
지하철 이용 패턴 기반 가성비 맞춤 타겟팅 광고 전략

4팀
이성우
이아름
윤여진
최혜미
이예림



목차

1 분석 배경

기존 광고의 한계
지하철 타겟팅 전략 제안

2 데이터 수집

메인 데이터
기타 데이터

3 문제 정의

데이터의 이질성
타겟의 차별성

4 데이터 전처리

결측치 및 중복값 제거
시간대별 집계
특성별 그룹화

5 분석 방법론

주거/비주거 구분
시장규모 분석
연령대 및 폐르소나

6 이용 패턴 분석

일별 이용 패턴
순유입 데이터 산출
활동성 지수 분석

7 결론 및 한계점

산출물 및 인사이트
한계점 및 개선방향
프로젝트 성과

8 프로젝트 회고

주제 선정 및 갈등 해결
협업과 소통의 성과

01 분석 배경



기존 광고 플랫폼의 타겟팅 한계

현재 지하철 광고는 불특정 다수를 대상으로 하여 광고 효율성이 낮음

지하철 이용자 특성 기반 타겟팅 전략

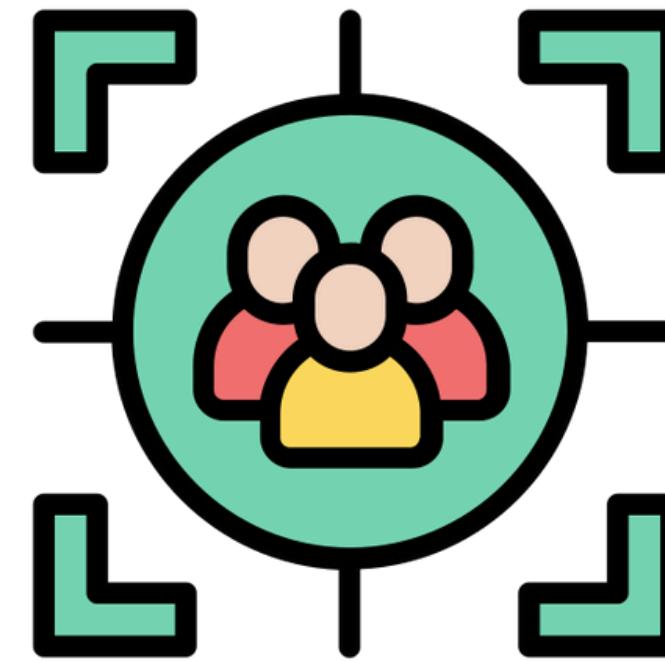
지하철 이용 패턴, 연령대별 특성, 역별 성향 데이터를 분석하여 가성비 높은 맞춤형 광고 전략 제안

01 “광고주는 왜 고민하는가”



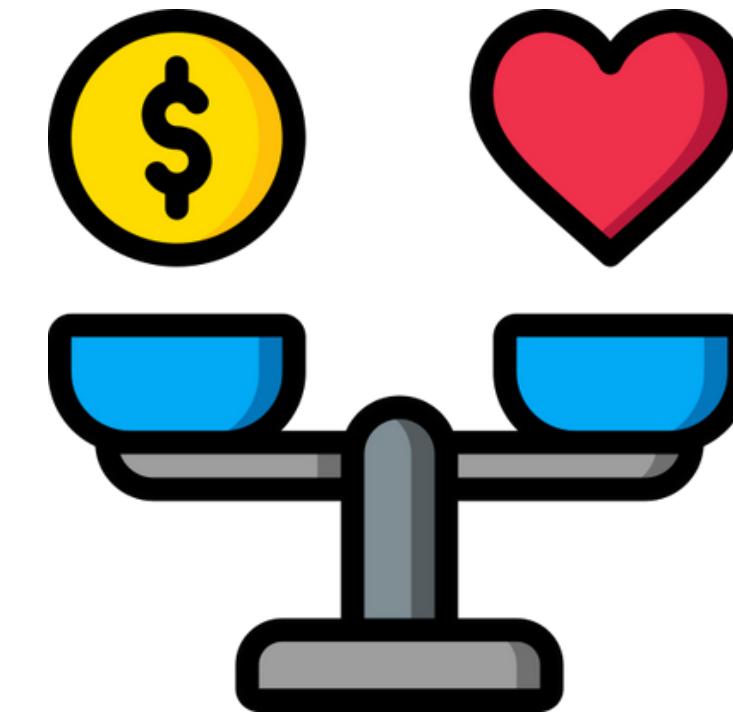
비싼 광고비

유동인구가 많은 핵심 상권이
항상 최적의 선택일까?



명확한 타겟 필요

역별 이용자 특성은
브랜드 타겟과 일치할까?



저예산 고효율

제한된 예산 내에서 효율을
극대화할 수 있는 대안은 무엇일까?

02 데이터 수집

메인 데이터



서울시 지하철 역별 승하차 데이터
25년 상반기 1~8호선 역별

이용자수 통계

(<https://data.seoul.go.kr/dataList/OA-22470/F/1/datasetView.do>)

기타 데이터



위도 및 경도 정보

각 역의 지리적 위치 좌표를 활용

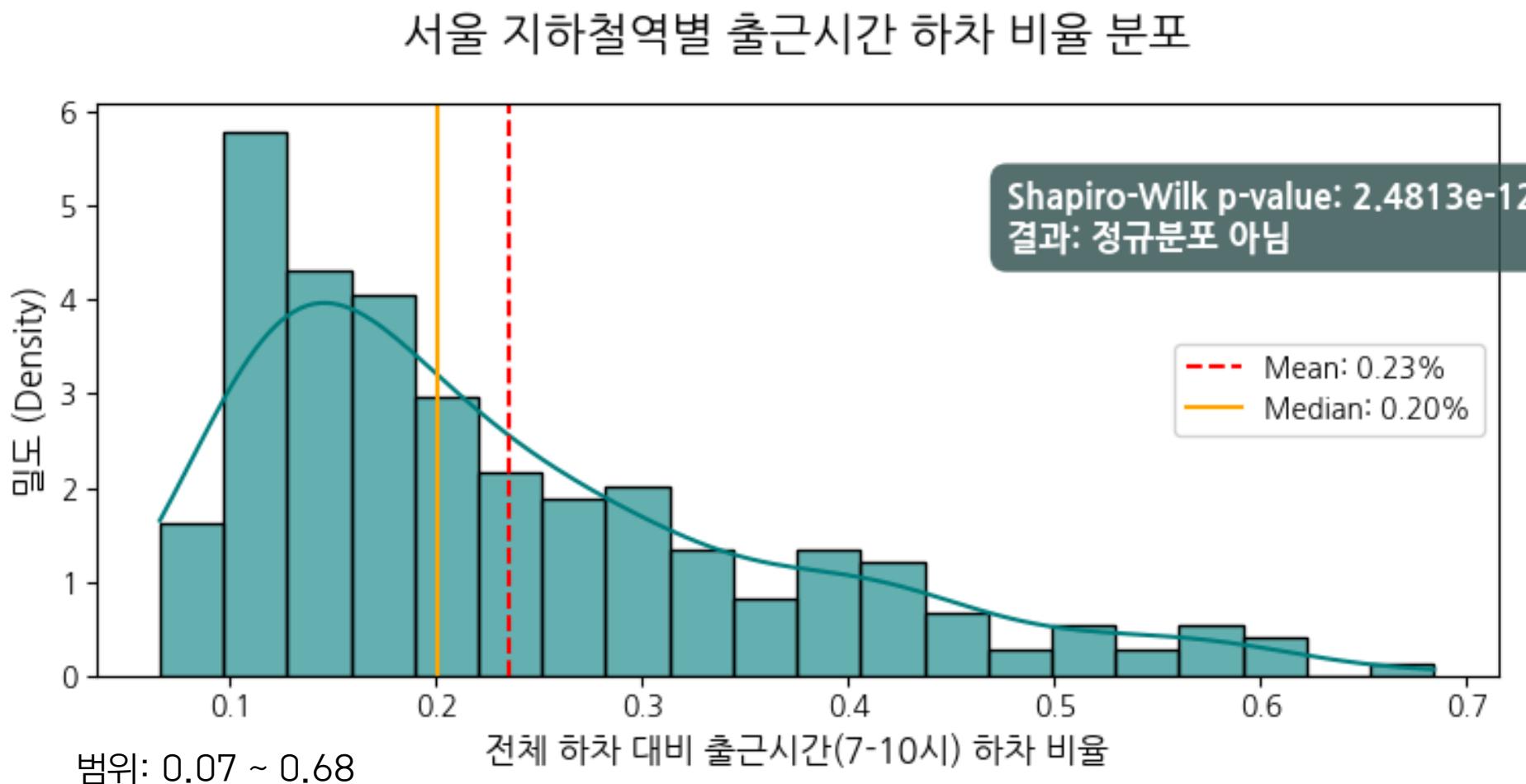


역 등급 정보

광고 대행사에서 책정한 역별 등급(SSA~C)을
통해 광고 효율성과 비용 대비 효과 검증에 활용

03 데이터의 이질성 확인

“지하철역은 모두 똑같지 않다”



분석 개요

지표: 오전 출근 시간대(07-10시) 하차 비중

분석목적:

- 역별 이용 패턴의 이질성 존재 여부를 탐색적으로 확인

통계 분석 결과

- Shapiro-Wilk 정규성 검정 : P-value < 0.05
- *CV(변동계수) : 55%

출근시간 하차 비율은 0.07~0.68의 넓은 범위와 CV 55%로 평균 대비 변동성이 매우 크게 나타나 일부 역은 높은 집중도를, 일부는 낮은 비율을 보여 역 간 이용 패턴에 뚜렷한 이질성이 존재할 가능성을 확인하였다.

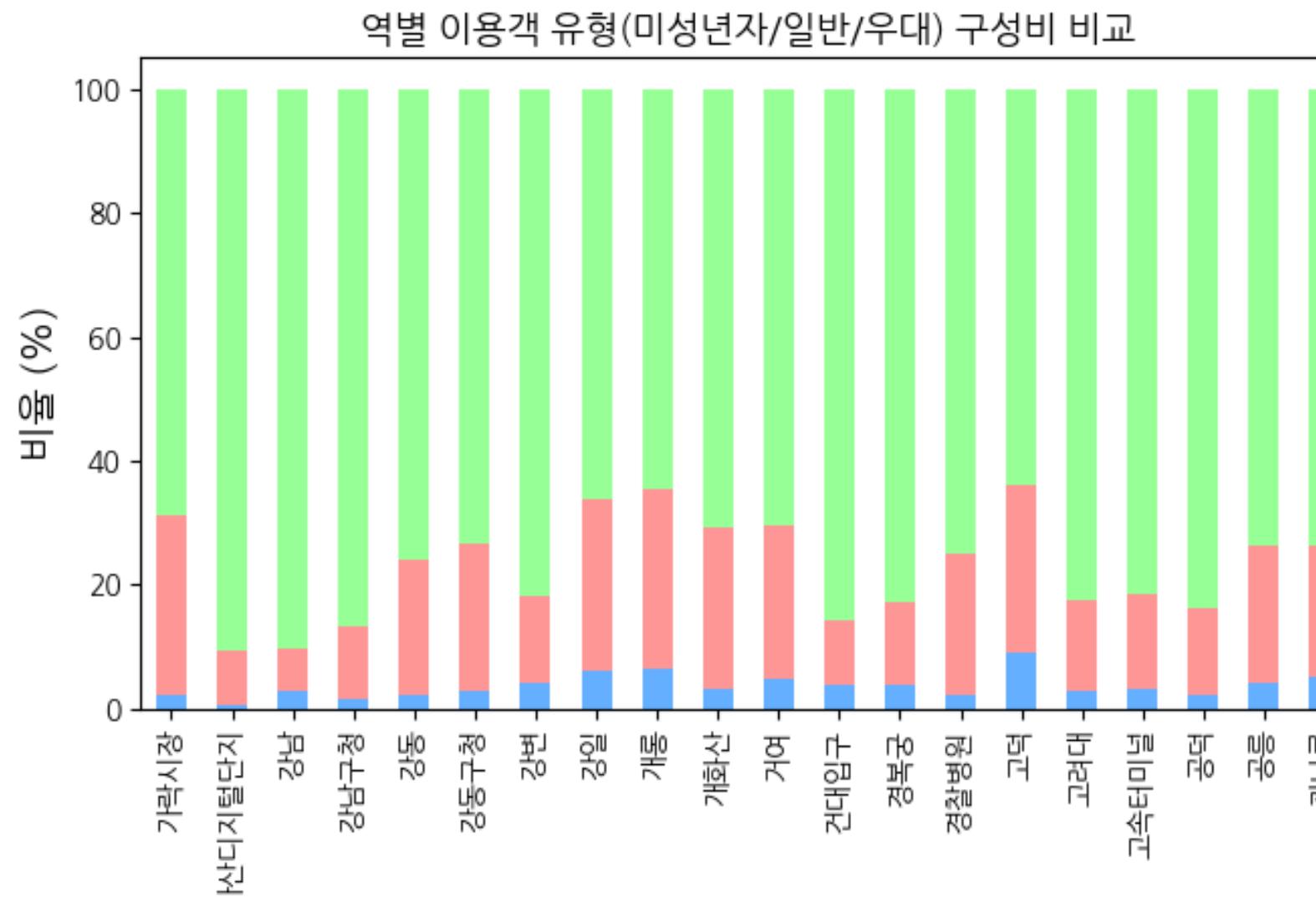


$$*CV(\%) = \frac{\text{표준편차}(\sigma)}{\text{평균}(\mu)} \times 100$$

역마다 특성이 다르기 때문에,
모든 역에 똑같은 광고를 하는 것보다
역별 맞춤형 광고를 하는 것이 훨씬 효율적!

03 타겟의 차별성 확인

“누가 타느냐가 역의 성격이다”



분석 개요

지표: 역별 이용자 유형(미성년자 / 일반 / 우대) 구성비

가설:

- 귀무가설: 지하철역과 이용객 구분은 독립이다. → 역에 따라 이용객 구성비 차이가 없다.
- 대립가설: 지하철역과 이용객 구분은 독립이 아니다. → 역의 특성에 따라 이용객 구성비가 통계적으로 다르다.

통계 검정 결과

- 카이제곱 독립성 검정
- P-value < 0.05

역에 따라 이용객 구성비는 유의하게 다르게 나타났다.



역마다 성격이 다르기 때문에,
모든 역에 똑같은 광고를 하는 것보다
역별 맞춤형 광고를 하는 것이 훨씬 효율적!

03 분석 접근 방향

역별 특성과 타겟



타겟에 맞는 광고 도출

역별 이용자수 규모를 알아보자



고효율 알짜역

해당역의 광고비나 등급



저비용으로 고효율을 내보자

04 데이터 전처리 프로세스

데이터 정제



결측치 및 중복값을 제거하여 데이터 품질 확보

괄호 제거



역명 및 데이터 내 불필요한 괄호 정보 삭제

시간대 집계



시간대별 정보를 일별 단위로 집계 및 통합

변수 생성



순유입, 활동성 지수 등 분석용 변수 산출

특성별 그룹화



주거/비주거, 연령대, 시장규모 등으로 분류

05 분석 방법론

01

역성격지수 및 로그 변환 활용

출퇴근 시간대 순유입 데이터를 활용하여 **로그 역성격지수를 산출** 기준값 0.4를 적용하여 **주거지역과 비주거지역**으로 분류

02

이용자 수 기반 시장 세분화

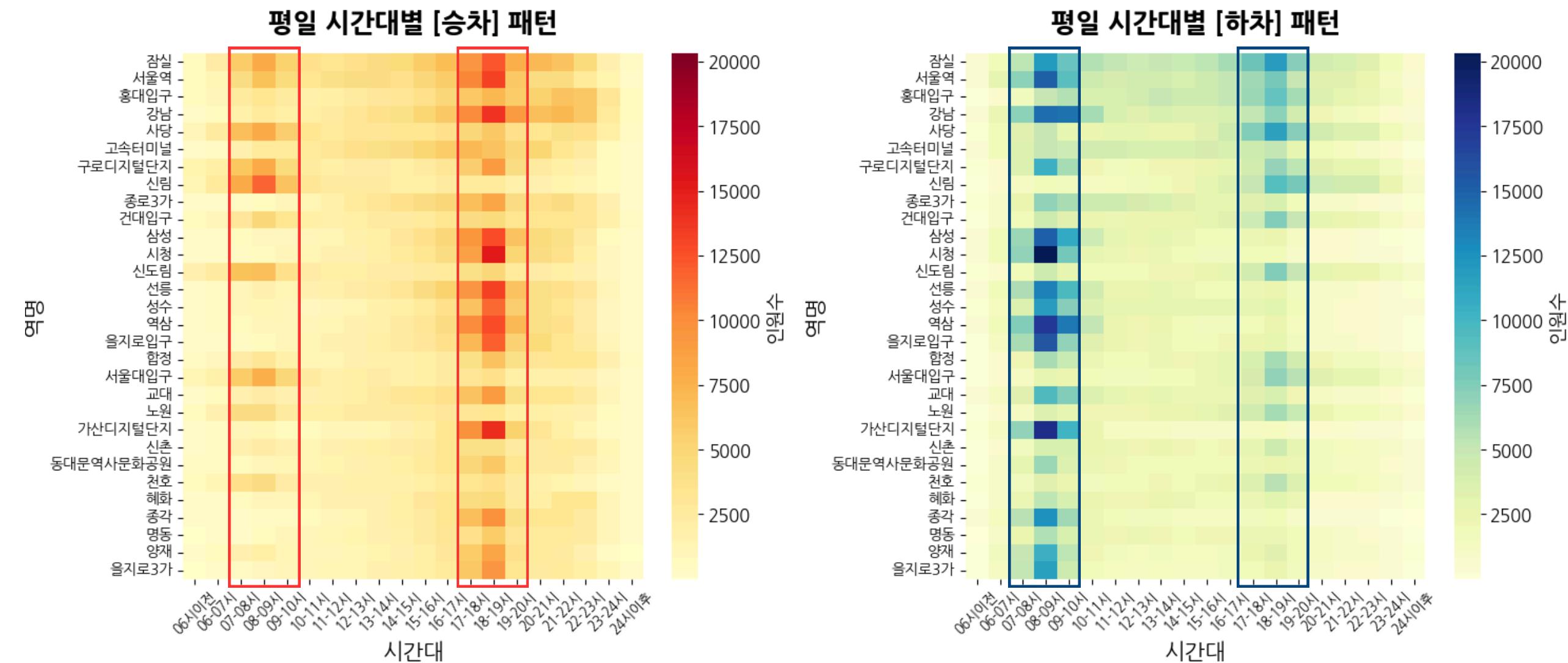
역별 이용자 수를 기준으로 시장 규모를 산정하고, **하위 20%~60% 구간의 역**을 선정하여 **비용 대비 효용이 높은 가성비 역** 도출

03

활동성 지수 기반 페르소나 설정

연령대별 이용 비율 **상위 30%** 역을 추출하고, 시간대별 순유입 데이터로 **활동성 지수를 계산**하여 4가지 페르소나 유형으로 분류

06 이용 패턴 분석 : “일별 이용 패턴”



시간대별 이용자 수 추이

상위 30개 역의 시간대별 이용자 수를 히트맵으로 분석한 결과, **오전 7-10시**와 **오후 5-8시**에 이용자가 집중되는 패턴을 확인

출퇴근 시간대 집중 현상

지하철 이용자들은 **출퇴근 시간대에 집중적으로 이동**하므로, 이 시간대를 기준으로 순유입 데이터를 산출하여 역별 특성을 파악

06 이용 패턴 분석 : “순유입 데이터 정의”

M_{on} = 출근 시간대 승차

M_{off} = 출근 시간대 하차

E_{on} = 퇴근 시간대 승차

E_{off} = 퇴근 시간대 하차

순유입 = 하차 - 승차

$$I_{station} = \frac{\left(\frac{M_{on}}{M_{off}} \right)}{\left(\frac{E_{on}}{E_{off}} \right)}$$

$$I_{log} = \ln(I_{station})$$

순유입 개념

순유입은 특정 시간대의
하차 인원에서 승차 인원을 뺀 값으로,
해당 역의 해당 시간대 이용자 유입량

양수일 경우 **유입지**, 음수일 경우 **유출지**로 간주

역성격지수

로그 역성격지수는

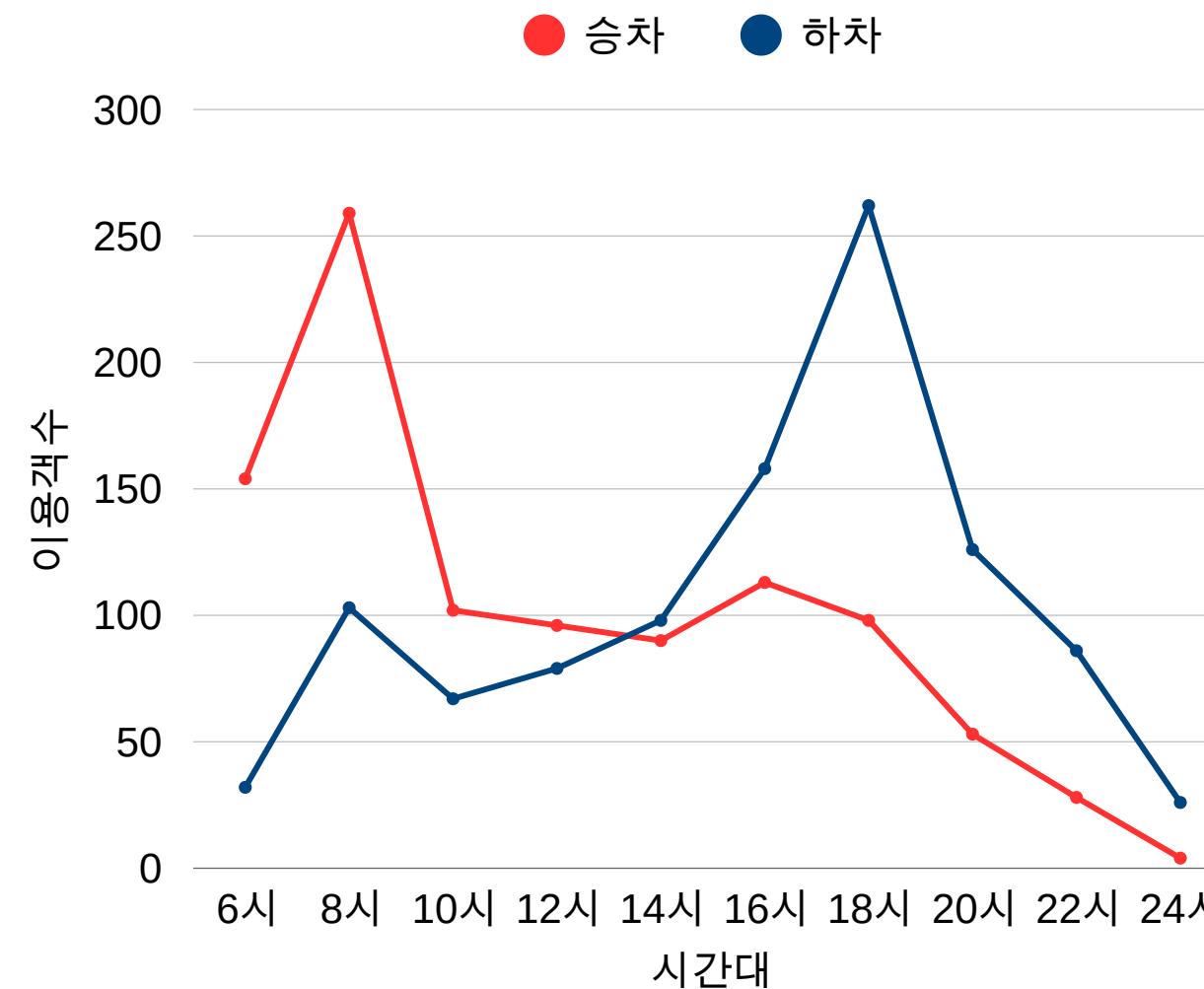
출근 시간대 승/하차 비율을 퇴근 시간대 승/하차 비율로 나눈 값에
자연로그를 적용한 지표

주거지역과 비주거지역을 효과적으로 구분 가능

+	출근 승차자 ↑ 퇴근 하차자 ↑	주거지
-	출근 하차자 ↑ 퇴근 승차자 ↑	비주거지

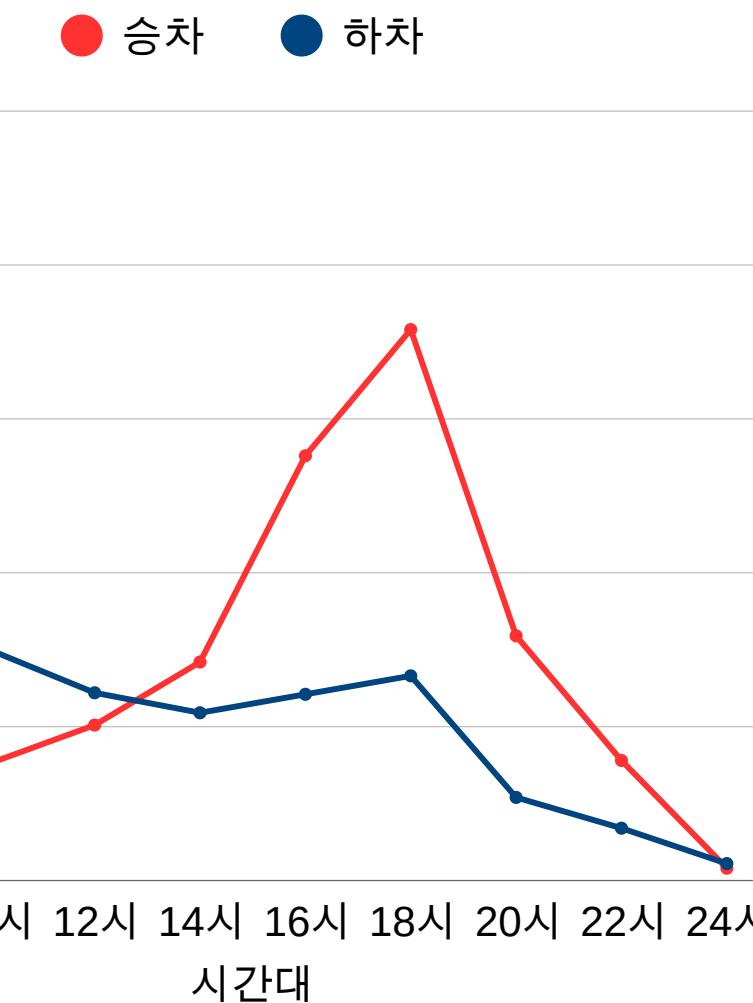
06 이용 패턴 분석 : “주거/비주거 구분”

로그 역성격지수를 기준으로 0.4를 임계값으로 설정하여 주거지역과 비주거지역을 구분
(출퇴근 시간대 순유입 비율이 약 1.5배 차이를 보이는 지점)



주거지역 특성

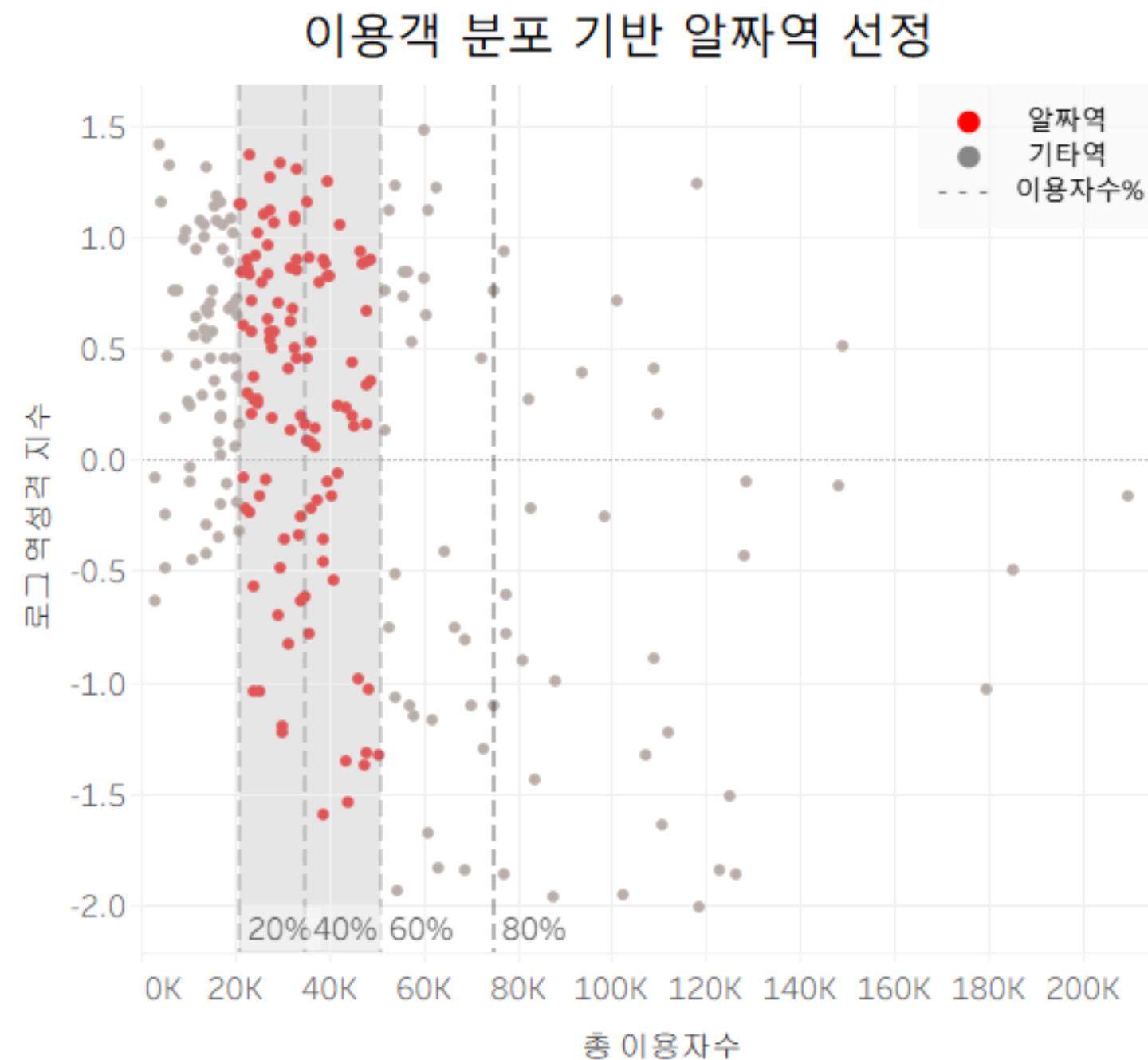
아침 시간대 순유출 발생
저녁 시간대 순유입 발생



비주거지역 특성

아침 시간대 순유입 발생
저녁 시간대 순유출 발생

06 이용 패턴 분석 : “시장 규모별 구분”



06 이용 패턴 분석 : “연령대별 구분”

연령대 매핑

일반

성인 이용객 (20~65세), 외국인, 직원포함

미성년자

외국인 포함

우대

만 65세 이상, 장애인, 국가유공자



연령대별 그룹화

각 연령대별 이용자 수 산출
전체 역 중 **상위 30%** 역 추출하여 그룹화



데이터프레임 구축

타겟 맞춤형 분석 기반 마련

06 이용 패턴 분석 : “페르소나 분석”

01

시간대별 분석

순유입 데이터 집계하여
출퇴근 외 시간대를 포함한
역별 이용 패턴 분석

02

활동성 지수 산출

분석한 이용 패턴을 기반으로
전체 역에 활동성 지수 산출
역의 활동 특성 수치화

03

페르소나 부여

활동성 지수 기반으로
‘페르소나’ 부여

04

4가지 페르소나

변화가

핫플레이스

생활밀착형

순수 주거형

06 이용 패턴 분석 : “활동성지수 및 페르소나”

낮 활동성

$$\frac{\text{평일 } 11\sim16\text{시 하차 합}}{\text{평일 일일 전체 하차 합}}$$

심야 활동성

$$\frac{\text{평일 } 22\sim24\text{시 승차 합}}{\text{평일 일일 전체 승차 합}}$$

주말 유동성

$$\frac{\text{휴일 } 11\sim17\text{시 하차 평균}}{\text{평일 } 11\text{시} \sim 17\text{시 하차 평균}}$$

휴일 나들이 지수

$$\frac{\text{휴일 } 12\sim18\text{시 하차 합}}{\text{휴일 일일 종 하차}}$$



변화가

낮 활동성 ↓
심야 활동성 ↑

핫플레이스

주말 유동성 ↑
휴일 나들이 지수 ↑

생활밀착형

낮 활동성 ↑
주말 유동성 ↓

순수주거형

심야 활동성 ↓
낮 활동성 ↓

07 분석 결과(평일)

* 시중 광고 등급 체계: SSA(최고등급) ~ C(최하등급), 출처: <https://brandp.co.kr/>

평일 일반 알짜역

역명	페르소나구분	광고등급
한양대	생활밀착형	A
한강진	핫플레이스	A
여의나루	핫플레이스	A
동대입구	핫플레이스	SA
안암	생활밀착형	A

평일 미성년자 알짜역

역명	페르소나구분	광고등급
종합운동장	핫플레이스	-
하남시청	생활밀착형	-
온수	순수주거형	SA
한양대	생활밀착형	A
남한산성입구	순수주거형	-

평일 우대 알짜역

역명	페르소나구분	광고등급
길동	생활밀착형	A
하남시청	생활밀착형	-
미아	생활밀착형	A
신정네거리	순수주거형	SA
중계	순수주거형	A

07 분석 결과(주말)

* 시중 광고 등급 체계: SSA(최고등급) ~ C(최하등급), 출처: <https://brandp.co.kr/>

주말 일반 알짜역

역명	페르소나구분	광고등급
논현	변화가	-
강남구청	변화가	SA
학동	변화가	SA
동대입구	핫플레이스	SA
청담	변화가	SA

주말 미성년자 알짜역

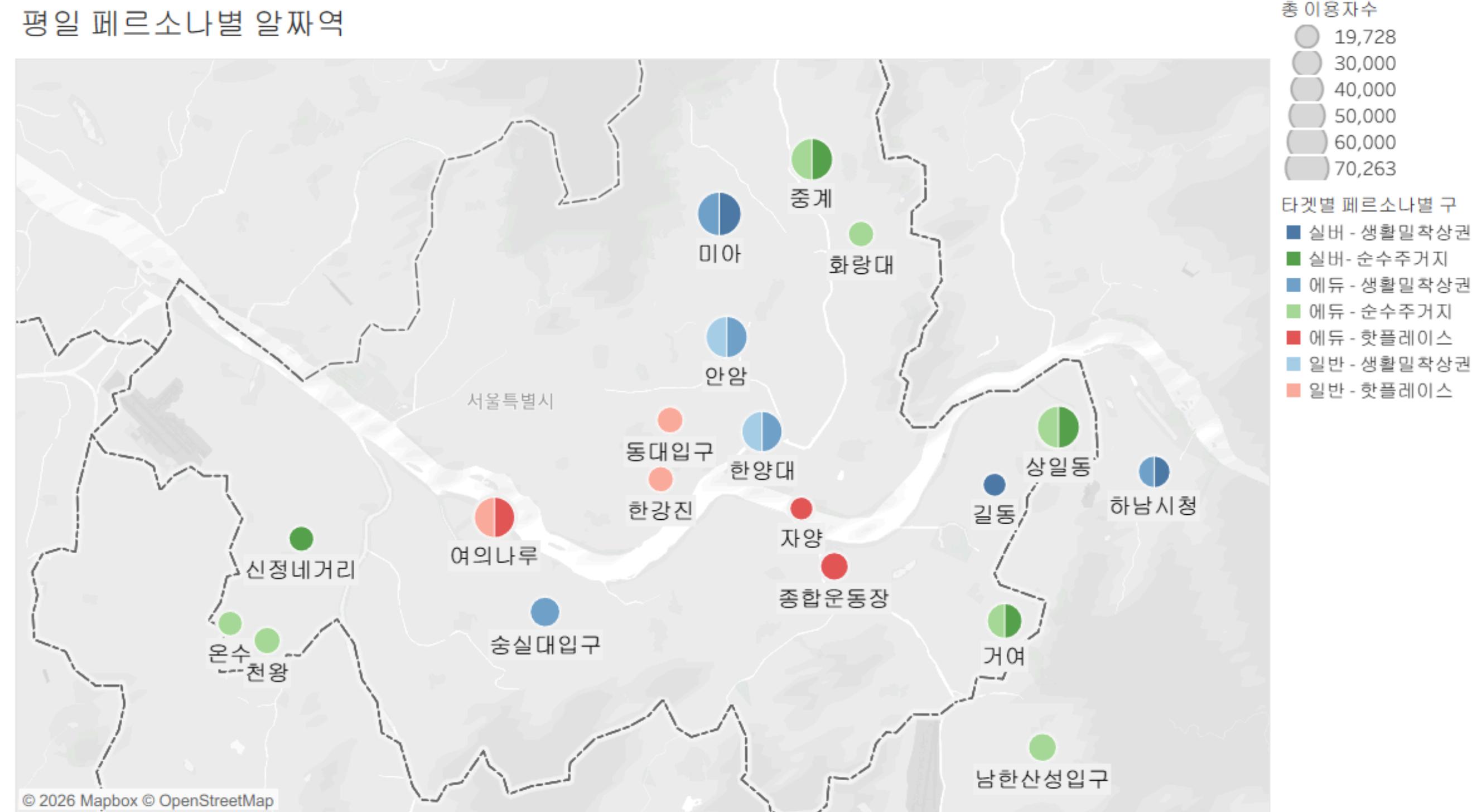
역명	페르소나구분	광고등급
이촌	핫플레이스	A
월드컵경기장	핫플레이스	A
상일동	순수주거형	-
하남시청	생활밀착형	-

주말 우대 알짜역

역명	페르소나구분	광고등급
하남시청	생활밀착형	-
미아	생활밀착형	A
신정네거리	순수주거형	SA
거여	순수주거형	A
상일동	순수주거형	-

07 분석 결과 : 평일 시각화 자료

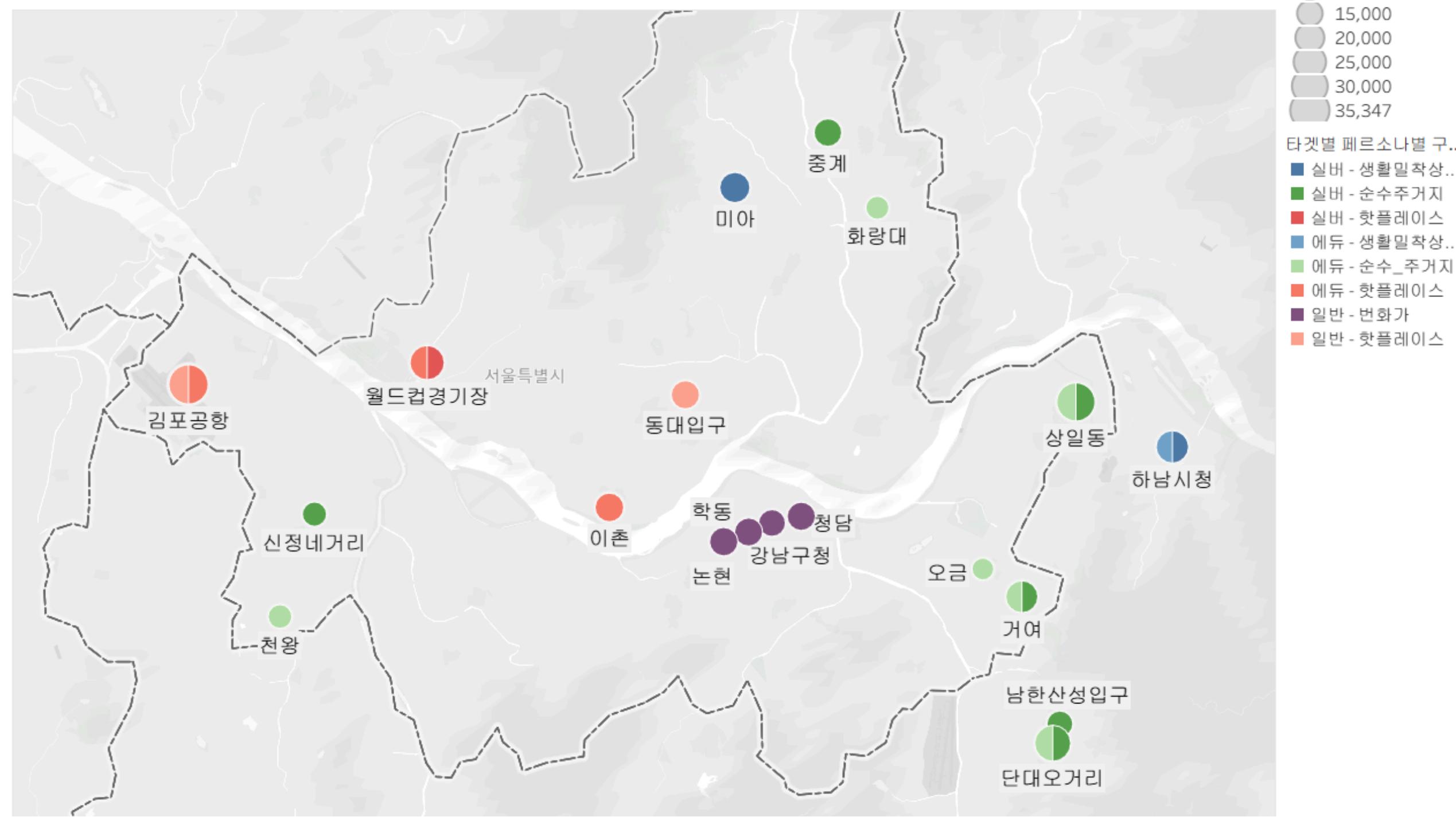
평일의 일반, 에듀, 실버의 알짜역 분포



07 분석 결과 : 주말 시각화 자료

주말의 일반, 에듀, 실버의 알짜역 분포

주말 페르소나별 알짜역



07 광고 추천 제안

타겟층	맞춤 광고	추천 역
일반	금융, 공연, 마트, 가전제품 등	한양대, 한강진, 여의나루 등
미성년자	학원, 인강, 게임, 연예인 등	온수, 한양대 등
우대	건강기능식품, 병원, 헬스케어, 보청기, 안마기 등	길동, 미아, 중계 등

07 분석의 한계점

본 연구는 데이터 기반 지하철 광고 전략을 제시하였으나,
분류 기준 설정 및 데이터 수집 과정에서 기인한 몇 가지 한계점이 존재함.

분류 기준의 주관성

시장 규모 구분 시 명확한 통계적 근거보다는 주관적 판단에 기반하여 임의로 하위 20%~60% 구간을 설정하여 추가적인 검증이 필요함.

데이터 구조적 한계

타교통수단으로 환승 인원 구분 불가

단순 하차 ≠ 실제 체류

본 데이터셋에서는 환승 정보를 따로 분류하지 않았기에, 환승객이 많은 역은 상업지로 잘못 판단되었을 가능성이 있음.

지하철 데이터만 사용

실제 상권 유동 인구와 차이 발생 가능성이 존재함.

비용 정보 부재

광고 집행 비용 데이터를 사전에 확보하지 못해 실제 비용 대비 효과 분석이 제한적이었으며, 추가적인 연구를 통한 보완이 필요함.

08 프로젝트 회고



주제 선정과 갈등 해결

프로젝트 초기 주제 선정 과정에서 팀원 간 의견 차이가 있었으나, 멘토링과 소통을 통해 합의점을 도출함.

결과적으로, 공통의 목표를 갖고 업무에 몰입할 수 있었음.

협업과 소통의 성과

원활한 소통과 개인 역량에 맞춘 역할 분담, 노션/디스코드에서의 정기적인 피드백과 멘토링, 그리고 캔바 등 협업 도구 활용을 통해 효율적인 작업이 가능했음.

감사합니다