

K-LEAGUE 예매 분석

- 인천 유나이티드 -

데이터분석가 이진우

목차

1. 프로젝트 개요
2. 프로젝트 구성
3. 데이터 전처리
4. 데이터 분석
5. 프로모션 제안 및 회고

프로젝트 개요



K LEAGUE

K리그는 한국프로축구연맹 산하의 대한민국 프로축구 리그입니다. 1983년 슈퍼리그라는 명칭으로 출범하였으며, 현재까지 AFC 챔피언스 리그 우승 횟수가 가장 많은 리그이기도 합니다.

코로나 19

코로나19로 인해 피해를 입었던 관광·스포츠·공연업계의 제한이 해제되면서 관객이 조금씩 늘어나는 추세입니다. 2023년에는 더 많은 관객이 방문 할 것으로 예상하고 있습니다.

월드컵

작년 월드컵의 뜨거운 열기로 K-리그에도 관심이 많아진 지금, 적절한 전략을 통해 관객을 늘리는 방안을 검토해보았습니다.

프로젝트 구성

비즈니스 분석	데이터 전처리	데이터 분석	제안 및 회고
<ul style="list-style-type: none">• K LEAGUE• 인천 유나이티드	<ul style="list-style-type: none">• MEMBERSHIP.CSV• RESERVE_TICKET.CSV• D95E264436CF42DE.XLSX• SCHEDULE2022120221019212245.XML	<ul style="list-style-type: none">• 구매 패턴 분석• 고급좌석 이용 고객 분석	<ul style="list-style-type: none">• 프로모션 제안• 회고

인천 유나이티드

2003년에 창단 된 인천광역시의 K리그1 소속 프로축구단

전술력

시즌 초반 주로 하위권에 머물지만 적극적인 여름 이적시장 활용으로 하반기에 좋은 성적을 거둡니다.

생존왕

부족한 재정과 잦은 선수 이탈에도 승강제 도입 이후 한 번도 강등하지 않은 유일한 팀으로 하반기로 갈수록 팬들의 응원이 뜨겁습니다.



데이터 전처리

membership.csv

고객번호, 나이, 직업, 좋아하는 선수

job, zipcode가 결측치이고 좋아하는 선수가 무고사
회원가입 기본 설정으로 추정됨

-> zipcode는 0, job은 '기타'로 분류

reserve_ticket.csv

주문번호, 고객번호, 주문시간, 경기시간
취소유무, 티켓번호, 예매좌석 정보

중복컬럼 제거(detail_cancel_date_time)

예매,취소 시간컬럼 병합

상대팀 정보 추가

스카이라운지구역 결측치 발견

-> seat_Attribute_5 컬럼을 'sky'로 분류

d95e264436cf42de.xlsx

일자별 인천유나이티드 순위

2022년 경기 순위만 남기고 제거

schedule2022120221019212245.xml

경기시간, 경기정보, 관객 수

한글 인코딩 문제가 발생

-> 디코딩 후 인코딩을 진행하고 다시 디코딩하여 문제 해결

승부차기 진행하지 않아 home_team_tk, away_team_tk 컬럼 삭제

게임시간 컬럼을 datetime형식으로, 요일 컬럼을 int로 변경

데이터 분석



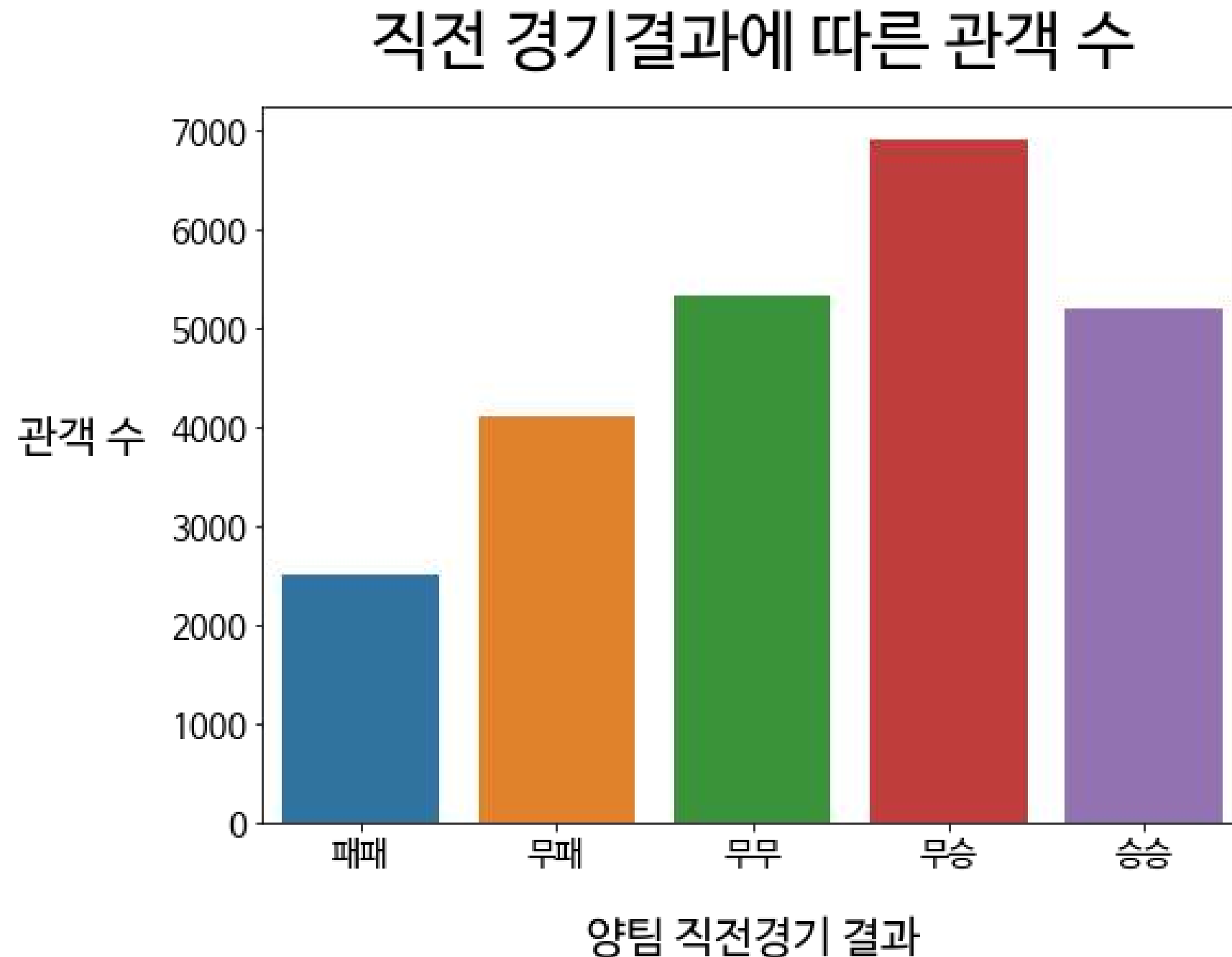
가설 1

직전 경기 결과와 순위가 예매에 영향을 미친다.



직전 경기 결과

두팀 모두 직전 경기를 패배 했을 때 관중의 수가 적어지는 경향이 있다.

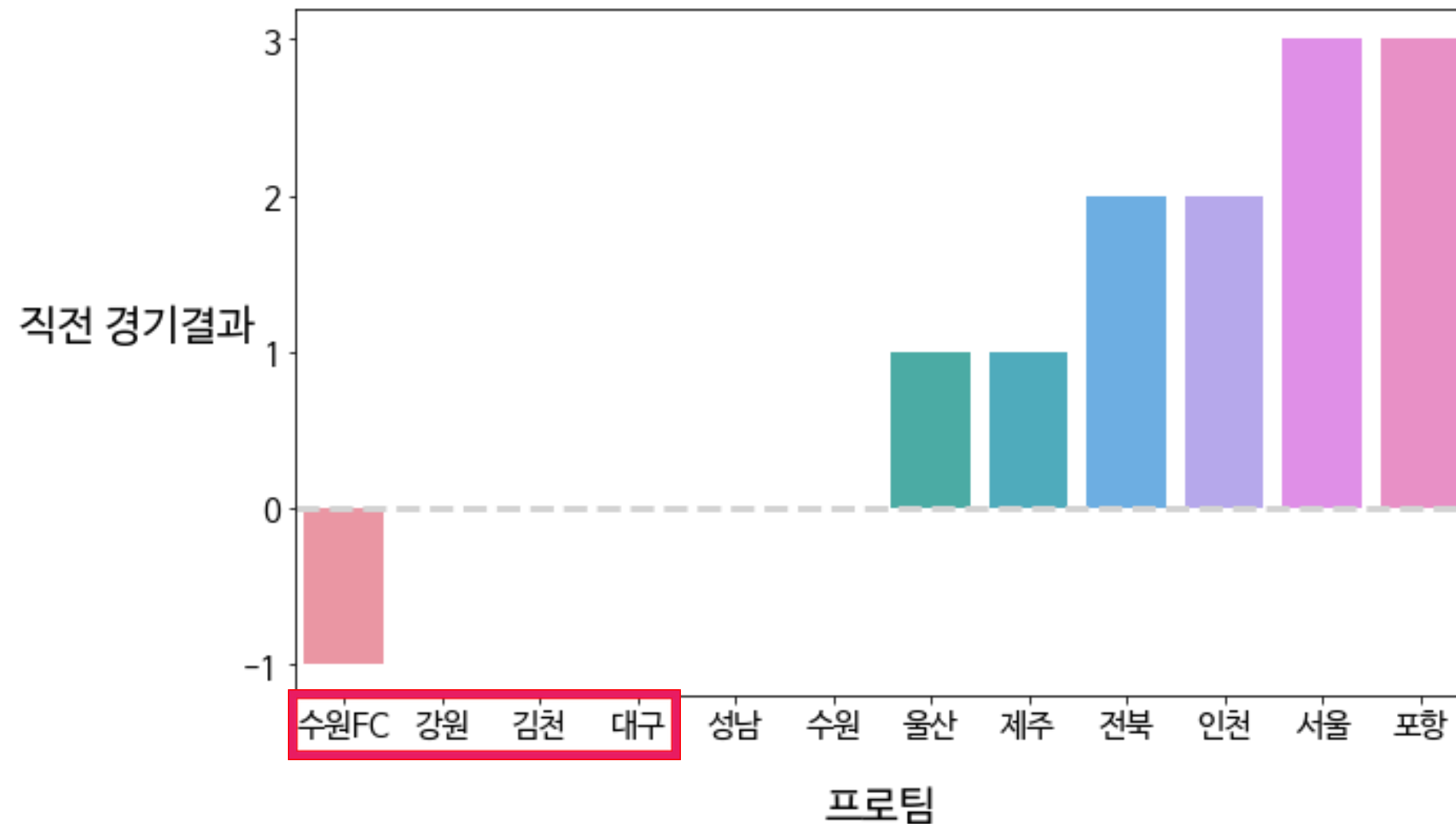


직전 경기 점수 합계

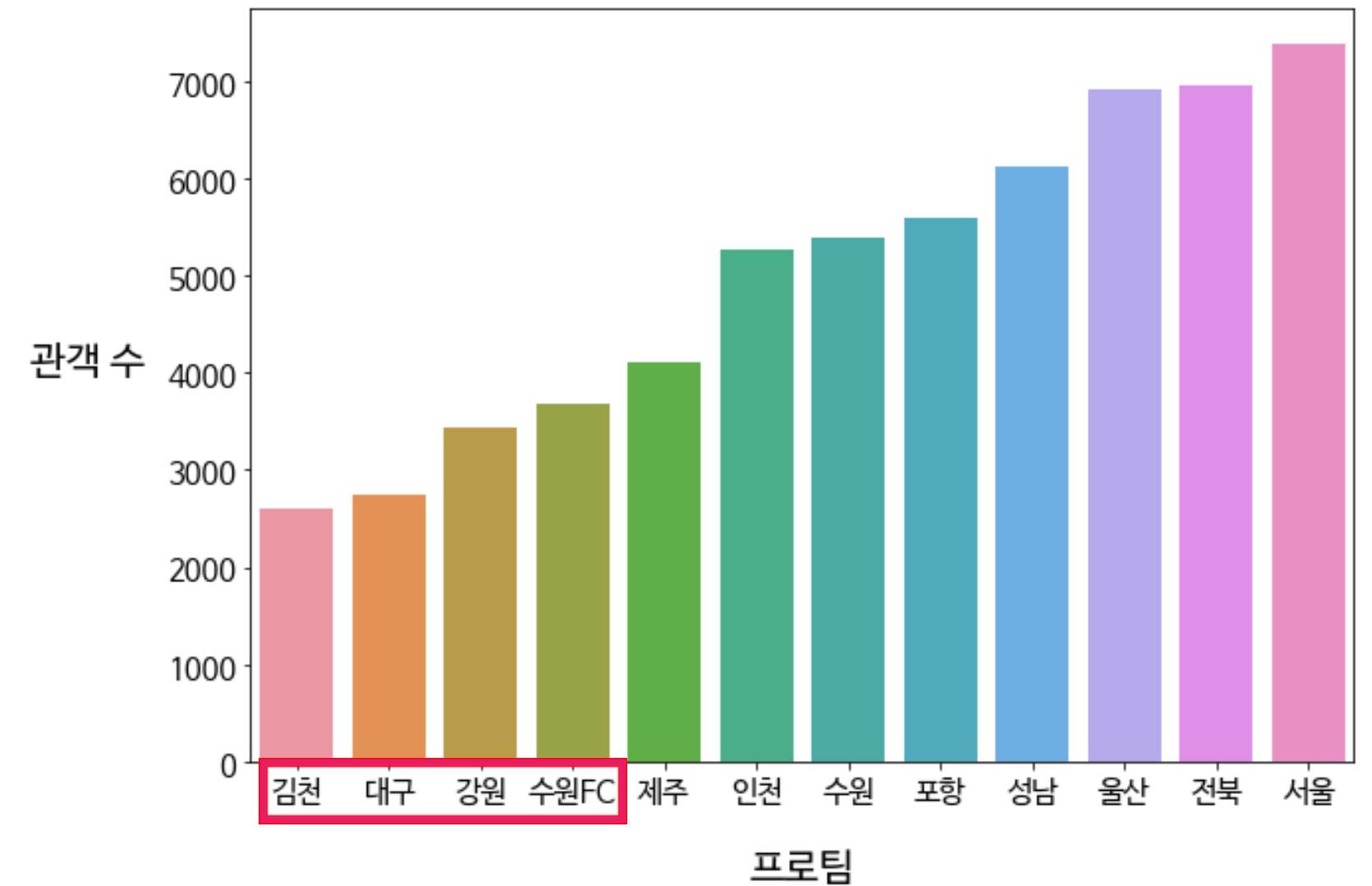
직전경기 : 승리 +1 / 패배 -1 / 무승부 0

직전경기 하위권 팀의 관객 수도 하위권

직전 경기결과 점수



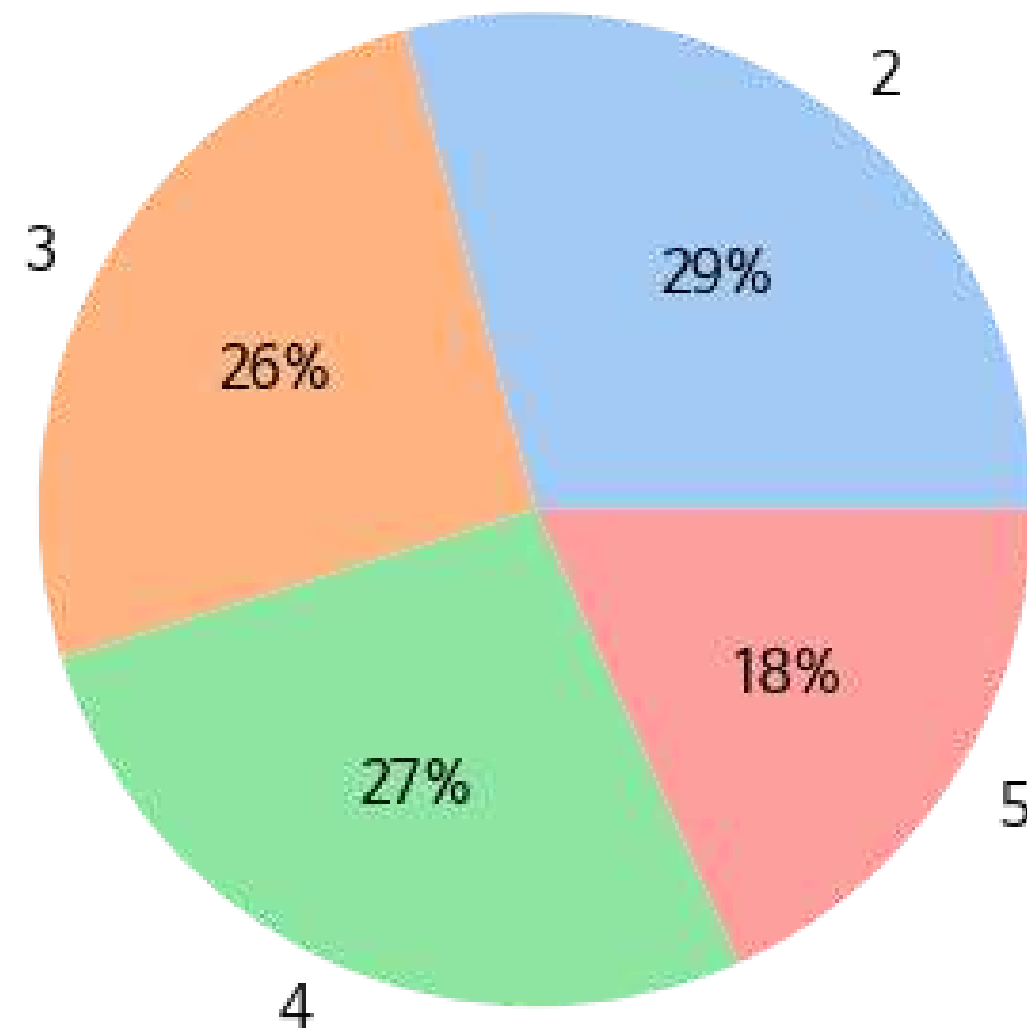
프로팀 별 평균 관객 수



인천유나이티드 순위에 따른 관객 수

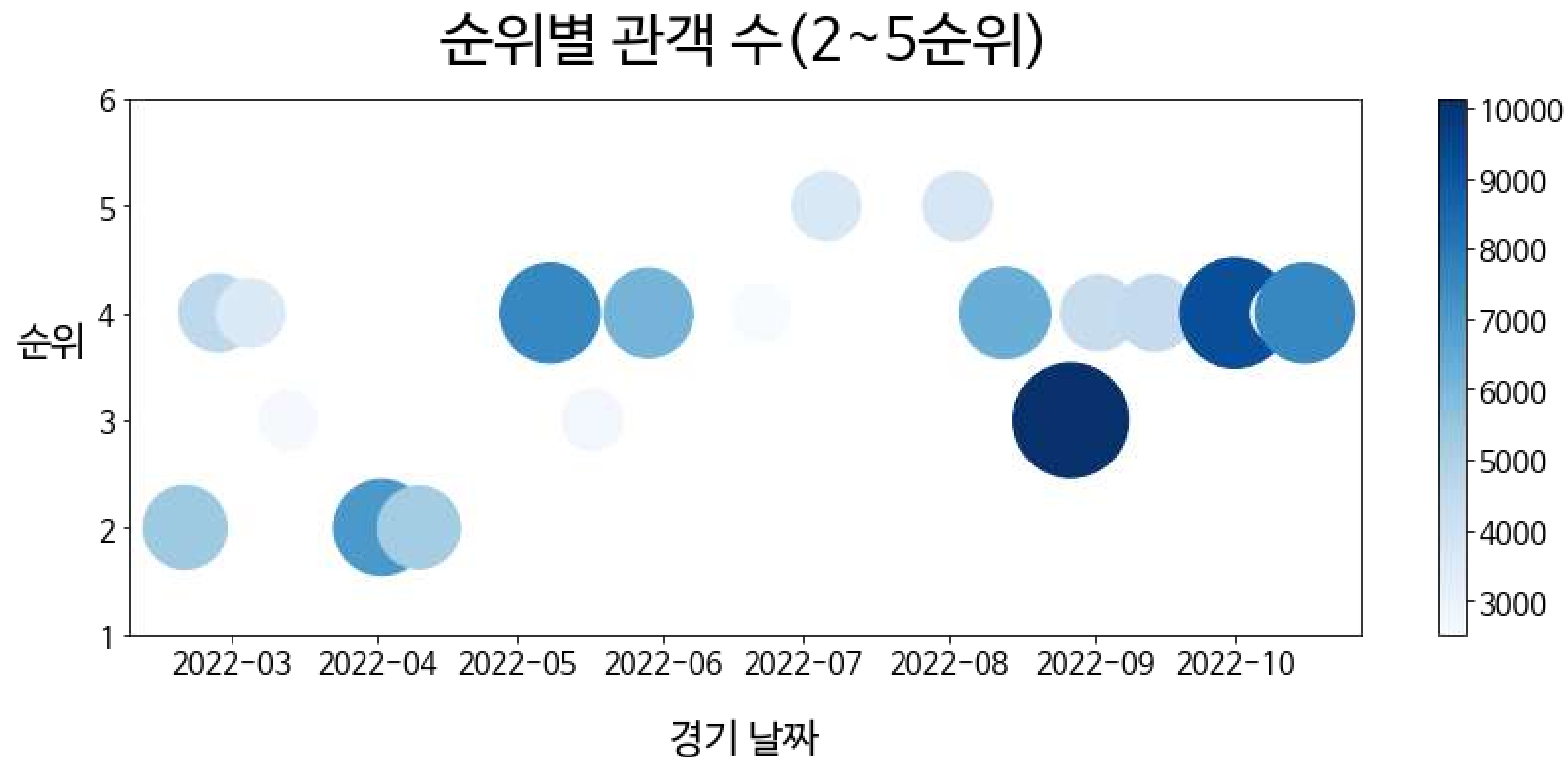
순위가 높을 때 평균 관객수가 높은 편이다.

순위별 평균 관객 수 비율



인천유나이티드 순위에 따른 관객 수

구체적으로 살펴보면 높은 순위보다는 순위의 변화가 있을 때 주로 관객 수의 변화가 관찰됨
2022년의 자료만으로 순위에 따른 관객 수를 단정짓는 것은 어려워보임



가설 1 결론

직전 경기 결과와 순위가 예매에 영향을 미친다.

직전 경기 결과

- 직전 경기결과가 안좋은 팀간의 경기는 관객 수가 적을 것으로 예측할 수 있다.
- 인천이 패배한 경우 관람객 위로 및 재방문 독려를 위해 일부 고객에게 쿠폰 제공

순위

- 순위와 관객 수는 상관관계가 없는 것으로 보인다.

가설 2

새로운 패턴을 가진 고객이 있을 것이다.



패턴1. 경기 예매 시간

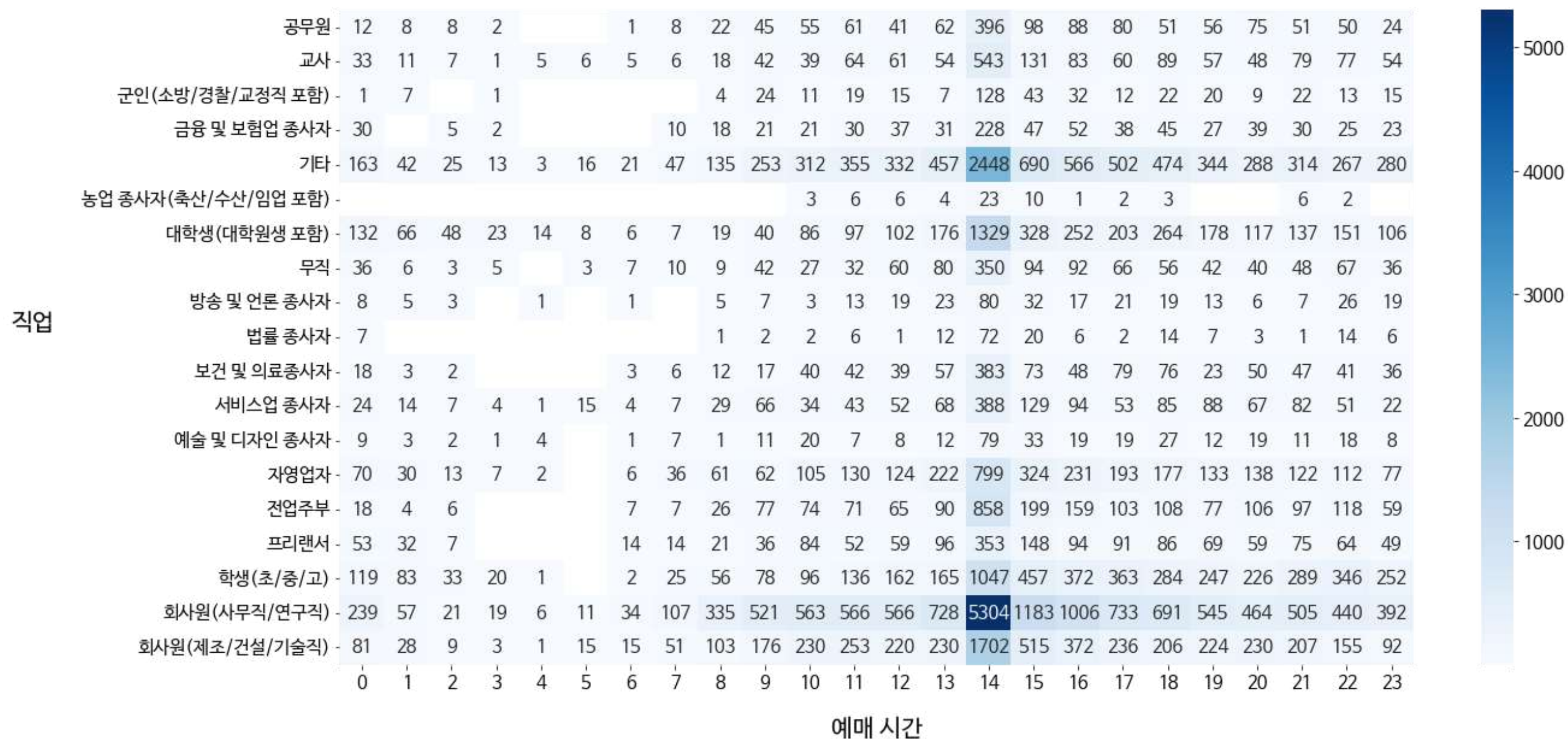
경기를 예매하는 시간은 월, 화, 수요일 14시



패턴1. 경기 예매 시간

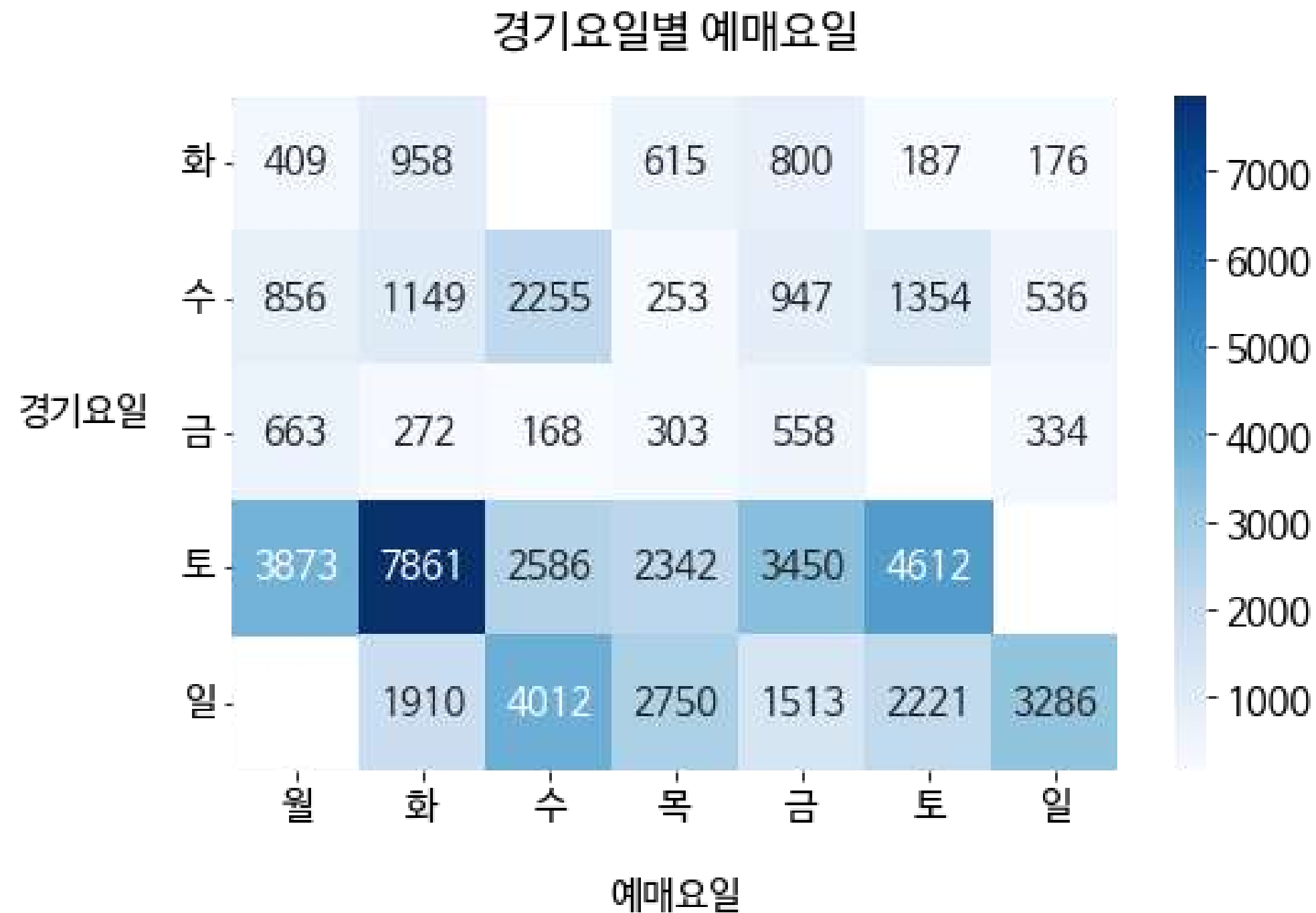
주로 직업이 **회사원**이거나 **기타인** 회원

직업별 경기에매 시간



패턴1. 경기 예매 시간

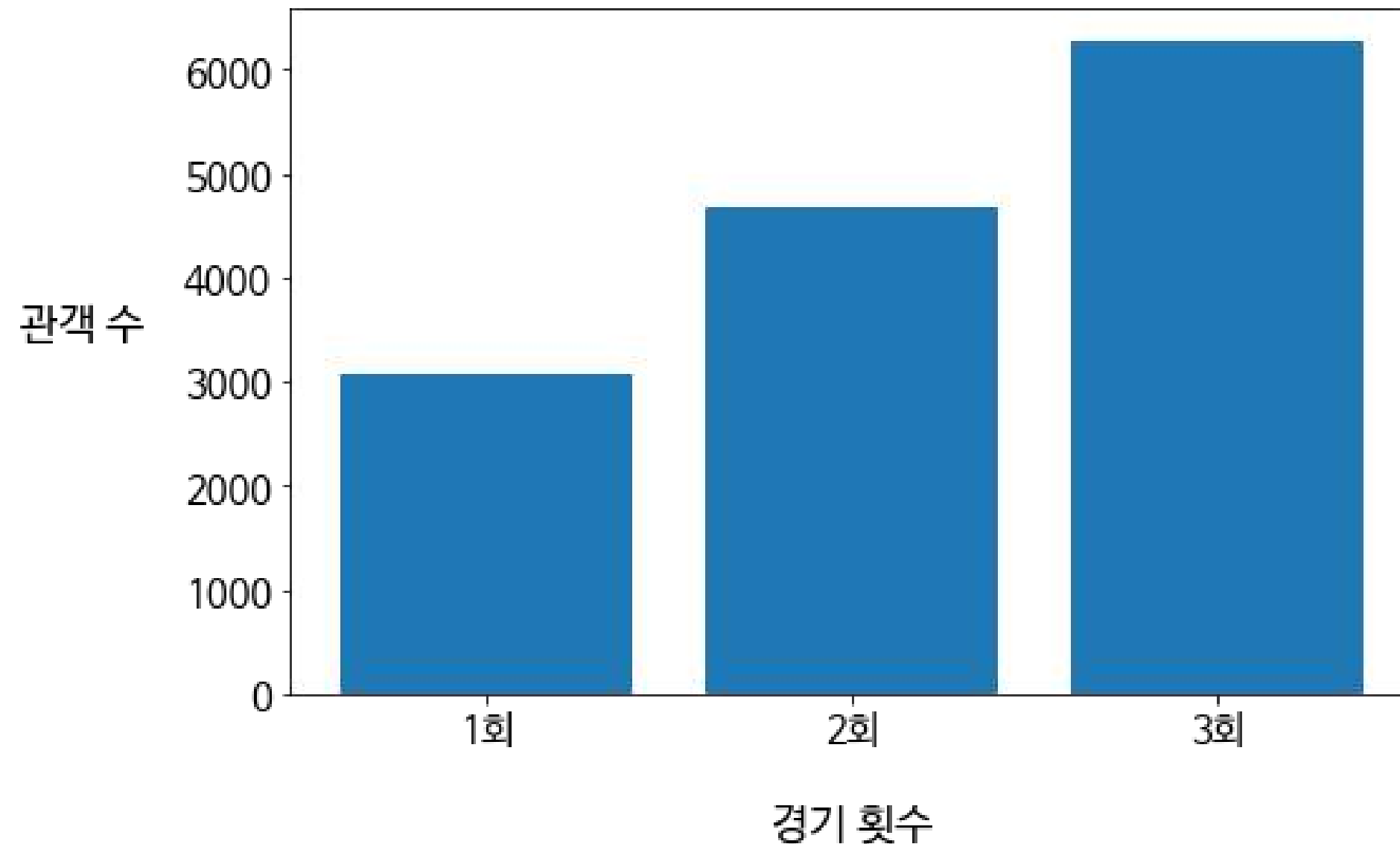
예매는 주로 경기 **당일**이나 경기 **3일 후** 이루어짐



패턴2. 경기 횟수가 잦을수록 관객 수가 많다.

월 경기 횟수가 잦을수록 경기장을 이용하는 평균 관객 수가 늘어난다.

경기 횟수별 평균 관객 수

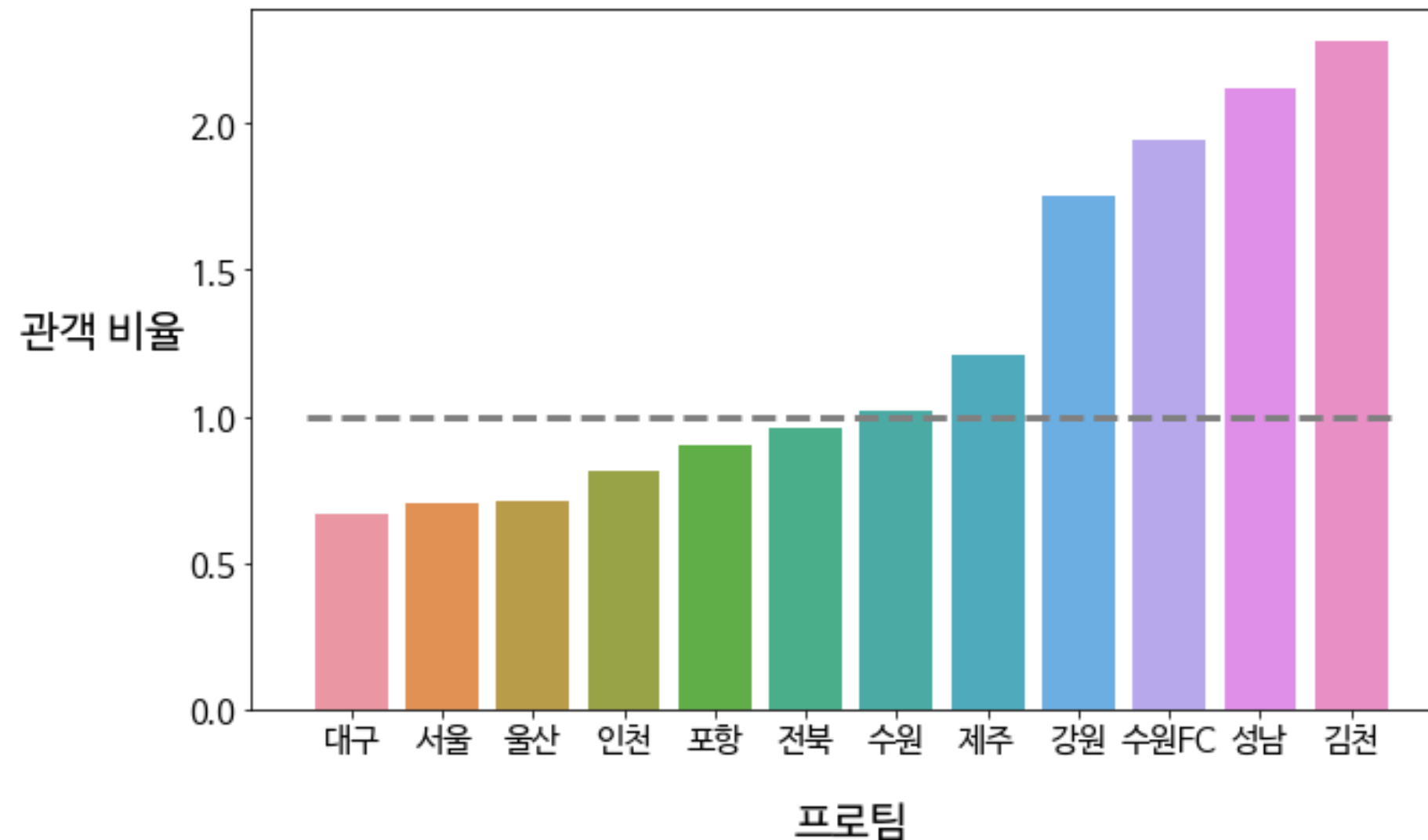


패턴3. 홈경기 관객 비율이 높다.

홈경기의 관객 수가 원정 경기의 관객 수 보다 많다.

관객 비율 = 원정경기 평균관객수 / 홈경기 평균 관객 수

원정경기 관객 비율



가설 2 결론

경기 예매 피크 타임

- 월요일, 화요일, 수요일 14시 경 많은 회사원, 기타 고객이 예매한다.
- 이는 경기 당일과 경기 3일 후에 예약하는 패턴 때문인 것으로 파악된다.
- 이 시간에 고급좌석 이용 경험이 있는 고객에게 고급좌석 얼리버드 할인 쿠폰을 지급하여 노출효과를 올린다.

월별 경기 횟수

- 월별 경기 수가 많을수록 관객 수도 늘어난다.
- 관객 수가 많을수록 좌석의 쾌적함이 떨어지므로 경기가 많은 달에 고급좌석 할인을 진행

당일예매 고객

- 당일예매 고객은 지속적으로 당일예매를 이용한다. 이는 경기 수요의 예측을 방해하고 운영비용을 추가로 발생시킨다.
- 온라인 예약 시 할인 받을 수 있는 쿠폰을 지급하여 온라인 예매를 이용하도록 유도한다.

홈경기 관객 비율

- 홈경기의 관객비율이 높다는 것은 홈팬들이 많이 온다는 뜻일 수 있지만 원정팬들의 진입장벽이 높다는 의미로도 해석될 수 있으므로 원정팬을 위한 행사도 진행할 필요가 있다.

가설 3

구매유도에 호응하는 고객이 있다.

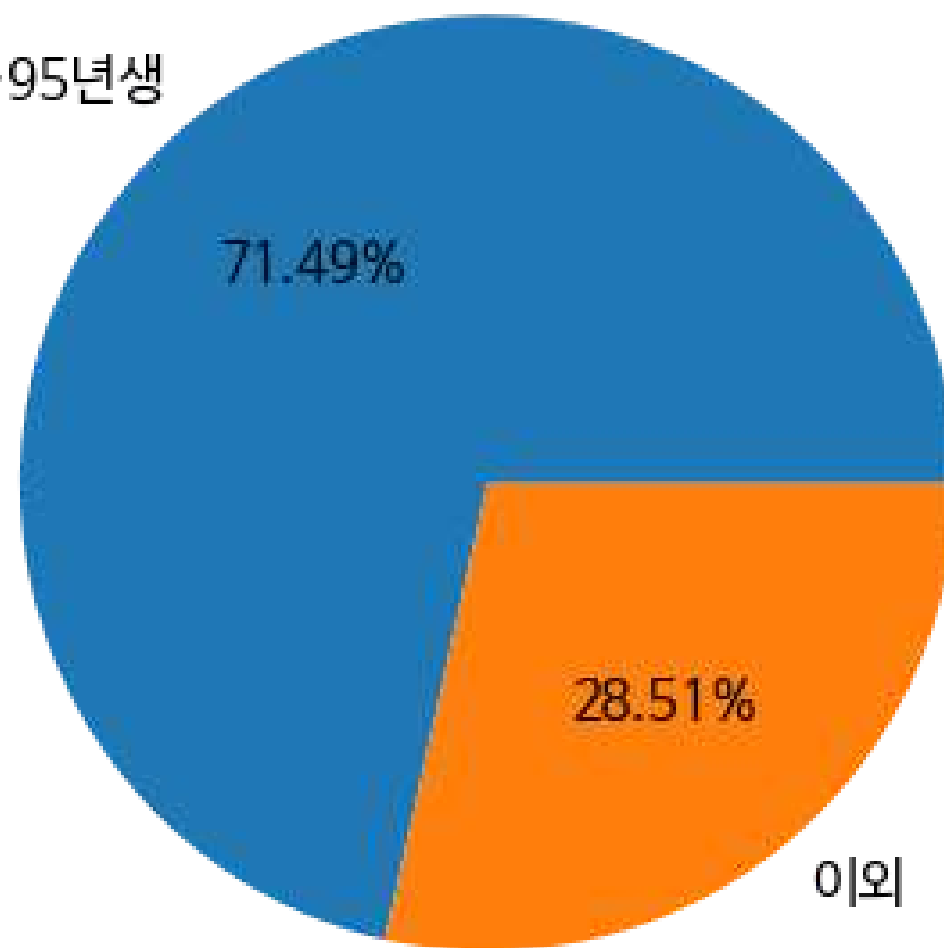


구매유도에 응하는 고객

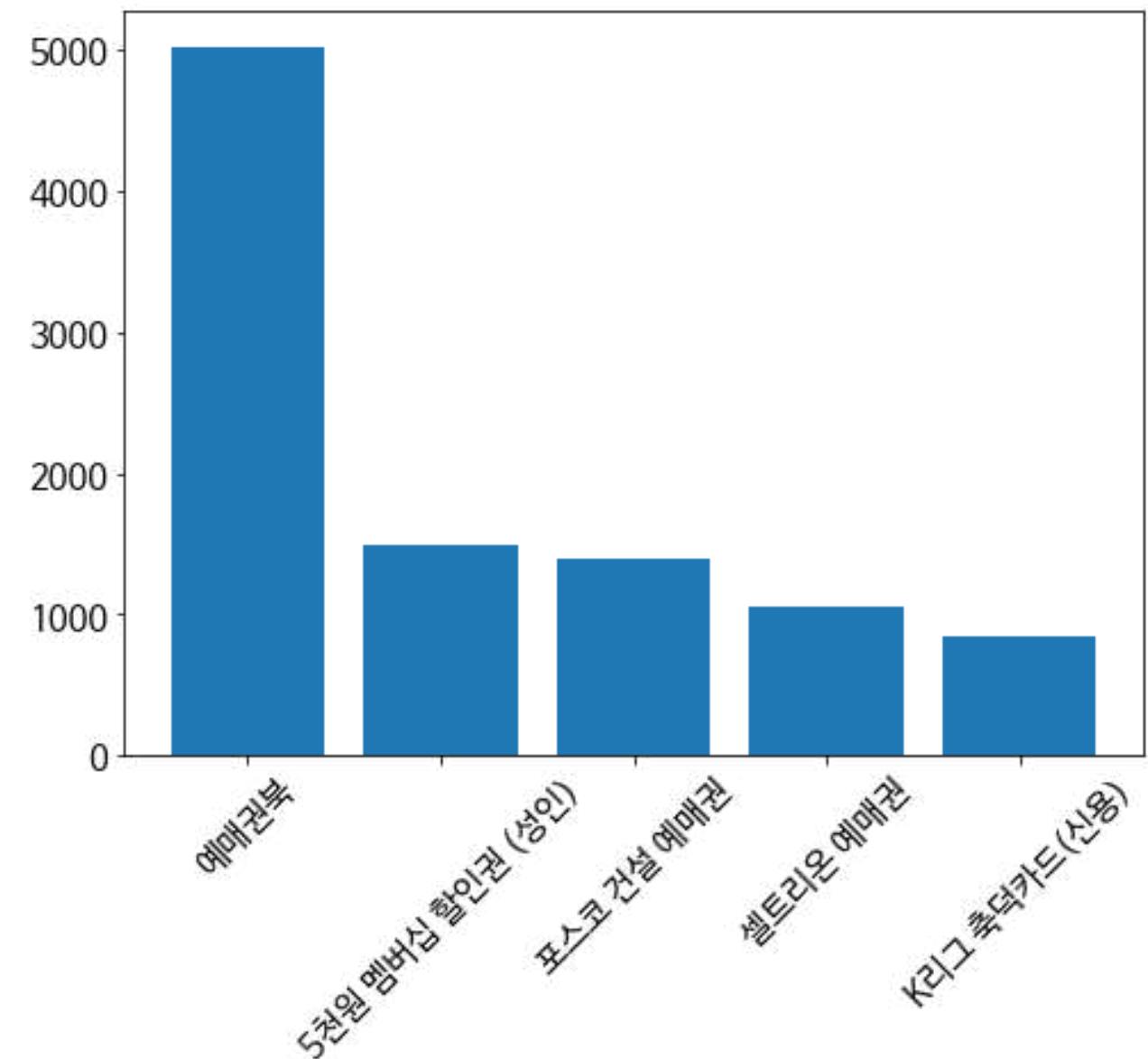
할인권 사용 고객의 70% 이상이 **75~95년생**
가장 인기가 많은 할인은 **예매권북**

할인권 사용 나이 비율

75년생~95년생



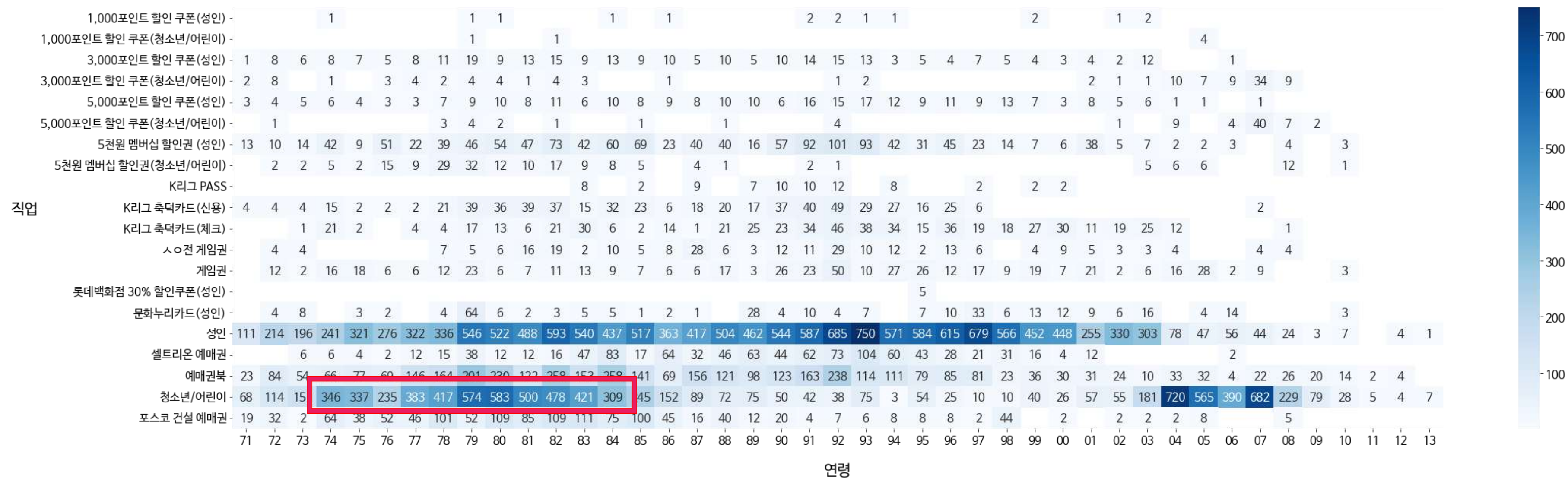
할인권 사용순위



구매유도에 응하는 고객

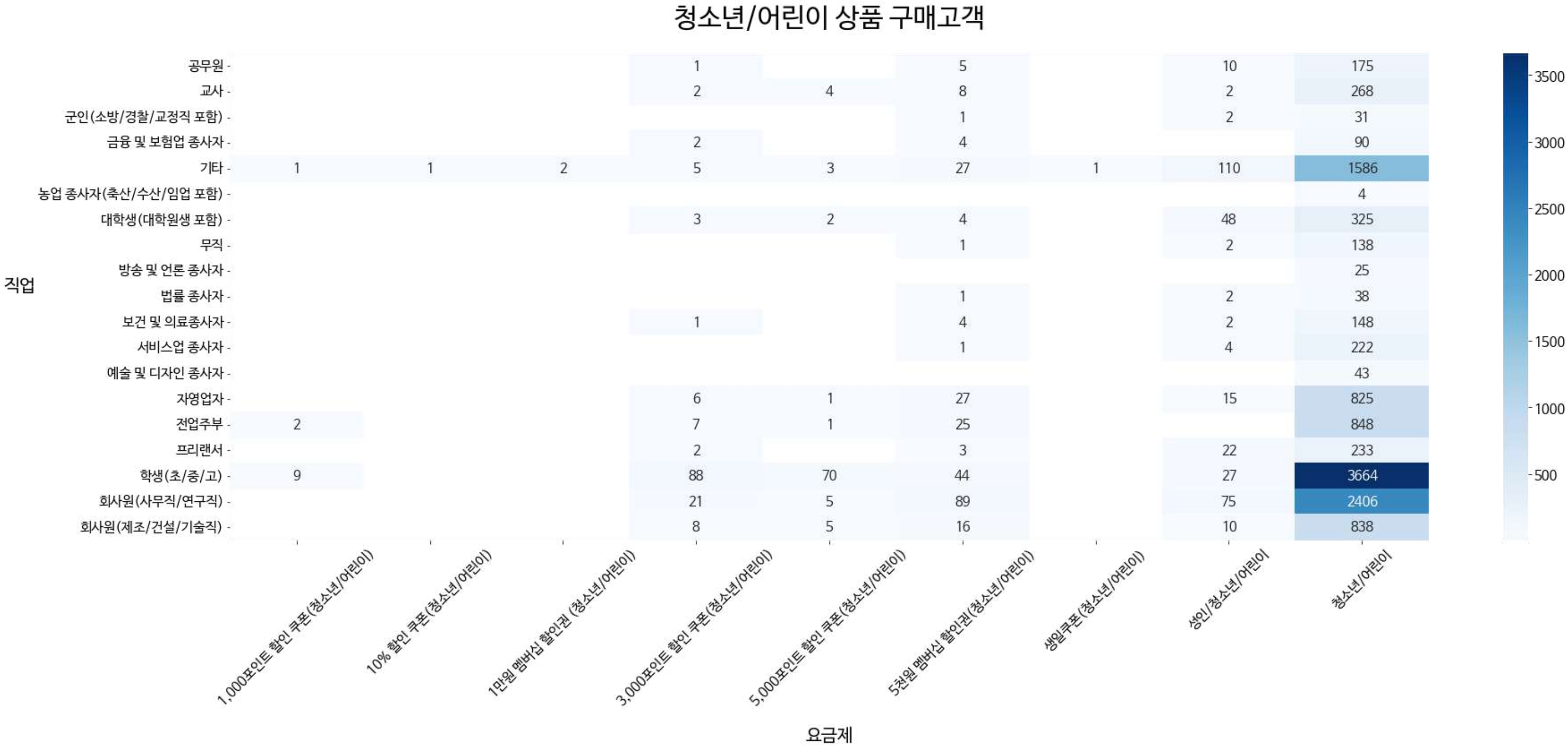
70년대 중반~ 80년대 중반 청소년/어린이 티켓 구매

Top4 구역 연령별 예매상품



구매유도에 응하는 고객

청소년 / 어린이 티켓 구매자들의 쿠폰 사용률이 현저히 낮음



가설 3 결론

구매유도에 호응하는 고객이 있다.

호응고객

- 쿠폰 사용자의 70%가 75년생~95년생이다.
- 가장 인기있는 할인은 예매권북이다.

비호응고객

- 70년대초반~ 80년대 초반 청소년/어린이 티켓 구매자들은 쿠폰 사용비율이 적다.
- 미성년자 관객의 비율은 전체의 약 24%를 차지한다.
- 가족할인 혜택을 제공하여 구매호응을 유도할 수 있다.

가설 4

고급좌석 예매 고객 분석



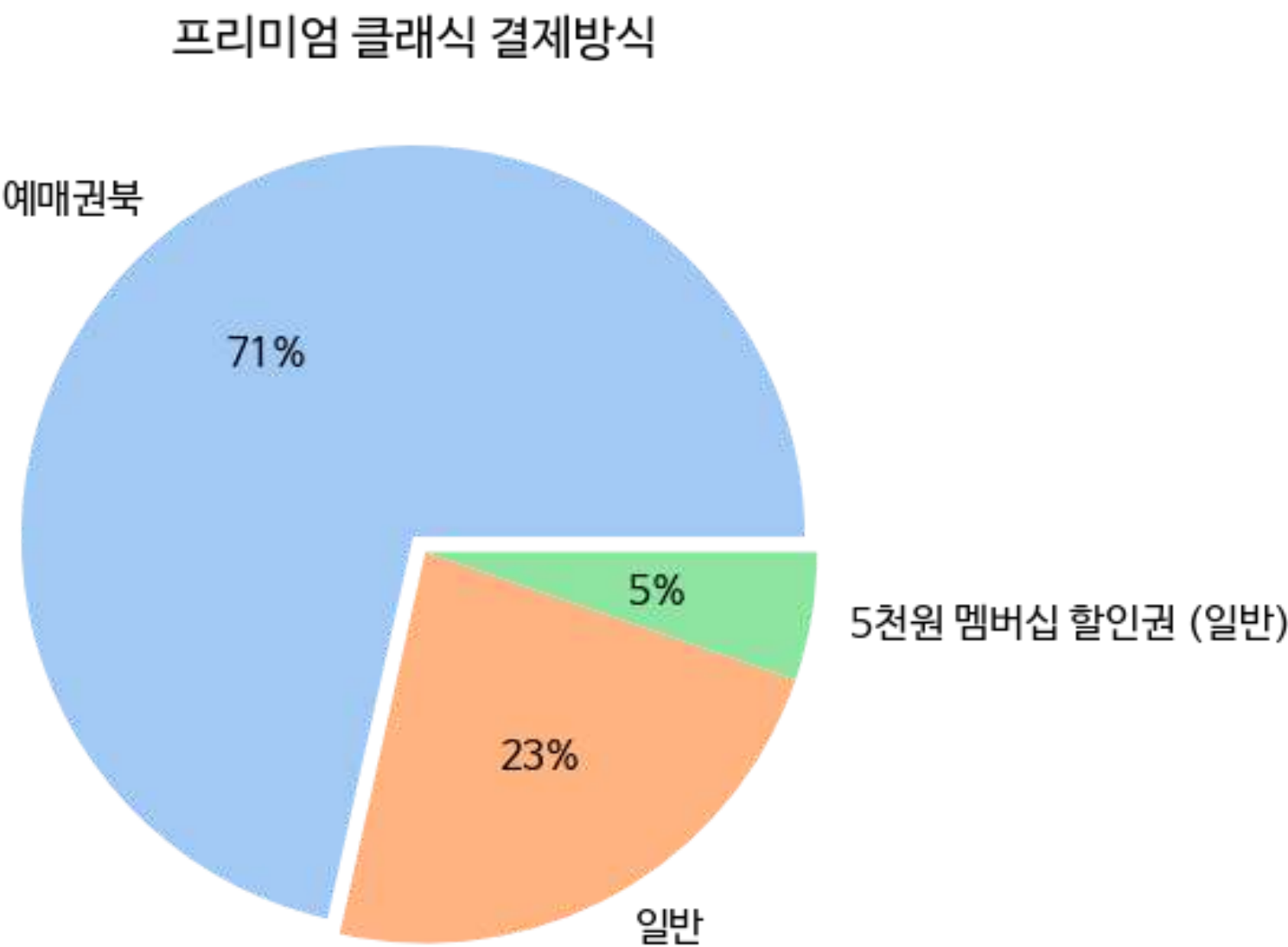
고급좌석 이용자

주로 70~90년대생 직장인이 이용



고급좌석 이용자 (프리미엄 클래식)

프리미엄 클래식 고객의 70% 이상이 **예매권북**을 사용한다.
해당 고객군은 **청소년 / 어린이** 구매 이력이 적다.

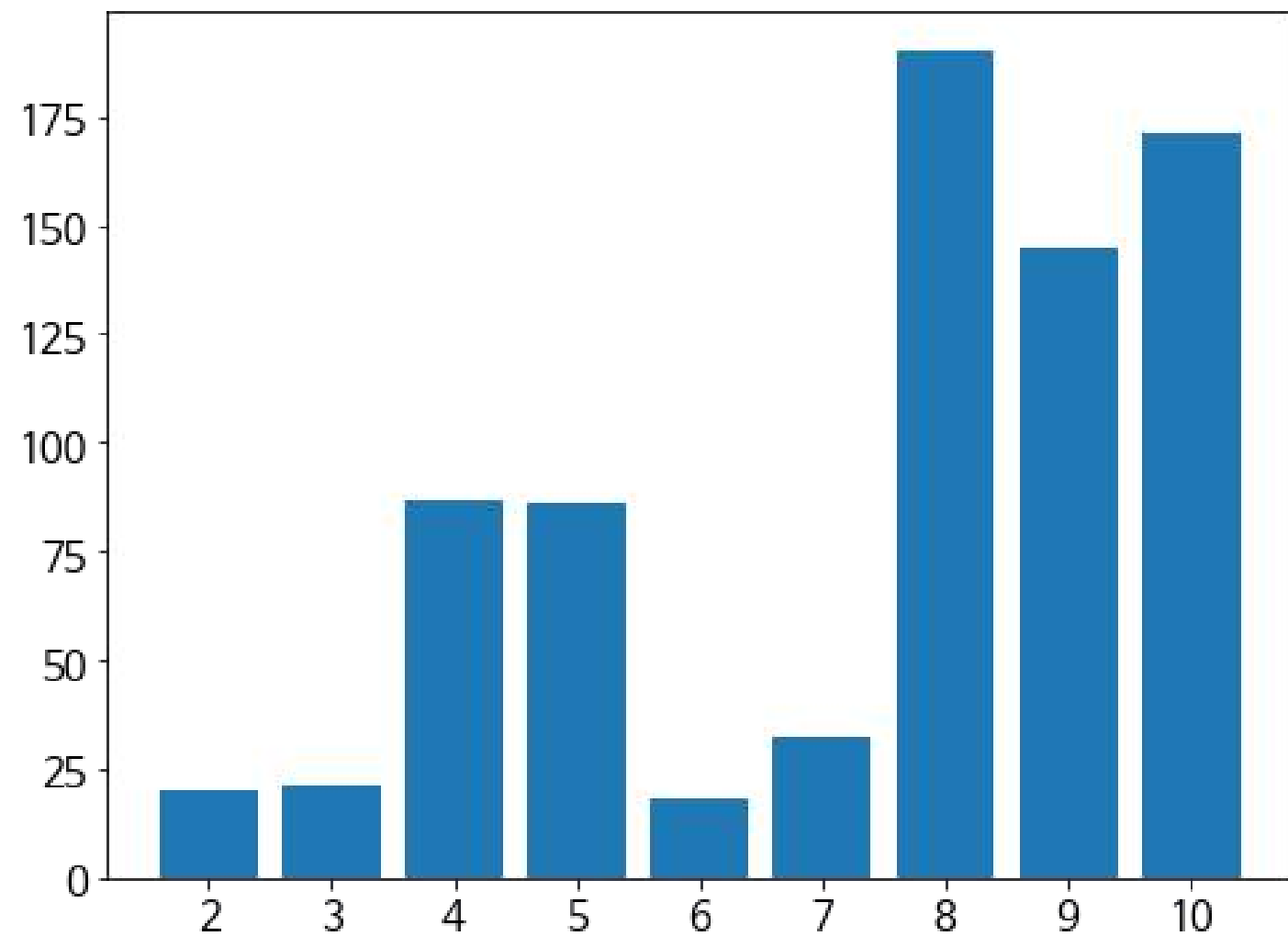


프리미엄 골드석 50% 할인쿠폰	2
3,000포인트 할인 쿠폰(일반)	9
5,000포인트 할인 쿠폰(일반)	10
5천원 멤버십 할인권 (성인)	11
K리그 축덕카드(체크)	11
K리그 축덕카드(신용)	26
청소년/어린이	26
성인	67
5천원 멤버십 할인권 (일반)	87
일반	451
예매권북	770

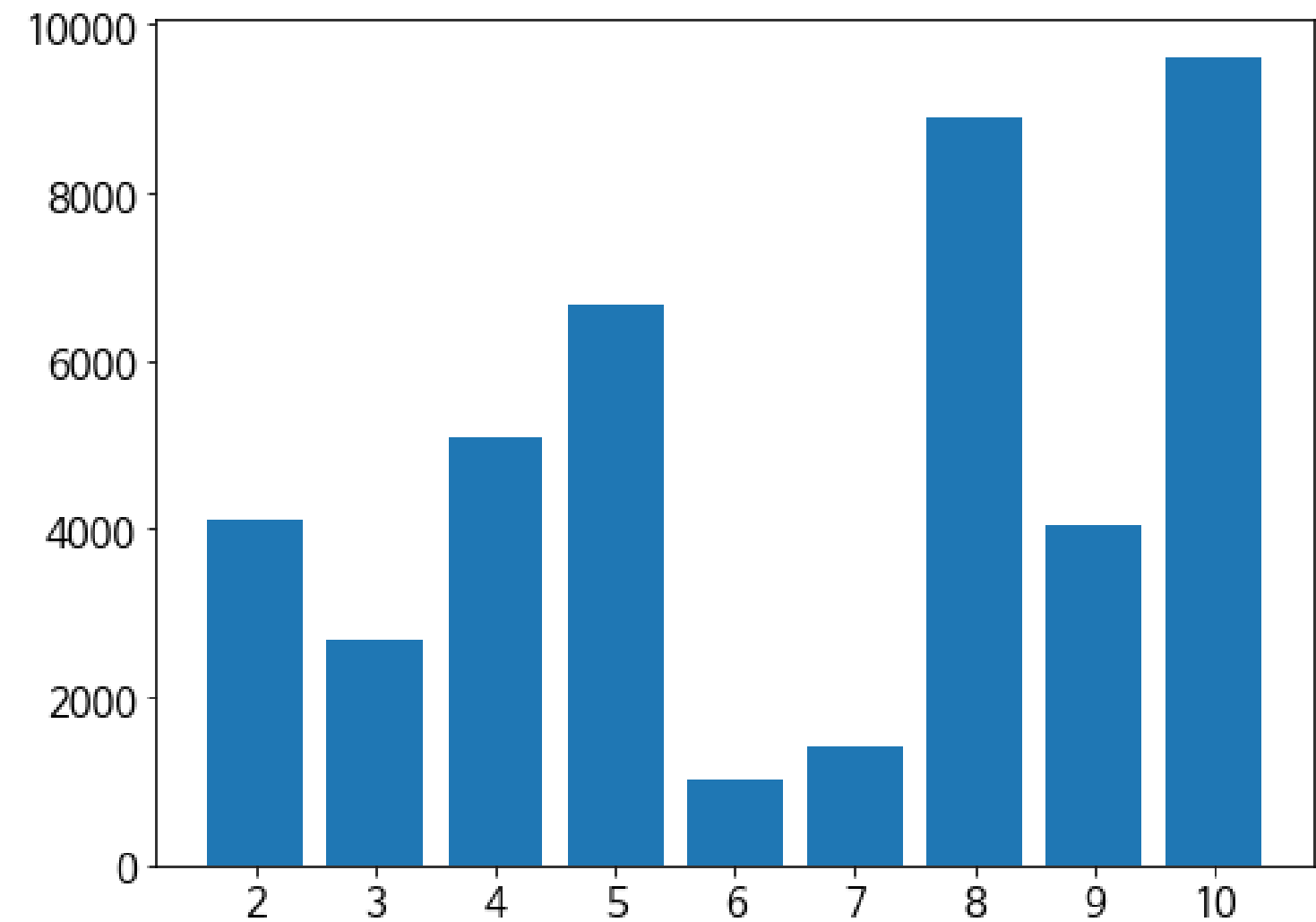
프리미엄 클래식 이용 추이

주로 **시즌 후반**에 사용하며, 전체 관객 수와 고급좌석의 수요가 비례하는 경향이 있다.

월별 예매권북 사용량 (프리미엄 클래식)



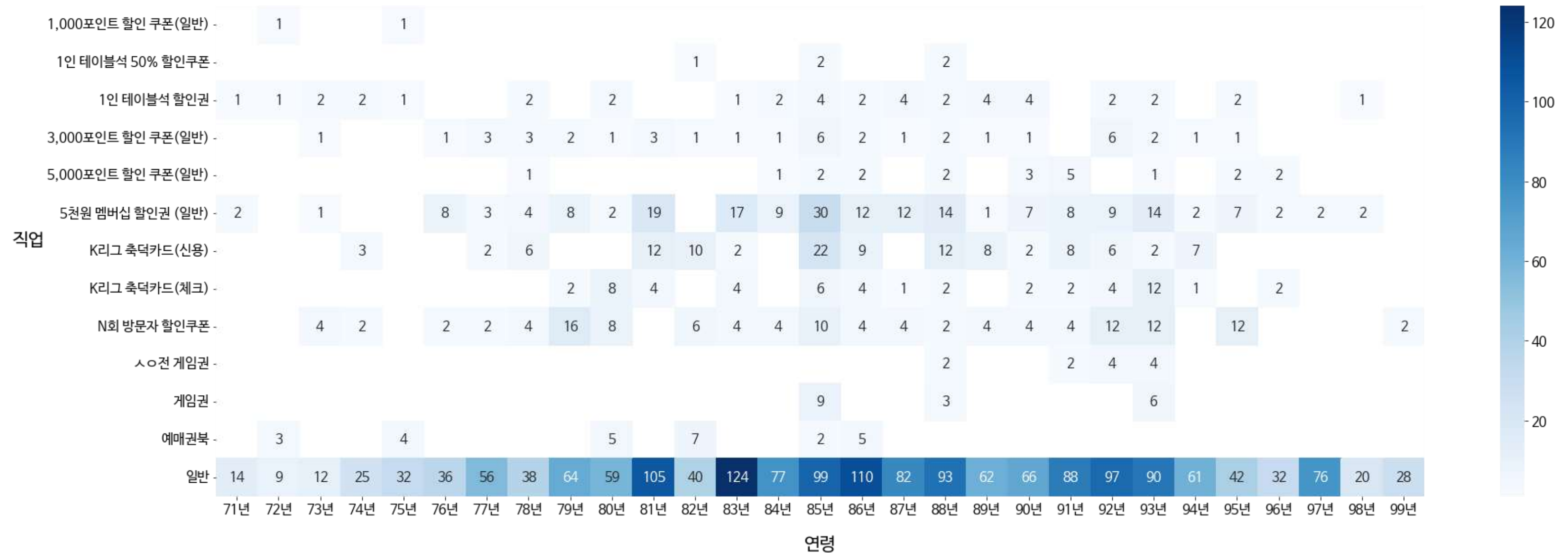
월별 예매량 (전체)



테이블석

프리미엄석에 비해 예매권북 사용이 현저히 낮고 **일반티켓 이용자가 많으며**,
프리미엄 클래식구역과 마찬가지로 **청소년 / 어린이 구매 이력이 적다.**

70~90년생 테이블석 이용자



가설 4 결론

고급좌석 예매 고객 분석

프리미엄 클래식

- 주로 70년~90년대생이고, 미성년자 티켓 구매 이력이 적은 것으로 보아 자녀가 없는 고객층으로 예상된다.
- 예매권북 사용률이 높고, 하반기에 예매권북을 사용하는 경향이 있다.
- 전체 관객 수에 비례하여 고급좌석의 수요도 늘어나는 경향이 있다.
- 하반기에는 예매권북으로 인한 수요가 있으므로 상반기에 고급좌석을 할인 판매하여 예매권북 소지자도 참여할 수 있도록 진행한다.

테이블석

- 예매권북 사용률이 적고 일반티켓 수요가 있는 편이다.
- 프리미엄 클래식과 마찬가지로 자녀가 없는 고객층으로 판단된다.
- 테이블석은 부수적인 이익을 창출하기에 좋은 좌석이므로 예매권북 홍보 및 할인행사 진행 필요

프로모션 제안

- 인천 유나이티드가 패배한 경우 경기관람의 부정적인 감정을 해소하기 위해 일부 고객에게 할인 혜택을 제공한다.
- 월, 화, 수 14시 경 70~90년대 직장인을 대상으로 마케팅을 진행한다.
- 자녀가 없는 경우 고급석 또는 테이블석을 할인을 제공한다.
- 자녀가 있는 경우 미성년 추가 할인이나 가족할인 등을 제공한다.
- 하반기에는 관객이 늘어나고 고급석의 수요가 좋은 편이므로 상반기에 할인 행사를 통해 고객의 경험을 증진한다.
- 테이블석은 예매권북 수요가 적으며 일반 예매가 주를 이루므로 즉흥적 마케팅을 통해 사용량을 늘린다.

프로젝트 회고

고급좌석 예매 예측모델이라는 프로젝트 목표에 도달하지 못하여 아쉽다. 하지만, 데이터를 직접 뜯어보고 들여다 보는 것도 중요하다는 생각으로 열심히 했고 모델링을 하지 못한 것 치고는 만족한다.

생각은 많았으나 생각대로 되지 않는 부분들도 많았고 생각나는 대로 풀어가다 보니 정리하기가 쉽지 않았다. 중간에 한 번 프로젝트를 정리하는 과정이 많은 도움이 되었다. 분석을 할 때 꼭 큰 줄기를 기억하면서 진행해야 한다는 것을 다시 한 번 느꼈다.

또 무턱대고 데이터를 보기 시작했는데, 데이터를 보면 볼 수록 도메인 지식을 잘 알아야 한다는 생각이 많이 들었다. 도메인 지식은 데이터분석의 지도와 같다는 생각이 든다.

파이썬이 아닌 다른 기술들도 사용하고 싶었으나, 파이썬만 사용하기에도 벅찼다. 생각을 이루기 위해서는 꾸준한 공부와 새로운 기술을 연마할 필요성을 느꼈고 몇 가지 책을 구매하였지만 공부하지는 못했다. 시간을 계획적으로 사용하지 못한 것 같아 아쉽다.