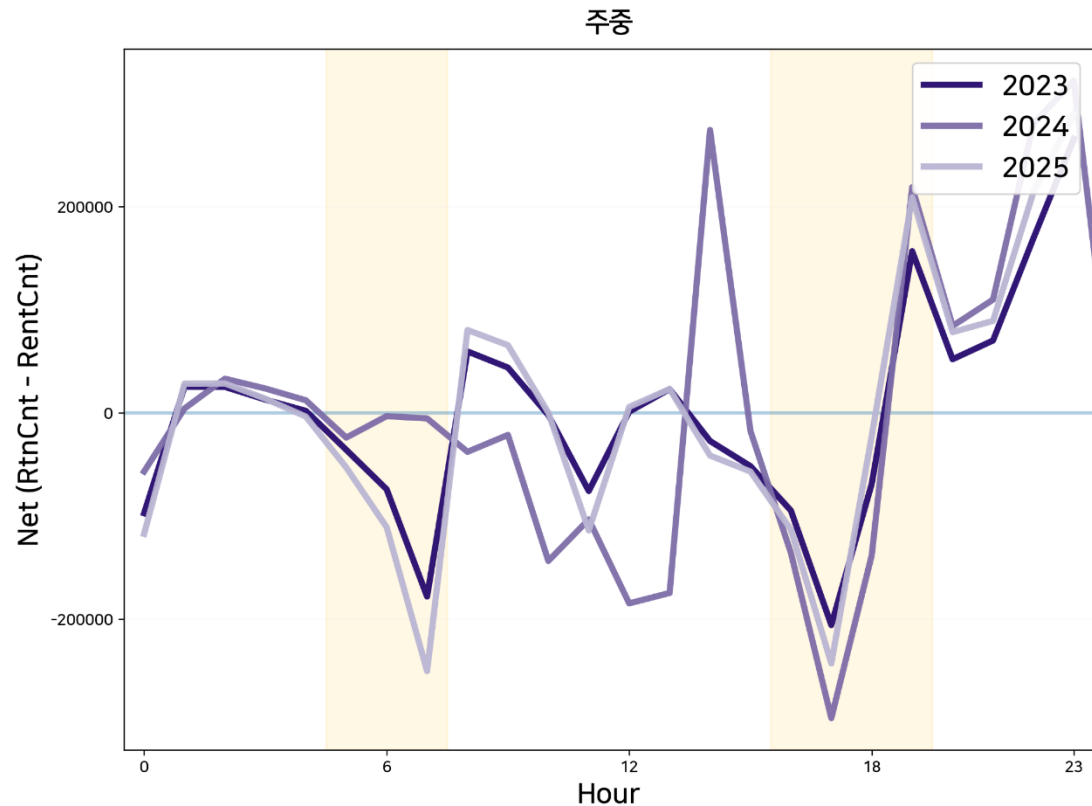
수요 패턴의 변화는 주말에 발생하며 정기권 대비 타 이용권 변동계수는 **약 46% 높음**

요일별 이용 분포



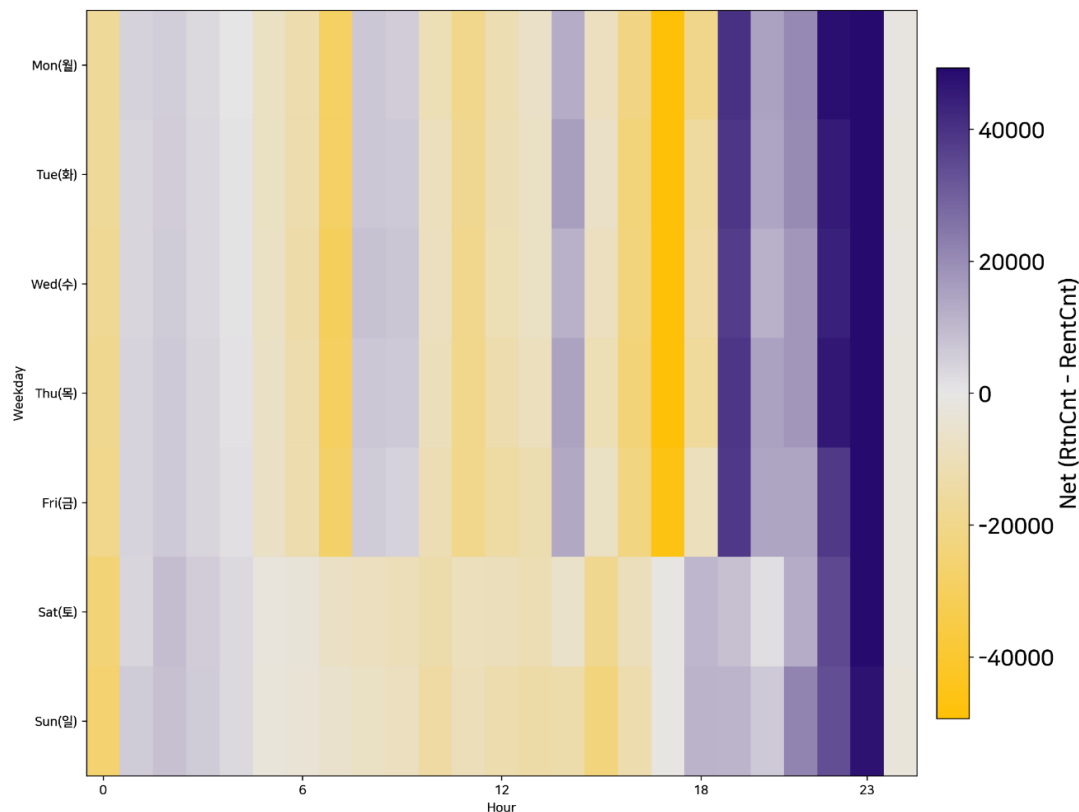
주중 **출퇴근 시간대**에 불균형이 심화되는 경향을 보임

시간대별 따릉이 Net(반납-대여) 분포



요일x시간대 조합에서 퇴근 이후 불균형이 더 강하게 나타남

요일x시간대 Net 평균

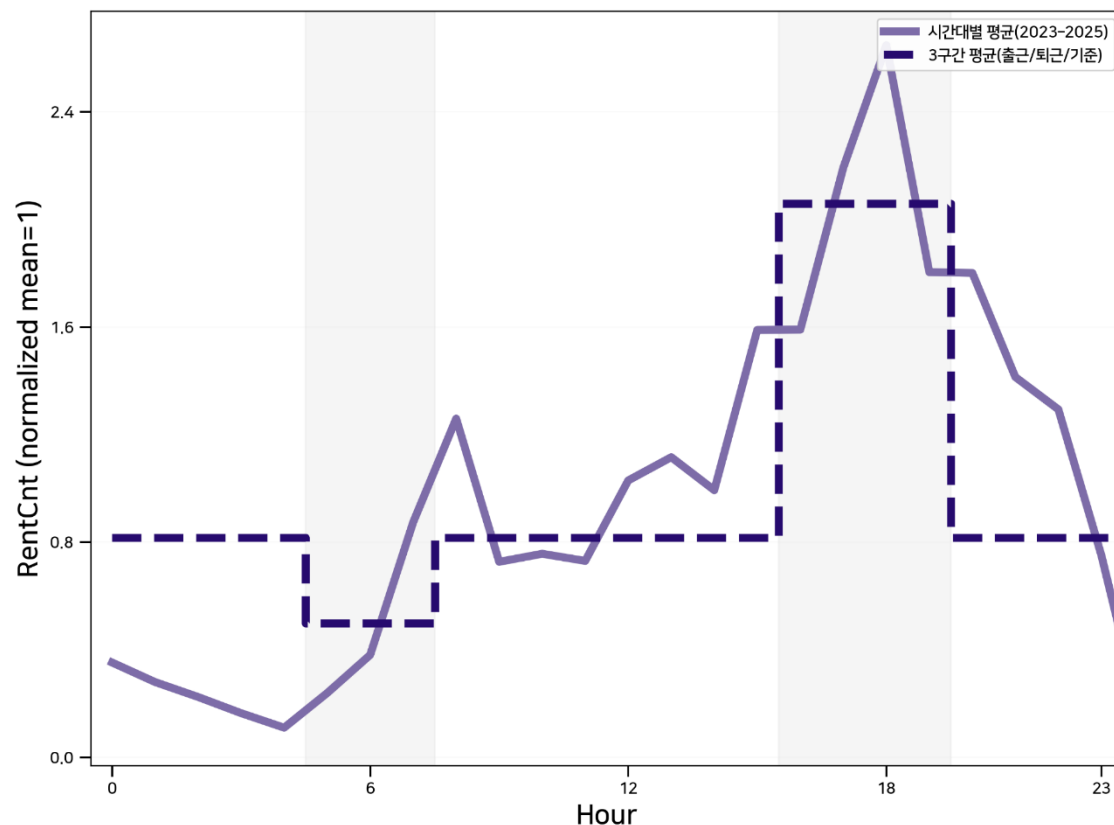


특정 요일, 퇴근 이후(저녁~야간)에 불균형 집중
평균이 아니라 요일x시간 조합에서 구조가 분리됨을 확인

➔ 시간 변수를 연속형으로 두기보다 구조적으로 처리할 필요

출근·퇴근·기준시간 3구간으로 시간대 더미 구조화

시간대 민감도 분석

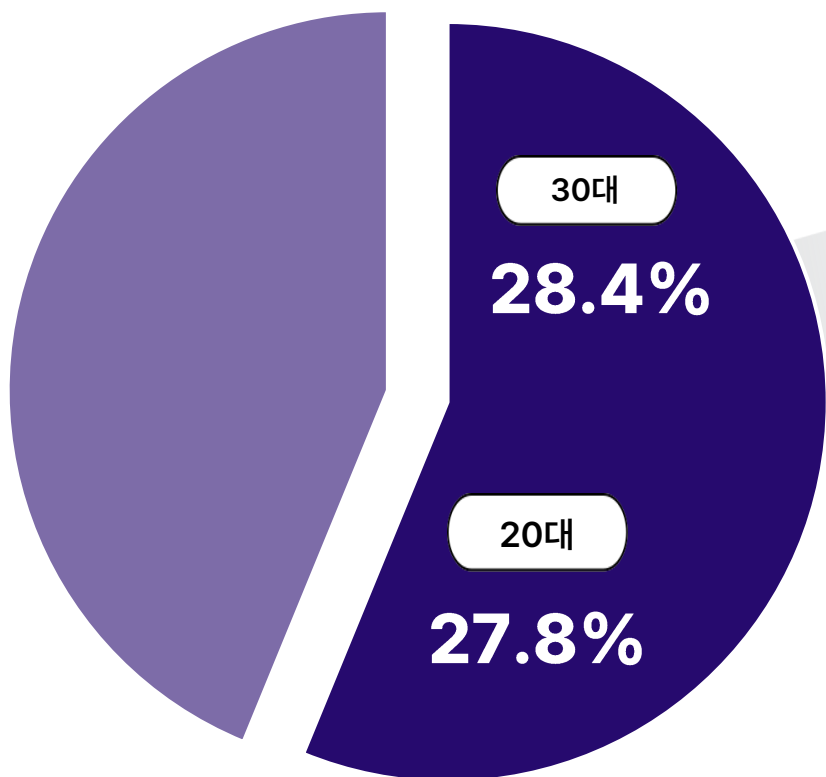


기간에 따라 경계가 조금씩 달라지기는 하나
민감도 분석을 통한 근사 경계로 시간대 분할

연속형 혹은 24개 시간 더미 대신 3구간 더미
출근(5-7시), 퇴근(16-19시), 기준시간

20대와 30대 인구 이용 비율이 전체 이용자의 절반 이상

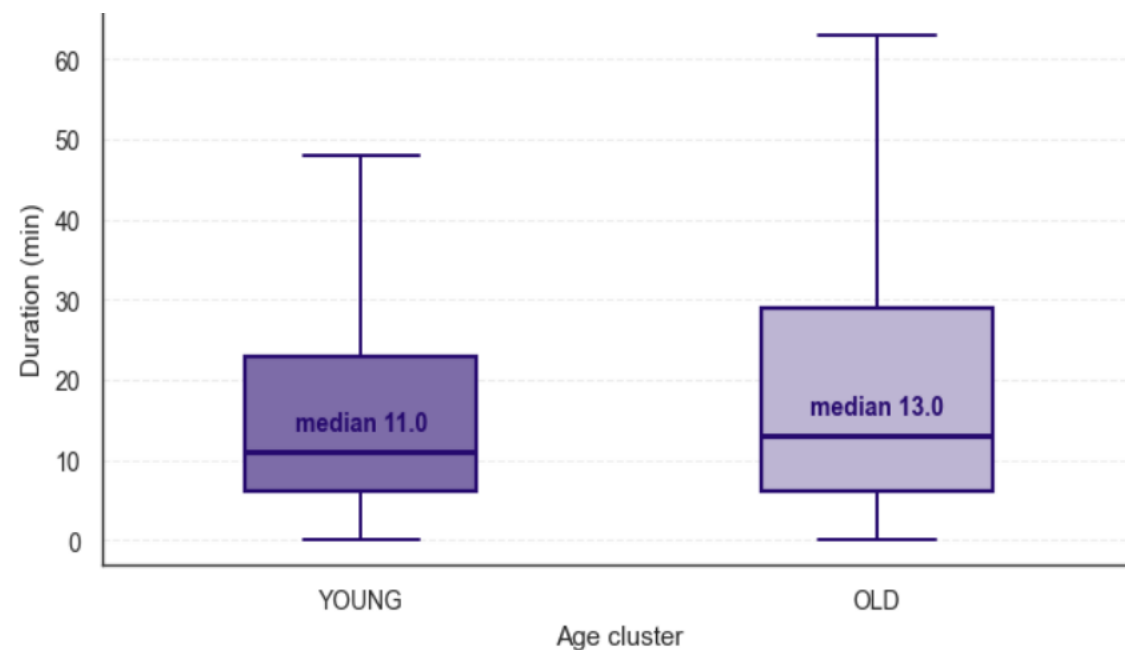
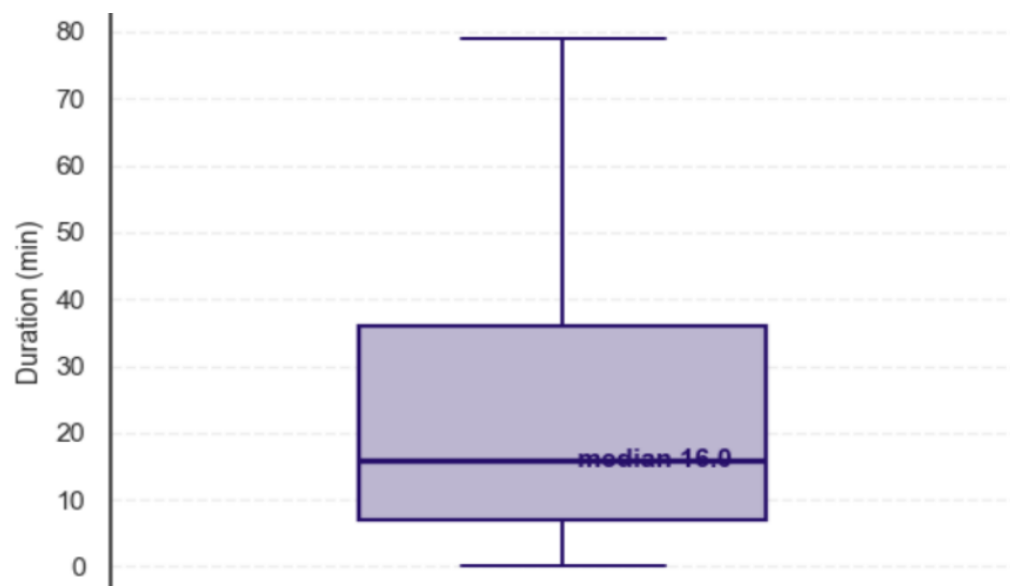
따릉이 이용자 연령대 비율



연령대	비율 (%)
10대	6.4%
20대	27.8%
30대	28.4%
40대	18.3%
50대	13.2%
60대	4.9%
70대 이상	0.9%

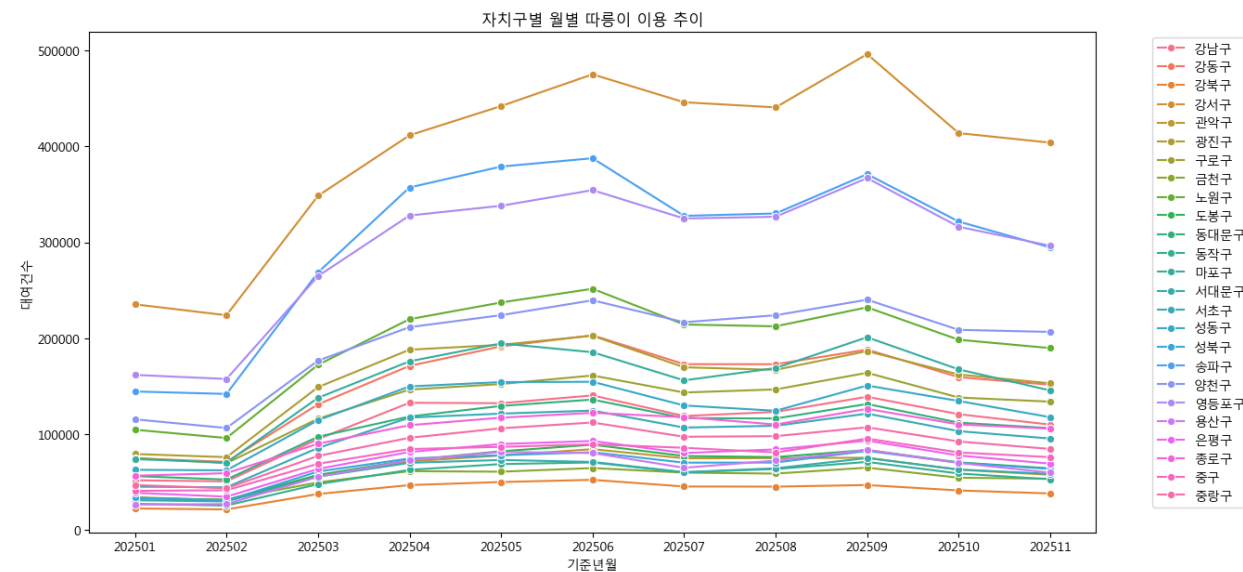
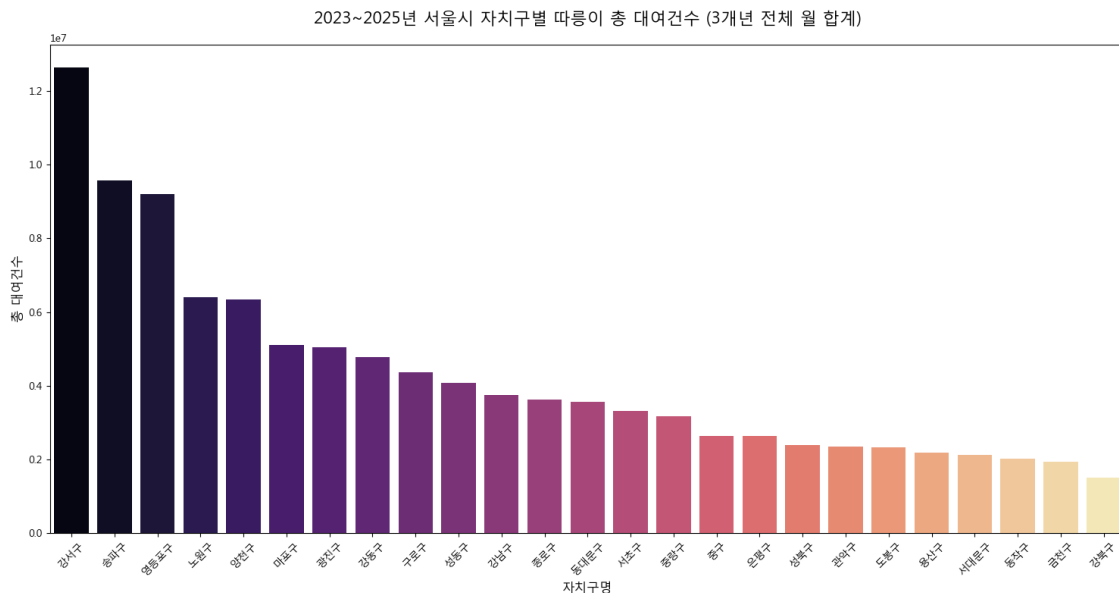
주로 단거리 이동수단으로 사용하며 저연령층의 탑승 분포가 고른 편임

따릉이 이용 시간 분포 및 클러스터링



자치구별 대여건수 편차를 확인하여 대여 영향 요인에 대한 자치구별 분석 필요성을 파악함

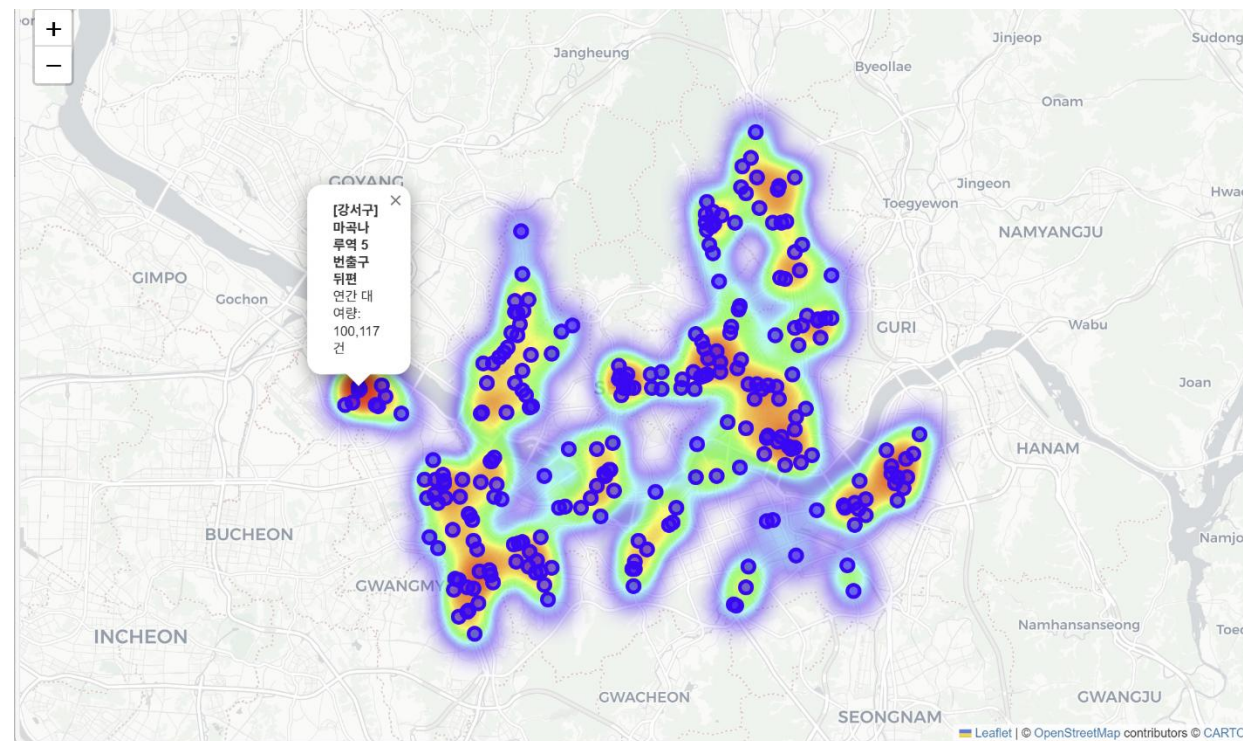
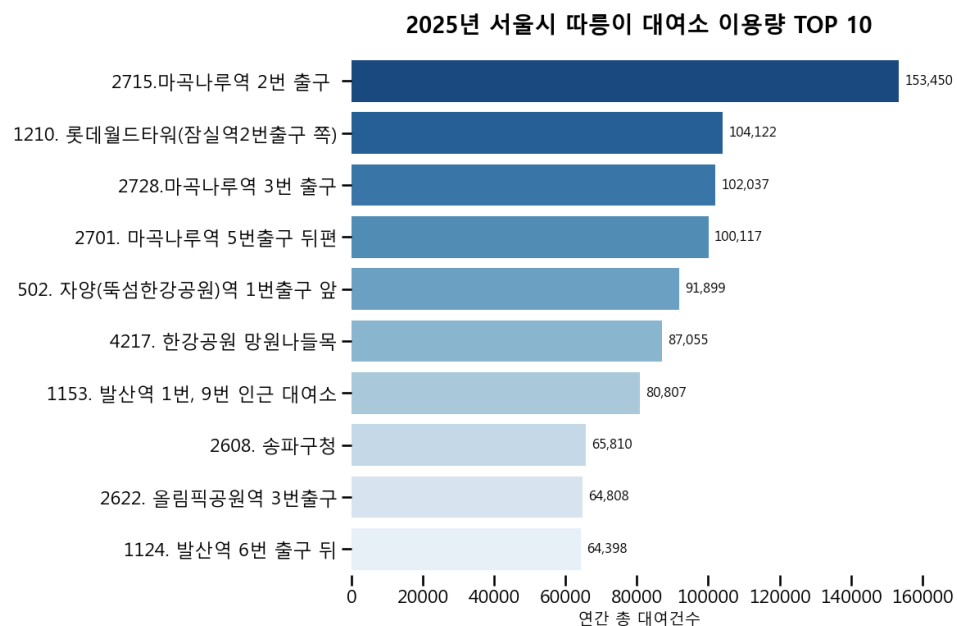
자치구별 총 대여건수



- 강서, 송파, 영등포구가 연간 이용량 최상위, 따릉이 이용이 가장 활발한 지역으로 나타남
- 전체 자치구에서 5~6월과 9월에 이용량이 정점을 찍는 비슷한 경향을 보이며 동절기(1~2월) 이용량은 최하

자치구별 이용량 상위 대여소는 특정 **지하철역 인근**에 몰려 있는 경향을 보임

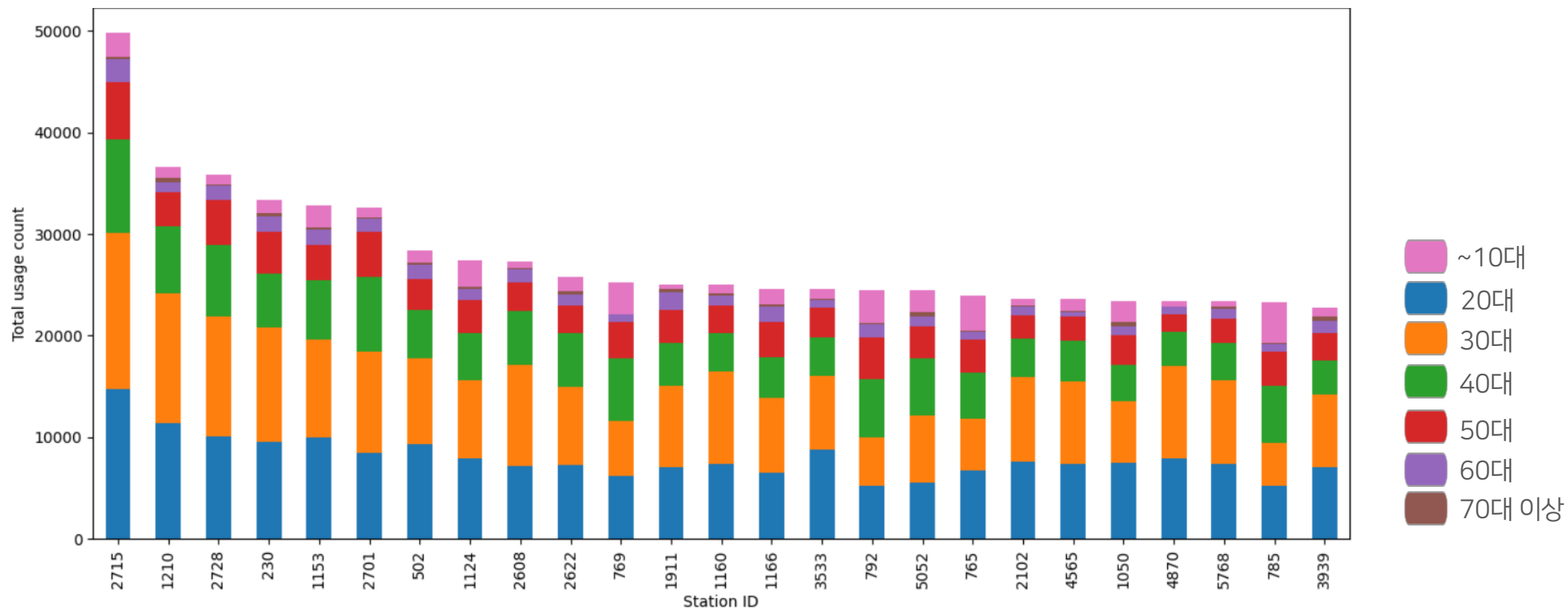
이용량 최상위 대여소 시각화



- 강서구 수요량 가장 많은 이유는 **마곡나루역** 이용 밀집
- 상위 대여소는 구 전체에 고르게 퍼져 있기 보다는 특정 지하철역 인근에 모여 있음
(예: 강서구 마곡나루역, 송파구 잠실역, 종로구 광화문역, 광진구 건대입구역)

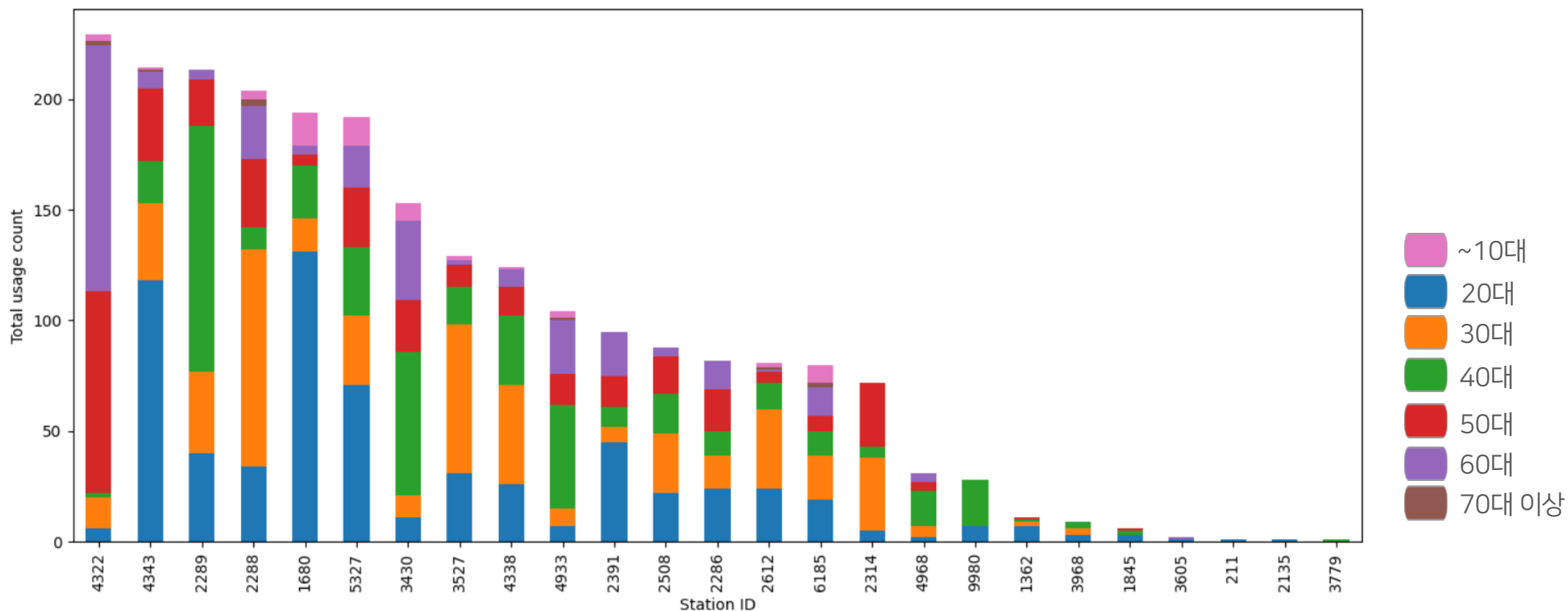
이용률 상위 25개 정류장의 경우 20대부터 40대까지의 이용률이 압도적

연령별 따릉이 이용 시간



이용률 하위 25개 정류장의 경우 비교적 다양한 연령대가 사용

연령별 따릉이 이용 시간



행정구역 및 연령대 분석 인사이트

가설 도출

상위 25개

주로 20, 30, 40대가 사용
마곡나루역, 영등포구청역 등 지하철역 근처 정거장 다수

하위 25개

육안상으로 고연령층 비중이 높은것으로 추정
왕십리 자이아파트, 안골마을 입구 등 거주지 근처 정거장

유의성 검정



상위-하위 25개 정거장 연령 비율 검정

Mann-Whitney U test 결과 p-value 0.23

하위 25개 정거장의 고연령 비율이 높다는 가설 기각

연령대별 변수의 유의성은 찾기 어려움

1. 따릉이 이용자는 대부분 출퇴근시 따릉이를 이용할 것이다
2. 하위 25개 정류장은 비교적 고연령층 이용 비중이 높을 것이다

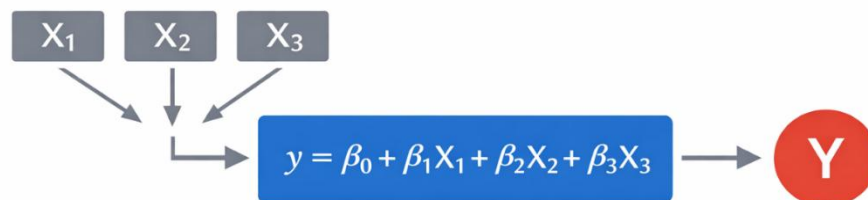
출퇴근과 따릉이 이용률의 상관관계 파악 위해
지하철역과 따릉이 정거장의 거리를 새로운 변수로 제안

이용률 영향 요인

복합적인 변수 중 핵심 영향 요인을 정량적 도출을 위해 다중선형회귀를 채택함

분석 근거 및 목표

- 따름이 이용률은 다양한 외생 변수가 복합적으로 작용하므로, 각 요인이 이용 행태에 미치는 개별적 영향력을 정량적으로 분리하여 파악하고자 함
- 독립 변수와 이용률 간의 상관관계 방향(+, -) 및 크기를 수치화하여 데이터 해석력 확보



- 인간의 행동을 분석하는 사회과학 및 정책 분야에서는 높은 결정계수(R^2)값보다 변수의 통계적 유의성 확보(F-통계량, 개별 변수들의 p-value)와 상관성을 규명하는 것이 목적
- 규명된 상관성을 바탕으로 시공간적 수요 불균형을 야기하는 지표 식별

회귀 변수 및 전처리

변수명	전처리	목적
log_rentals (Y)	로그 변환을 통해 왜도 제거 및 회귀 모델 정규성 확보	데이터의 왜도 제거 및 회귀 모델의 정규성 확보
기온_중심화_제공	평균중심화 및 제곱화	비선형적 관계 모델링 (적정 온도 이후 감소 효과 반영)
강수여부, 풍속	원본 유지	기상 요인이 대여량에 미치는 직접적 영향력 측정
지하철역거리(m)	Standard Scaling	역세권 영향력의 비선형적 감소 특성 반영 및 수치 표준화
총생활인구	Standard Scaling	자치구별 인구 규모 차이에 따른 편향 제거 및 통제 변인 보정
TIME_BLOCK	One-Hot Encoding	시간대별(출퇴근 등) 고정 효과 통제
요일(Day of Week)	One-Hot Encoding	평일/주말의 상반된 패턴 및 요일별 미세 수요 차이 정밀 분석

다중선형회귀 결과

변수명	2023년	2024년	2025년
주말	-0.2040	-0.2053	-0.1834
공휴일	-0.1787	-0.2305	-0.2249
강수 여부	-0.2821	-0.2286	-0.2517
기온	+0.0130	+0.0107	+0.0087
저녁 피크 타임	+0.2022	+0.2136	+0.2033

Ex) 2025년도 이용률

이용률

$$= 1.7932 + (0.2033 \times \text{저녁 피크 타임}) - (0.2517 \times \text{강수 여부}) + (0.0087 \times \text{기온}) \\ - (0.1834 \times \text{주말 여부}) - (0.2249 \times \text{공휴일 여부}) \dots$$

- 주말 및 공휴일은 모두 (-) 관계를 보임. 이는 따릉이가 주말 여가용보다는 평일 출퇴근 등의 목적의 이동 수단으로 더 활발히 이용되고 있음을 뜻함
- 강수 여부: 이용률을 가장 크게 감소시키는 요인이며, 비가 오면 따릉이 이용률이 급격히 줄어듦
- 기온: 수치는 작지만 (+) 관계로, 날씨가 따뜻해질수록 이용자가 늘어나는 경향 확인
- **저녁 피크 시간대: 이용률에 가장 긍정적인 영향을 주는 요인.** 저녁 시간대에 이용률이 눈에 띄게 증가하여 퇴근 및 여가 수요가 핵심임을 보여줌

자치구별 이용 행태에 영향을 미치는 개별 요인에 따라 **상이한 운영 전략을 고안**해야 함

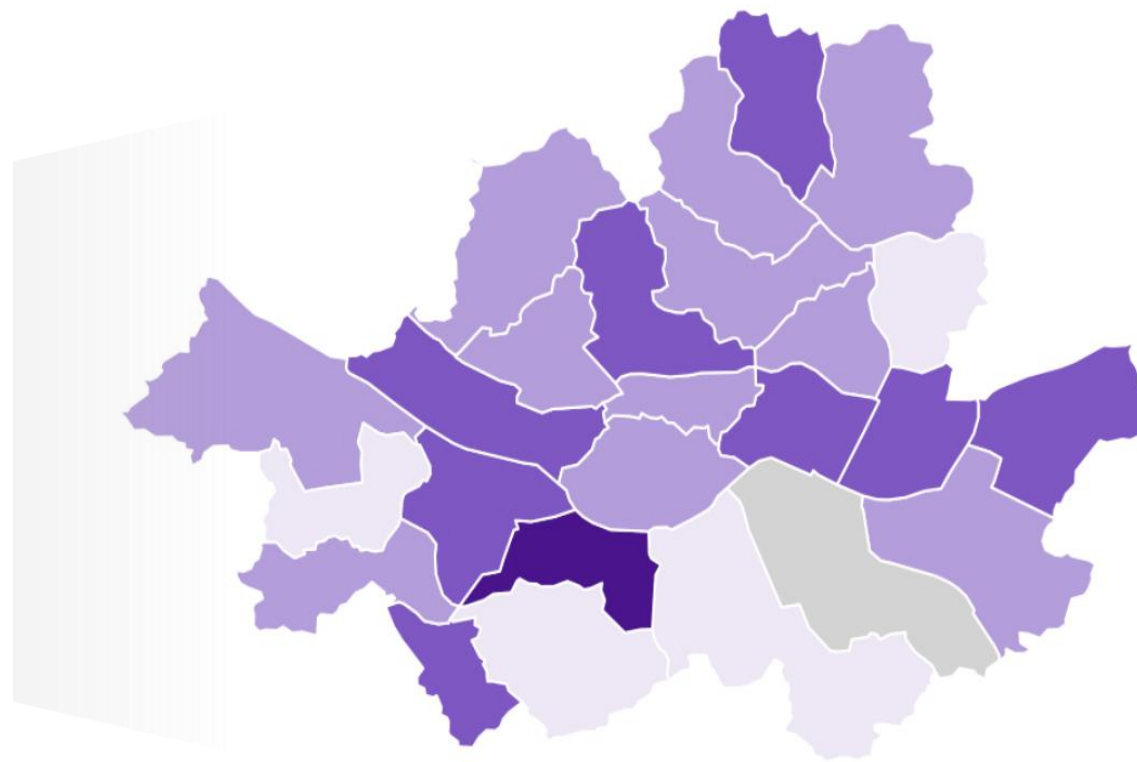
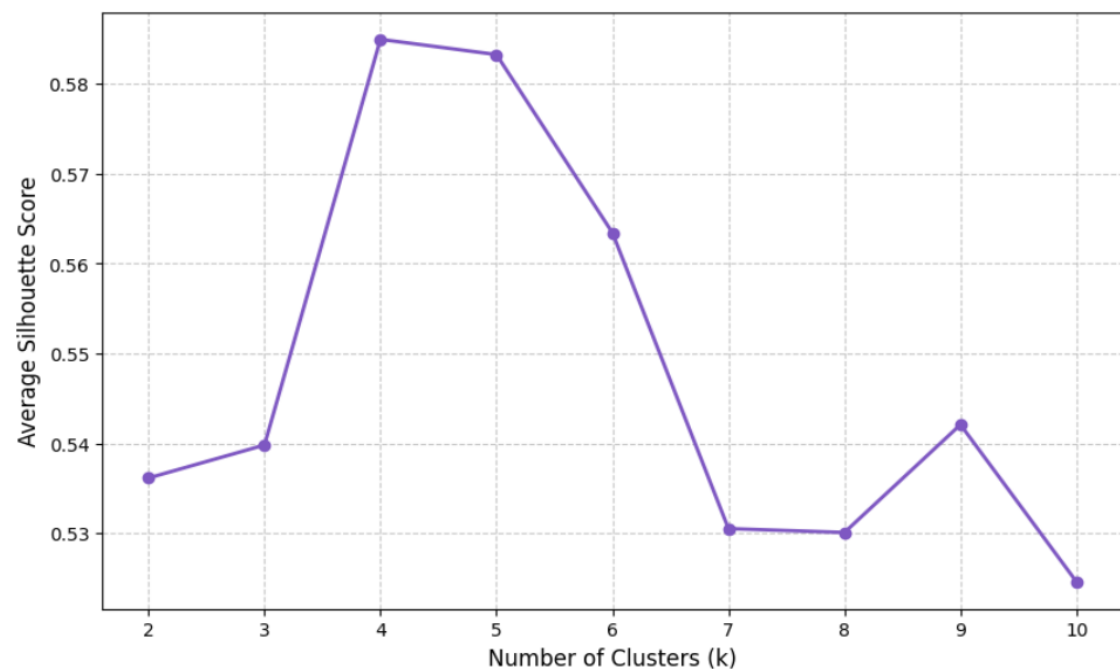
자치구별 공간, 환경적 요인 회귀 분석

변수명	2023년	2024년	2025년
주말	-0.2040	-0.2053	-0.1834
공휴일	-0.1787	-0.2305	-0.2249
강수 여부	-0.2821	-0.2286	-0.2517
기온	+0.0130	+0.0107	+0.0087
저녁 피크 타임	+0.2022	+0.2136	+0.2033
동작구	+0.3038	+0.3449	+0.3074
마포구	+0.2035	+0.1766	+0.1596
서초구	-0.0999	-0.1124	-0.1394
관악구	-0.1909	-0.1986	-0.2036

- 이용량 중앙값인 강남구 대비 상대적 증감을 회귀모델로 표현
- 2023~2025년 3개 년도 모두 동작구, 마포구, 종로구 등은 (+)의 높은 상관계수를 보임
- 해당 지역들은 지리적 여건이나 인프라 덕분에 다른 변수(기온, 시간대 등)가 동일할 때도 따릉이가 훨씬 활발하게 이용되는 이용 효율이 높은 지역

계층적 군집 분석을 위한 실루엣 점수 (K=4) 0.58로 나타남

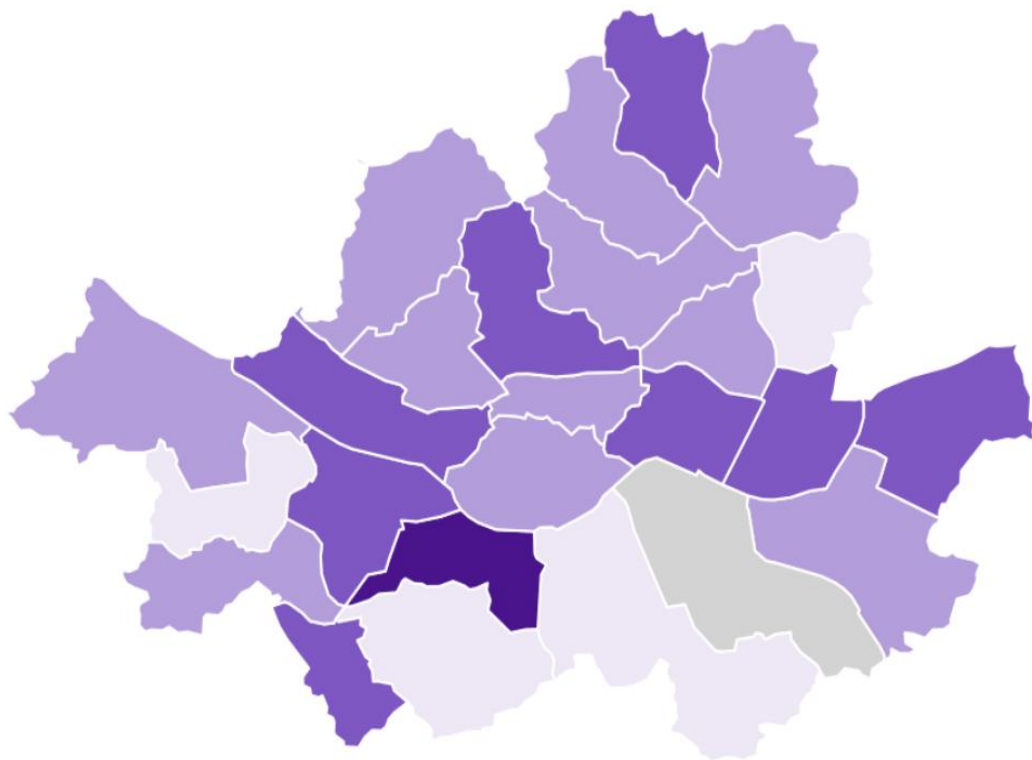
실루엣 점수 그래프



영향력 수치 기준 계층적 군집 분석 결과

자치구별 4가지 군집을 도출함

자치구별 계층적 군집 분석



군집	성격	기본 수요	저녁 피크 반응성	기상 민감도	해당 자치구
Cluster 1	효율형	최고치	높음	보통	동작구 (단독)
Cluster 2	성장형	양수(+)	높음	보통	마포, 성동, 영등포, 종로 등
Cluster 3	정체형	음수(-)	보통	보통	강서, 송파, 노원, 중구 등
Cluster 4	취약형	최저치	보통	보통	관악, 서초, 양천, 중랑

*타 지역 대비 반응성을 기준으로 높음, 보통으로 표기
강남구를 기준으로 하여 강남구는 군집에서 제외

영향력 수치 기준 계층적 군집 분석 결과 자치구별 4가지 군집이 도출됨

자치구별 계층적 군집 분석



군집	성격	기본 수요	저녁 피크 반응성	기상 민감도	해당 자치구
Cluster 1	효율형	최고치	높음	보통	동작구 (단독)
Cluster 2	성장형	양수(+)	높음	보통	마포, 성동, 영등포, 종로 등
Cluster 3	정체형	음수(-)	보통	보통	강서, 송파, 관악, 서초 등
Cluster 4	취약형	최저치	보통	보통	관악, 서초, 양천, 중랑

*타 지역 대비 반응성을 기준으로 높음, 보통으로 표기
강남구를 기준으로 하여 강남구는 군집에서 제외

영향력 수치 기준 계층적 군집 분석 결과 자치구별 4가지 군집을 도출됨

자치구별 계층적 군집 분석



군집	성격	기본 수요	저녁 피크 반응성	기상 민감도	해당 자치구
Cluster 1	효율형	최고치	높음	보통	동작구 (단독)
Cluster 2	성장형	양수(+)	높음	보통	마포, 성동, 영등포, 종로 등
Cluster 3	정체형	음수(-)	보통	보통	강서, 송파, 노원, 중구 등
Cluster 4	유악형	최저치	보통	보통	관악, 서초, 양천, 중랑

- 인구가 밀집되어 있고 자전거 이용이 생활화된 지역으로 저녁 퇴근 시간대 반응성이 매우 높음

- 운영 전략: 퇴근 시간 및 여가 시간 수요에 집중한 배치 필요

*타 자치구 대비 반응성은 낮은 편임(보통)으로, 강남구를 기준으로 하여 강남구는 군집에서 제외

영향력 수치 기준 계층적 군집 분석 결과 자치구별 4가지 군집을 도출됨

자치구별 계층적 군집 분석



- 대여 건수는 많을 수 있으나, 회귀 분석상 기본 수요는 음수를 기록
- 강서, 송파처럼 이미 자전거가 과잉 공급되어 '대당 효율'이 떨어지고 있거나, 단거리 교통수단으로서의 매력력이 정체된 상태로 추측
- 운영 전략: 이용률이 현저히 낮은 대여소에 대한 점검 선행

Cluster 3

정체형

음수(-)

보통

보통

강서, 송파,
노원, 중구 등

Cluster 4

취약형

최저치

보통

보통

관악, 서초,
양천, 중랑

*타 지역 대비 반응성을 기준으로 높음, 보통으로 표기
강남구를 기준으로 하여 강남구는 군집에서 제외

영향력 수치 기준 계층적 군집 분석 결과 자치구별 4가지 군집을 도출됨

자치구별 계층적 군집 분석



군집	성격	기본 수요	저녁 피크 반응성	기상 민감도	해당 자치구
Cluster 1	최저치	가장 낮음	가장 낮음	가장 낮음	관악, 서초, 양천, 중랑 등
Cluster 2	최저치	가장 낮음	가장 낮음	가장 낮음	관악, 서초, 양천, 중랑 등
Cluster 3	최저치	가장 낮음	가장 낮음	가장 낮음	관악, 서초, 양천, 중랑 등
Cluster 4	최저치	가장 낮음	가장 낮음	가장 낮음	관악, 서초, 양천, 중랑 등

- 기본 수요가 가장 낮으며(-0.14), 외부 환경 요인에 대한 저항력이 강함
- 관악, 서초처럼 경사가 급하거나 광복 도로가 많아 물리적으로 자전거를 타기 힘든 환경에 기인할 것으로 추론
- 운영 전략: 외부 요인에 대한 추가 분석을 통한 배치 전략 고안

*타 지역 대비 반응성을 기준으로 높음, 보통으로 표기
강남구를 기준으로 하여 강남구는 군집에서 제외

결론 및 제언

날씨

기상 조건에 민감하여 강수 시 이용량이 약 30% 급감하며, 기온은 약 22°C까지 이용량이 비례하여 증가하다가 그 이상에서는 감소함

이용권, 시간대

정기권 중심의 고정 수요가 압도적인 비중(82%)을 차지하며 일주일 중 평일 퇴근 시간대의 수요가 가장 높음

연령, 행정구역

주로 20~30대 인구가 특히 마곡나루역, 잠실역 등 주요 지하철 거점 대여소에 이용 밀집 현상이 발생함 또한 자치구별 이용 편차가 두드러짐

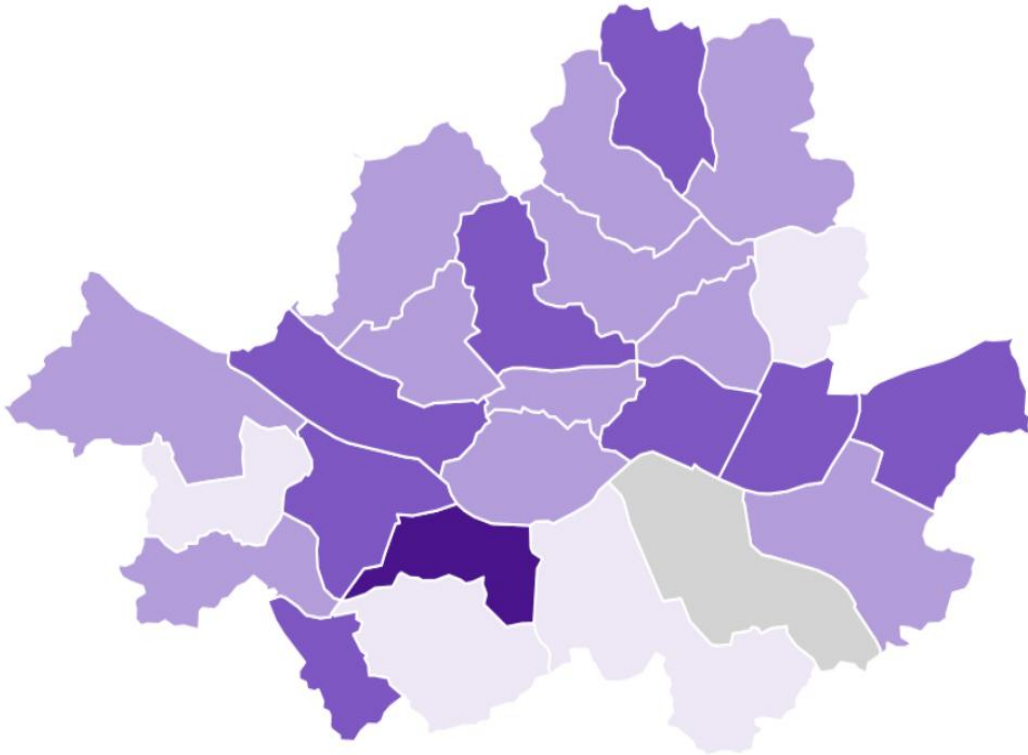
다중선행회귀 분석 결과

따릉이 배치 문제는 단순 운영 문제가 아닌 구조적인 **시간·공간적 불균형 문제**임을 확인

자치구별 맞춤형 전략 도입을 통해 효율적인 운영과 실질적인 적자 관리를 제언함

분석 의의 및 한계점

군집별 상이한 운영 전략 고안



한계점

- 공간 단위 거시적 분석 방향성
같은 구 안에서도 대여소별 고도나 주변 도로 환경이 다르지만, 이를 '구'라는 하나의 수치로 평균화했다는 한계가 존재
- 수집 변수의 한계
대여량에 영향을 줄 수 있는 지형적 경사도, 자전거 도로의 연속성, 인근 지하철역의 승하차 인원수 등을 직접 변수로 포함하지 못하고 자치구 더미 변수를 통해 간접적으로만 추론함

확장 가능성

외부 데이터와 비선형 머신러닝 모델을 결합한다면 더욱 정밀한 실시간 재배치 알고리즘으로 발전시킬 수 있을 것임

E.O.D