

공유오피스 결제 전환율 향상을 위한 데이터 분석 및 전략 제안 보고서

: 시계열 예측과 앙상블 모델을 활용한 최적 전략 도출

0. 목차

1. 분석 배경

- 공유 오피스 동향
- 분석 주제

2. EDA

- 데이터 셋
- 1차 전처리
- 가설 정의
- 현황 분석

4. 예측 및 세부 분석

- 시계열 예측 분석
- 앙상블 모델을 활용한 예측 분석
- 주요 변수 세부 분석

5. 제안

- 신규 요금제 기획
- 무료체험 타당성 검토

5. 제언 및 시사점

- 시사점
- 한계점
- 추가 분석 제안

6. 참고자료

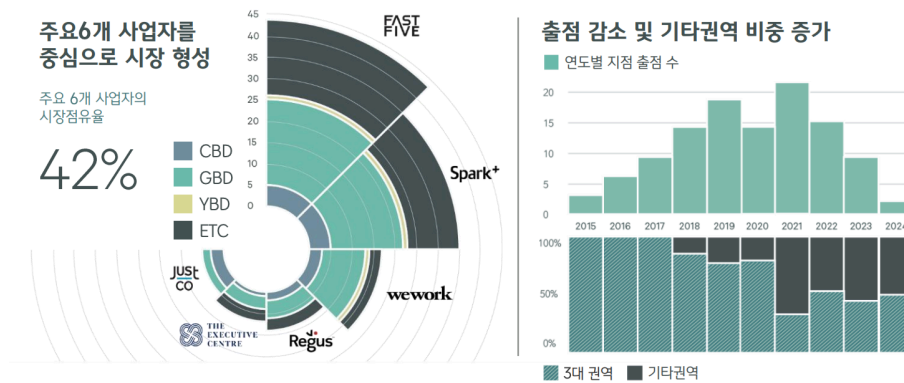
작성자
DA 4기 1팀
박준혁 이은지 장은희 조재윤

1. 분석 배경

1-1. 공유 오피스 동향

코로나19 팬데믹 이후 급격한 성장을 이루었지만 스타트업 및 벤처 업계 위축, 위워크 파산 신청으로 인한 불안감 증대 등으로 국내 공유 오피스 출점이 지속적으로 감소하고 있다. 서울 내 공유 오피스는 패스트파이브, 스파크플러스, 위워크 순으로 시장 점유율을 차지하고 있으며, 시장을 대표하는 기업인 패스트파이브는 매출 우위에도 수익성 개선이 과제로 지적되며 2018년 이후 6년째 적자가 이어지고 있다.

2025년 2월 기준 전년 대비 1~3개월 계약에 대한 선호도가 3% 가까이 증가하는 등의 단기 계약 선호도가 증가함에 따라 패스트파이브는 일주일 단위의 초단기 오피스 임대 서비스를 신설하기도 했다. 불안정한 대내외 정세로 인한 경제적 사회적 불확실성과 변동성이 커지면서 비즈니스 환경의 유연성을 요구하는 기업들의 수요가 증가하고 있다. 이러한 변화는 공유 오피스의 지속적인 발전과 혁신을 촉진할 것으로 예상된다.



2025년 기준 공유 오피스별 점유율

코로나19 팬데믹 당시 재택근무 일상화로 빠르게 늘었던 서울 시내 공유 오피스들이 최근 출점 수가 지속적으로 감소하는 등 부진한 흐름을 보이고 있다. 고금리 시기 스타트업 시장의 투자 위축으로 인한 수요 감소, 1세대 공유 오피스 미국 위워크 파산, 임대인 위주 시장으로의 개편 등이 그 원인으로 꼽힌다.

- 건물주 우위 시장서 공유오피스 매력 급감...올해 출점 2곳뿐 / 2024-09-18

패스트파이브·스파크플러스, 낮은 공실률 유지·수익 안정화에 총력...공용 공간 줄이고 가격 낮춰 승부

글로벌 최대 규모 공유오피스 기업 위워크가 코로나19 여파를 이기지 못하고 지난해 파산을 선언하면서 국내 공유오피스 기업들의 생존에도 관심이 쏠렸다. 파산보호 조치가 북미에 있는 위워크 사무실에만 적용된다고는 해도 국내 시장 역시 영향을 피하기가 어려울 것이라 시각이었다. '운영할수록 손해'인 공유오피스 사업 모델의 한계가 명확해졌다는 평가도 이어졌다.

- 위워크 파산 4개월, 토종 공유오피스는 어떻게 살아남았나

1-2. 분석 주제

본 분석에서는 공유 오피스 출입 데이터를 바탕으로 결제에 영향을 주는 요인들을 파악하고 이를 통해 결제율을 높이는 것을 목표로 한다.

2. EDA

2-1. 데이터 셋 설명

데이터 개요

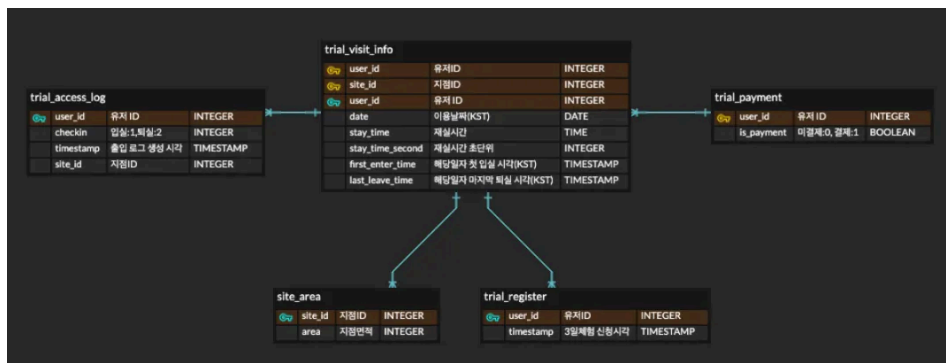
- 본 분석에 활용된 데이터는 공유 오피스 서비스 2년 8개월(2021.05.01 ~ 2023.12.31) 출입 데이터이다.
- 공유 오피스 서비스의 3일 체험 신청, 신청자 일자별 방문 및 출입 기록, 결제 여부, 지점별 면적 정보 등의 정보가 포함되어 있다.
- 지점 별 정보, 결제 후의 출입 데이터, 기업 및 개인 고객 구분 등의 정보의 부재로 인사이트 도출에 한계가 존재한다.

데이터 구성

본 데이터는 4개의 이벤트 데이터와 1개의 지점 정보 데이터프레임으로 구성되어 있다.

테이블명	설명	주요 컬럼
trial_register	3일 체험 신청	유저 id, 3일 체험 신청 일시
trial_visit_info	3일 체험 신청자 일자별 방문 기록	유저 id, 날짜, 지점id, 최초 입실 시각, 최종 퇴실 시각, 체류 시간
trial_access_log	3일 체험 신청자 출입 기록	유저 id, 지점id, timestamp, 출입 방향
trial_payment	3일 체험 신청자 결제 여부	유저 id, 결제 여부
site_area	지점별 면적	지점 id, 지점 면적

ERD (IE 표기법)



Database ERD

2-2. 1차 전처리

테이블 별 중복값, 결측치, 이상치 처리

• trial_visit_info

1. 중복값 : 집계된 중복값은 실 중복값으로 판단하여 마지막 데이터를 남기고 중복값 삭제 진행
2. 결측치 : `first_enter_time`, `last_leave_time` 2개의 컬럼에서 결측치가 확인되었으나, 데이터 대체 방법에 따라 추후 머신러닝 모델에 어떤 영향을 줄지 예측이 어려워 그대로 유지
3. 이상치 : `first_enter_time`이 00:00:00으로 조화되는 경우는 전날 입장하여 다음날 입장한 경우 자동으로 생성되는 데이터로 확인됨.

• trial_register

1. 중복값 : 집계된 중복값은 전산 오류 또는 재신청으로 인해 발생한 중복값으로, 마지막 데이터를 남기고 중복값 삭제 진행
2. 결측치 : 확인되지 않음
3. 이상치 : 서로 다른 날로 체험을 2회 신청한 경우는 `user_uuid`를 기준으로 마지막 데이터를 남기고 중복값 제거

• trial_payment

1. 중복값 : 집계된 중복값은 실 중복값으로 판단하여 마지막 데이터를 남기고 중복값 삭제 진행
2. 결측치 : 확인되지 않음

• trial_access_log

1. 중복값 : 집계된 중복값은 실 중복값으로 판단하여 마지막 데이터를 남기고 중복값 삭제 진행
2. 결측치 : 확인되지 않음

테이블 병합

- user_uuid 기준으로 trial_visit_info, trial_register, trial_payment 를 LEFT JOIN 방식으로 병합
- 컬럼 추가 : trial_access_log데이터는 date 와 user_uuid 로 그룹화 후, 각 그룹의 방문 횟수를 계산하여 count 컬럼 추가
- 이상치 : 총 방문 일수가 무료 체험 기간(3일)을 초과한 4일 이상인 데이터 제거
- 결측치 : count 컬럼에서 결측치가 존재하지만, 추후 머신러닝 전처리 과정에서 처리 방안 고려

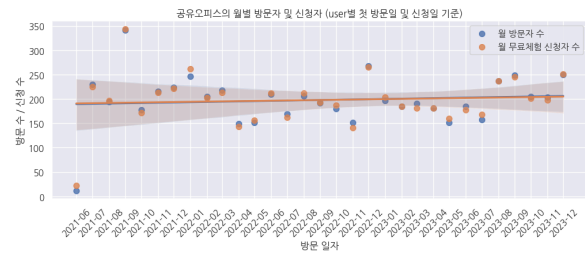
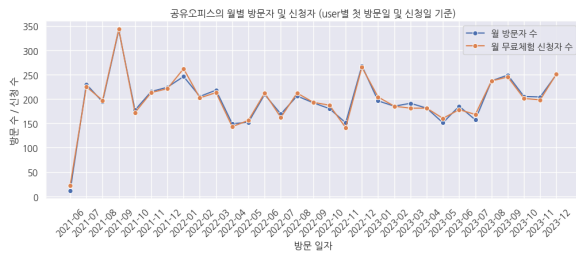
2-3. 가설 정의

1. 체류 시간에 따라 결제율에 차이가 있을 것이다.
2. 방문 일 수에 따라 결제율에 차이가 있을 것이다.
3. 방문 지점의 면적에 따라 결제율에 차이가 있을 것이다.

2-4. 현황 분석

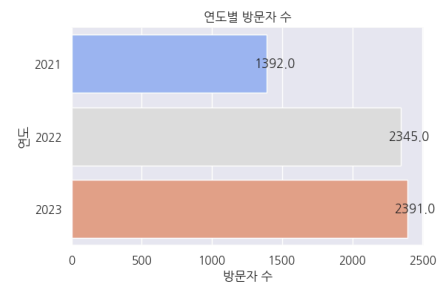
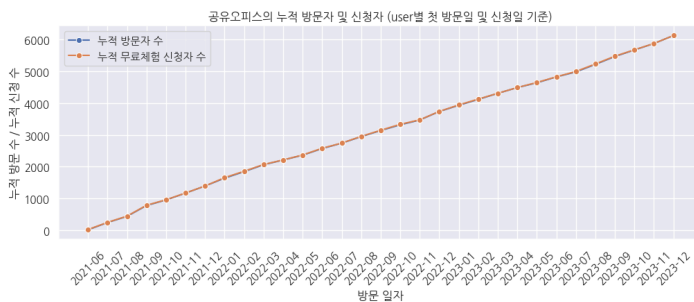
서비스 이용자 현황

1. 월별 방문자 및 신청자 현황 (user별 첫 방문일 기준)



2021년 6월부터 2023년 12월의 월별 방문자 및 신청자 수 현황을 살펴봤을 때, 2021년 8월을 제외한 **방문자 및 신청자 수는 150명에서 260명 박스권 내에서 유지**되는 것으로 관찰된다.

2. 누적 방문자/신청자 및 연도별 방문자 수



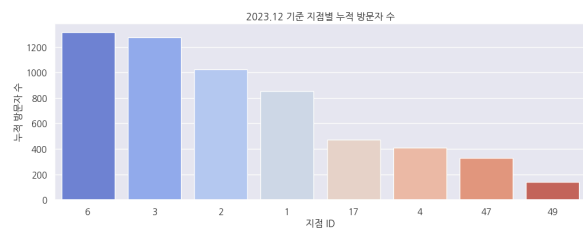
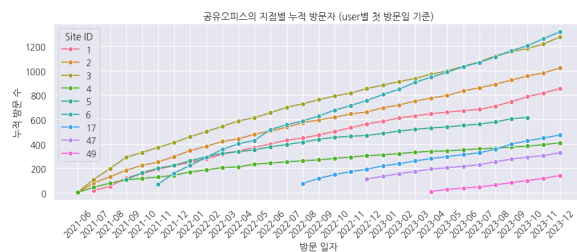
2023년 12월 기준 공유 오피스 전 지점의 누적 방문자 및 신청자는 약 6100명이며, 연도별 방문자 수의 경우 2021년 대비 2022년도 방문자 수 증감율은 68.4%, 2022년도 대비 2023년도 방문자 수 증감율은 1.9%로 방문자 수의 증가가 전년 대비 둔화되었음을 알 수 있다.

3. 지점별 오픈 일자

지점 ID	면적	오픈 일자
1	50	2021-07-02
2	100	2021-06-30
3	150	2021-06-30
4	100	2021-06-30
5	150	2021-09-03
6	150	2021-11-15
17	50	2022-08-08
47	50	2022-12-09
49	50	2023-04-22

지점 1, 2, 3, 4는 동일 시점에 오픈했으며, 그 후 5, 6, 17, 47, 49 순으로 지점을 오픈하였다. 2022년도 이후에는 모두 면적을 50인 지점만 오픈하였다. (오픈일자: 첫 방문일 기준)

4. 지점별 누적 방문자 현황



지점별 누적 방문자 현황을 살펴보면 각 지점별로 방문자 증가율이 상이한 것을 확인할 수 있다. **2023.12 기준 누적 방문자가 가장 많은 지점은 지점 6**이며, 3, 2, 1 순으로 누적 방문자 수가 많다.

5. 면적 별 방문 유저 수/결제를 비교

5-1) 면적 별 보유 지점 수

면적(평)	보유 지점 수(개)
50	4
100	3
150	2

5-2) 면적 별 평균 방문 유저 수

면적(평)	평균 방문 유저 수(명)
50	462
100	809
150	1,104

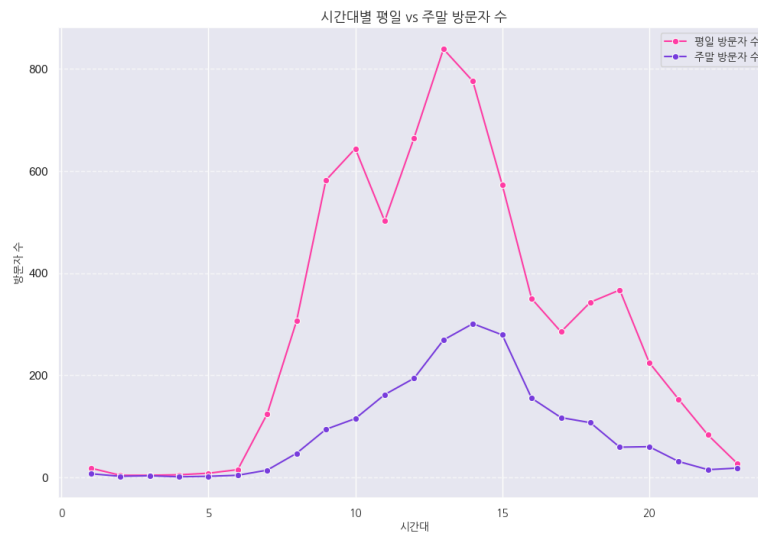
5-3) 면적과 결제 여부간의 통계적 검정 진행

면적에 따른 결제 여부가 유의미한 분포 차이가 있는지 확인하는 독립성 검정 수행

가설	귀무가설(H0): 지점의 면적에 따른 구독 전환율 차이가 없다. 대립가설(H1): 지점의 면적에 따른 구독 전환율 차이가 있다.
카이제곱 통계량	6.9730
p-value	0.0306
검정 결과	p값이 0.05보다 작으므로 귀무가설 기각. 두 그룹 간의 결제 전환율에는 유의미한 차이가 있음

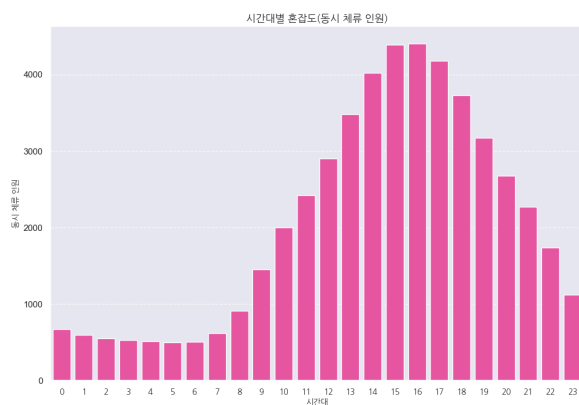
50평 대의 지점은 4개, 100평 대의 지점은 3개, 150평 대의 지점은 2개로 50평 대의 지점을 가장 많이 보유하고 있다. 면적 별 평균 방문 유저 수를 살펴보면 **150평 대 지점의 평균 방문 유저 수는 약 1,100명으로, 50평 대 지점보다 평균 방문 유저가 약 2.3배 높은 것으로 확인**된다. 통계적 검정을 통해 면적에 따라 결제 여부의 차이가 유의미하며, 면적이 넓은 지점일수록 결제 전환율도 높은 경향이 있음을 확인하였다.

6. 시간대별 평일/주말 방문자 수 비교



평일이 주말보다 평균 방문자 수가 더 높으며, 시간에 따른 방문자 수는 10시~15시가 가장 많고, 15시 이후로는 점차 감소하는 경향을 보인다.

7. 시간대별 혼잡도(체류 인원) 비교

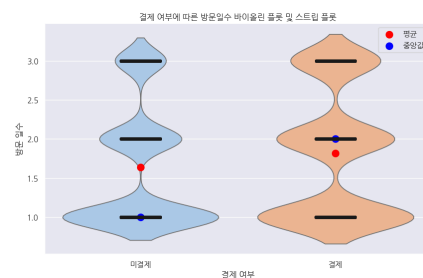


시간대별 동시 체류 인원은 15~16시가 가장 많으며, 이후 점차 증가하다가 다시 감소하는 경향을 보인다.

8. 결제 여부 별 무료 체험 기간 내 공유 오피스 방문 일수

결제 여부	중앙값	평균
결제	2.0	1.818571
미결제	1.0	1.637490

결제 여부에 따른 공유 오피스 방문 일수의 평균을 살펴봤을 때 결제한 경우는 1.81일, 미결제한 경우는 1.63일로, 방문 일수의 평균은 1~2일에 주로 분포해있으며, 중앙값과 평균을 모두 고려했을 때, 결제한 경우의 방문 일수가 더 긴 것으로 확인되었다.



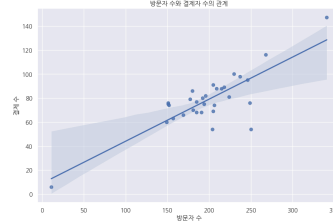
3. 예측 및 세부 분석

3-1. 시계열 예측 분석

분석 개요

- 목적

방문 수와 결제 수를 기준으로 산점도와 회귀선을 살펴본 결과, 방문자 수가 증가할수록 결제 수가 비례하여 증가하는 뚜렷한 선형 관계 (상관계수: 0.823682)를 확인하였다. 이러한 상관관계를 바탕으로, 공유 오피스의 핵심 성과 지표인 결제 수를 향상시키기 위해서는 방문자 수 증대가 중요한 요인임을 알 수 있다. 따라서 시계열 예측 모델을 활용하여 방문자 추이를 살펴보고, 향후 180일 기간 동안의 방문자 수 변화를 예측하고자 한다.



방문자 수와 결제자 수의 관계

- 사용 모델: Prophet

- Facebook이 개발한 시계열 예측 모델로, 계절성과 추세를 모델링하고 추세를 예측에 장점이 존재한다. 공유 오피스 방문자 데이터는 요일별, 월별 계절성과 특정 이벤트(휴일, 프로모션 등)에 따른 변동성을 보이는 특성이 있다. Prophet 모델은 이러한 다중 계절성과 추세, 휴일 효과를 효과적으로 모델링할 수 있으며, 이상치에 강건하고 누락된 데이터가 있는 시계열에도 적용 가능하다.

- 모델 평가를 위한 지표 선정

항목	설명
주요 평가 지표	SMAPE - MAPE의 대칭 버전으로, 예측값과 실제값의 차이를 대칭적으로 계산. - 대칭적으로 계산되어 예측값과 실제값의 차이를 공평하게 반영 - 상대적 오차를 공평하게 평가할 수 있어, 예측의 정확성을 더 균형 있게 파악
보조 평가 지표	MAPE - 예측값이 실제값에서 얼마나 퍼센트로 차이가 나는지를 평균적으로 나타냄

예측 기간 설정

- 1,2,3,4,5,6 지점: 중장기 계획 수립을 위해 180일 설정
- 17,45,47 지점: 데이터 보유 기간이 상대적으로 짧아 90일 설정

모델 평가 계획

- `cross_validation` 함수를 사용해서 교차 검증
 - initial (모델을 학습시키는 데 사용되는 데이터의 초기 기간) : 80%
 - horizon (initial 이후의 예측을 수행하고 평가할 기간) : 20%
 - period (교차 검증의 간격) horizon의 절반 기간으로 설정

1차 전처리 및 하이퍼파라미터 설정

- 전처리
 1. `first_enter_time`, `last_leave_time` 결측치 삭제
 2. 지점별 방문일자별 방문자 수 집계

3. date는 `ds`, 결과 변수인 방문자 수는 `y`로 컬럼 수정
- 하이퍼파라미터 설정
 1. `changepoint_prior_scale` (변화점에 대한 유연성 조절): 기본값인 `0.05` 설정
 2. 사용자 정의 계절성의 주기
 - 주간계절성 추가
 - `period` (계절성의 주기): 주간 주기를 반영하기 위해 한달 평균인 `7` 일을 지정
 - `Fourier Order` (푸리에 항의 수): 적당한 복잡성을 보도록 `5`로 지정
 - 월간계절성 추가
 - `period` (계절성의 주기): 월간 주기를 반영하기 위해 한달 평균인 `30.5` 일을 지정
 - `Fourier Order` (푸리에 항의 수): 적당한 복잡성을 보도록 `5`로 지정

2차 하이퍼파라미터 튜닝

1. `changepoint_prior_scale`: 변화점에 대해 더 민감하게 반응하도록 `0.1`로 조정
2. `weekly_seasonality` (주간 패턴 모델링): `True`로 설정하여 요일별로 반복되는 패턴을 예측에 반영
3. 사용자 정의 계절성의 주기 : 1차와 상동

모델 성능 및 결과

• 1차 성능 비교

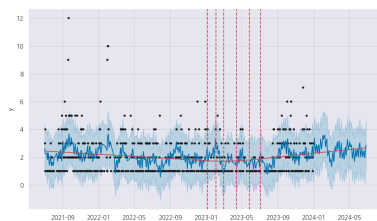
지점 / 평가지표	SMAPE	MAPE
지점 1	0.590	0.464
지점 2	0.490	0.532
지점 3	0.601	0.730
지점 4	0.429	0.454
지점 5	0.451	0.523
지점 6	0.551	0.602
지점 17	0.464	0.659
지점 47	1.153	0.858
지점 49	0.461	0.498

• 2차 성능 비교

지점 / 평가지표	SMAPE	MAPE
지점 1	0.587	0.463
지점 2	0.477	0.555
지점 3	0.575	0.743
지점 4	0.433	0.513
지점 5	0.469	0.611
지점 6	0.566	0.581
지점 17	0.494	0.787
지점 47	0.936	0.733
지점 49	0.464	0.506

모델 해석 및 결론

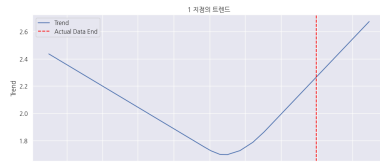
- 평가 지표가 상대적으로 낮은 지점은 해석에서 제외하고 `지점 1, 2, 4, 49` 4개 지점에 대해 해석 진행.



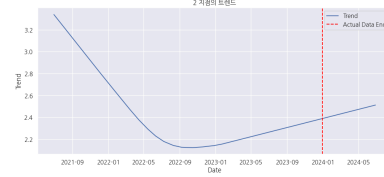
지점 1의 예측 시각화

- 지점 1의 예측 시각화를 샘플로, Prophet 모델에서 시사하는 바는 아래와 같다.
 - 붉은선은 트렌드라인으로, 시간이 지남에 따라 방문자 수가 어떻게 변화할지 보여준다.
 - 검은점과 파란선은 각각 실제데이터와 예측데이터이다.
 - 빨간 점선은 시간에 따른 추세의 변화점, 특정 시점에 트렌드가 변화하는 것을 의미함

• 트렌드 (Trend)



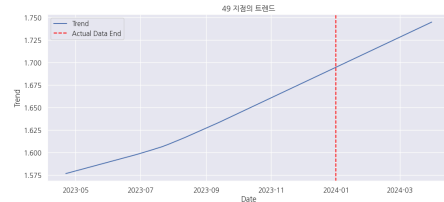
지점 1 트렌드



지점 2 트렌드



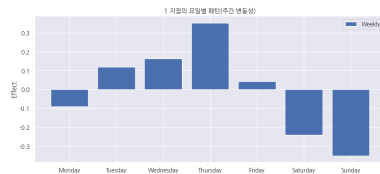
지점 4 트렌드



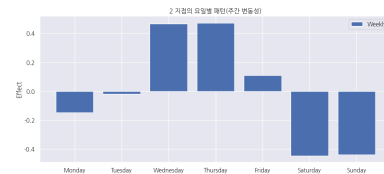
지점 49 트렌드

- 지점 1과 지점 2는 방문자가 각각 23년초, 22년도 중반까지 감소하다가 이후 증가 혹은 회복하는 경향을 보인다. 이후 180일 동안 방문자 수가 증가할 것으로 예측된다.
- 지점 4는 오픈 이후 방문자가 급격하게 감소한 후 방문자가 2명 이내로 계속 유지되는 경향을 보인다.
- 지점 49는 오픈 이후 방문자수가 꾸준히 상승세를 보이며, 이후 90일 동안도 방문자수가 계속 증가할 가능성이 높다.

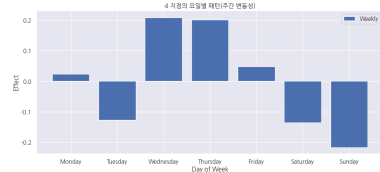
• 요일 별 패턴 / 주간 변동성 (Day_of_week)



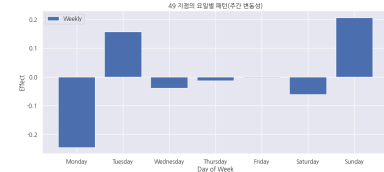
지점 1 요일 별 패턴 (주간 변동성)



지점 2 요일 별 패턴 (주간 변동성)



지점 4 요일 별 패턴 (주간 변동성)



지점 49 요일 별 패턴 (주간 변동성)

- 지점 1, 2, 4는 수목에 방문자가 집중되어 있으며, 주말은 상대적으로 낮은 방문율을 보인다. 해당 지점은 기업고객 또는 주간의 업무 또는 공부를 진행하는 방문자가 많은 것으로 예측된다.
- 지점 49의 경우 타 지점에 비해 화요일과 일요일의 방문자가 높게 나타나고 있다. 해당 지점의 주 방문자는 기업고객보다 프리랜서나 일반 고객일 것으로 예측된다.

• Prophet 모델을 통한 시계열 예측 분석 결론

방문자수가 증가 추세에 있는 1, 2, 49 지점은 예측과 같이 방문자가 향후에도 증가할 수 있도록 상품 개선 등의 관리가 필요하며, 방문자수가 보합세에 있는 4 지점은 방문자가 증가하지 않는 요인을 모색하고 개선하는 전략이 필요하다.

3-2. 앙상블 모델을 활용한 예측 분석

분석 개요

- 사용 모델: RandomForest, XGBoost, LightGBM, RUSBoost, CatBoost, HistGradientBoostingClassifier

항목	설명
목적	체험 고객의 구매 전환 예측 및 핵심 영향 변수 도출
주요 지표	AUC-PR (불균형 데이터에 대한 정밀도-재현율 균형 평가)
보조 지표	Accuracy
전처리 과정	사이클링 인코딩 · 원핫인코딩 · 오버 샘플링 적용, 2단계 파생 변수 생성

Feature Selection

모델 학습에 사용될 변수들을 선정하기 위해 **데이터 병합 후 피처 선택** 과정을 진행했다. 이 과정에서는 데이터의 특성과 모델 학습에 유용한 특성을 추출하는 데 초점을 맞추었다. 각 피처의 역할을 명확히 이해하고, 모델링에 불필요하거나 중복된 정보를 제거하며, 의미 있는 변수만을 선택하였다.

- 원본 데이터의 피처

원본 피처	설명
site_id	지점ID
date	방문 일자
stay_time	머무른 시간(timestamp 단위)
stay_time_second	머무른 시간(second 단위)
first_enter_time	입장 시간
last_leave_time	퇴장 시간
user_uuid	사용자 ID
is_payment	결제 여부
count	하루 방문 횟수
total_visit_count	3일 체험 동안 방문한 횟수 (1~3)
trial_date	체험 등록일

- 선택된 피처 및 처리 사항

선택 피처	변경/처리 사항	설명
site_id	유지	site_id 는 각 지점을 구분하는 카테고리형 변수로, 모델에 필요한 중요한 정보를 제공하므로 원핫인코딩 후 유지함.
date	유지	날짜 정보는 방문 패턴이나 특정 날짜에 따른 차이를 분석하기 위해 연도, 월, 일로 분리하여 처리함.
stay_time	삭제	stay_time 은 stay_time_second 와 동일한 의미를 가지며, 초 단위로 이미 정보가 제공되므로 중복을 피하기 위해 삭제함.
stay_time_second	유지	체류 시간을 초 단위로 나타내는 변수로, 모델 학습에 유용하므로 그대로 유지함.
first_enter_time	유지	입장 시간은 결제 여부에 영향을 미칠 수 있으므로, 시간 단위로 분리하여 처리함.
last_leave_time	유지	퇴장 시간도 결제 여부에 영향을 줄 수 있으므로, 시간 단위로 분리하여 모델에 유용하게 반영됨.
user_uuid	삭제	user_uuid 는 각 사용자를 구분하는 고유 ID로, 모델링의 설명복잡도가 올라갈수있어 삭제함.

선택 피처	변경/처리 사항	설명
is_payment	유지	결제 여부는 종속 변수로 설정되어, 모델이 결제 여부를 예측할 수 있도록 핵심 변수로 유지함.
count	유지	결측값이 있을 경우 중앙값으로 대체하여 결측값 문제를 해결하고, 하루 방문 횟수는 유용한 특성으로 유지함.
total_visit_count	유지	3일 체험 동안의 방문 횟수는 특정 기간 동안의 사용자 행동을 반영하는 중요한 변수로, 분석에 활용함.
trial_date	유지	등록일과의 관계를 활용할수있을것같아 유지함.

• 피처 선택의 이유

각 피처에 대해 선택 및 처리 과정은 모델링 목표에 맞게 중요한 정보는 최대한 보존하기 위해 변수를 선택하였다. stay_time 과 stay_time_second 는 동일한 정보를 제공하므로 stay_time 을 삭제하고, stay_time_second 는 초 단위로 구체적인 체류 시간을 제공하므로 유지했다.

first_enter_time 과 last_leave_time 은 방문 시각이 결제 여부에 영향을 줄 가능성이 높기 때문에, 시각적 특성을 반영할 수 있도록 각각 시간 단위로 분리하여 유지했다. user_uuid 는 고유 사용자 ID로, 모델에 큰 영향을 미치지 않으며, 모델의 설명력을 높이기 위해 삭제했다.

is_payment 는 결제 여부를 나타내는 종속 변수로, 결제 여부를 예측하는 데 사용되므로 유지했다. 또한, count 와 total_visit_count 는 각각 하루 방문 횟수와 3일 동안의 유니크한 방문 횟수를 나타내며, 사용자의 방문 패턴을 파악하는 데 중요한 특성으로 유지했다.

1차 전처리

• 결측치 처리

- first_enter_time , last_leave_time 각 520개의 결측치 삭제
- count 변수에서 2436개의 결측치를 다 삭제하면 데이터의 개수가 많이 줄어들어 중앙값으로 대체

• 피처엔지니어링

- 원핫인코딩
 - site_id를 원핫인코딩
- 파생 변수

구분	변수	설명
추가 변수	trial_diff	체험 등록일과 첫 방문일 차이 (1~3일) → date - trial_date
	weektype	date의 주말(1) vs 평일(0) 구분
	total_visit_count	사용자별 공유 오피스 방문 횟수 (중복 일자 제거)
	year , month	date 에서 시계열을 쪼개서 생성
	hour	등록일에서 시계열로 쪼개서 생성
삭제 변수	date	원본 데이터 제거
	first_enter_time	
	trial_date	
	last_leave_time	

2차 전처리

추가 변수	설명
hour_sin/hour_cos	시간 데이터 주기성 반영 (사이클링 인코딩)
unique_site_count	사용자가 방문한 공유 오피스 종류 수
area_pyeong	공유 오피스 면적 정보 추가

주요 모델별 특징

모델	장점	단점
CatBoost	카테고리형 데이터 처리 우수	긴 훈련 시간, 하이퍼파라미터 튜닝 복잡
RandomForest	과적합 방지, 안정적인 성능	느린 예측 속도, 높은 메모리 사용
XGBoost	병렬처리로 고속학습, 과적합 방지	높은 메모리 사용, 하이퍼파라미터 튜닝 복잡

모델 성능 비교 (2차 전처리 기준)

- 최종 6개 모델 비교

	RandomForest	XGBoost	LightGBM	CatBoost	RUSBoost	HistGradientBoost
1차 AUC-PR	0.54	0.46	0.54	0.47	0.48	0.47
1차 Accuracy	0.62	0.62	0.63	0.63	0.56	0.59
2차 AUC-PR	0.55	0.54	0.54	0.61	0.53	0.59
2차 Accuracy	0.62	0.62	0.62	0.65	0.59	0.64

- 모델 순위

모델	AUC-PR	Accuracy	순위
CatBoost	0.61	0.65	1위
RandomForest	0.55	0.62	2위
XGBoost	0.54	0.62	3위

 결론: CatBoost가 AUC-PR **0.61**로 가장 우수한 성능을 보임

베스트 모델 2개의 속성중요도 TOP 10

Catboost	RandomForest
day_sin	day_sin
month_sin	stay_time_second
stay_time_second	first_enter_hour_sin
first_enter_hour_sin	month_sin
last_leave_hour_sin	last_leave_hour_sin
total_visit_count	weekday_sin
weekday_sin	count
count	total_visit_count
trial_first_diff	date_year
date_year	trail_first_diff

두 모델의 속성 교집합인 date_month, date_day, stay_time_second, weekday, first_enter_hour 를 세부 분석에 사용하였다.

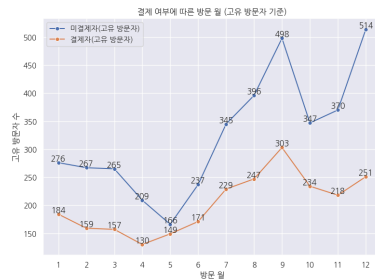
3-3. 주요 변수 세부 분석

분석 개요

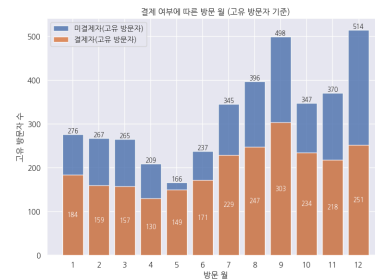
- 목적 : ML모델에서 속성 중요도가 제일 높은 5가지 위주의 세부 분석 진행하여 인사이트 도출 및 전략 제시

date_month : 결제 여부에 따른 date_month 파악

- 결제 여부에 따른 방문 월



결제 여부에 따른 방문 월 (고유 방문자 기준)



결제 여부에 따른 방문 월 (고유 방문자 기준)



월별로 결제 패턴이 다르게 나타나는 요인에 대한 가설

- 정부 또는 투자 기관의 예산 편성 및 집행 일정에 따른 영향

일반적으로 정부나 투자 기관은 연간 예산을 계획하고 집행하는데, 이 과정에서 자금 배분이 하반기에 집중될 수 있다. 많은 국가나 기관에서 회계연도가 상반기로 설정되어 있어 상반기에는 예산 검토 및 승인 절차가 진행되고, 하반기에 실제 집행이 이루어지는 경우가 많다. 스타트업 지원금도 이런 일정에 맞춰 지급되면, 기업들이 하반기에 자금 여유를 갖게 되고, 하반기가 될수록 유료 서비스를 체험하거나 결제할 가능성이 높아질 수 있다.

- 상반기 : 자금 확보 전 → 방문자 및 결제 전환율 하락
- 하반기 : 지원금 수령 후 → 자금 여유 생겨 방문자 및 결제 전환율 상승

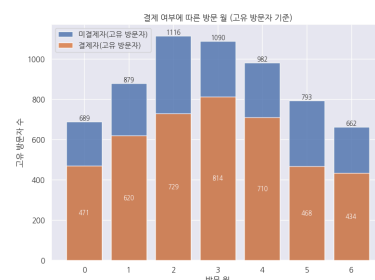
- 일반적으로 기업은 상반기에 오피스 이전이나 큰 환경 변화를 꺼리는 경향이 있다. 이는 연초에 사업 계획을 세우고 실행에 집중하며, 안정적인 운영을 우선하기 때문이다.

weekday : 결제 여부에 따른 weekday 파악

- 결제 여부에 따른 (고유) 방문 요일



결제 여부에 따른 방문 요일 (고유 방문자 기준)



결제 여부에 따른 방문 요일 (고유 방문자 기준)

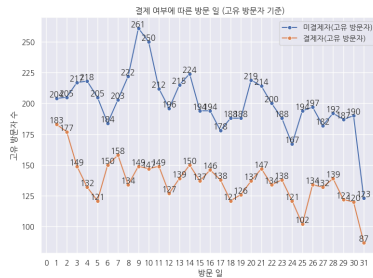


요일별로 결제 패턴이 다르게 나타나는 요인에 대한 가설

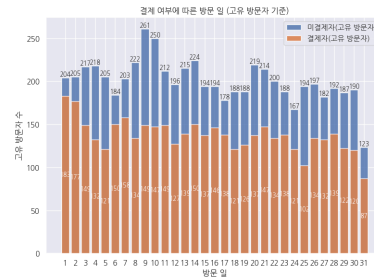
- **월요일 : 회원 비중 높음, 결제 전환율 증가**
 - 월요일 방문자는 **주로 회원**으로 구성되어 있다고 가정한다. 대부분의 회사가 월요일부터 출근하기 때문에, 이 날 방문하는 사용자들은 회원 중심이라고 가정한다. 이러한 사용자들이 결제 전환율이 **높은 경향**을 보인다.
- **화~금 : 방문자 증가**
 - 화요일부터 금요일까지는 방문자 수가 **증가**하는 경향을 보이며, 추가 방문자들은 **프리랜서나 일반 고객**일 가능성이 크다.
- **주말 : 월요일에 비해 방문자 수 증가, 결제 전환율 감소**
 - 주말에는 월요일에 비해 방문자 수가 증가하는 경향이 있지만, 전체적으로 **결제 전환율이 감소**하는 경향을 보인다.

date_day

- 결제 여부에 따른 (고유) 방문 일자



결제 여부에 따른 방문 일 (고유 방문자 기준)



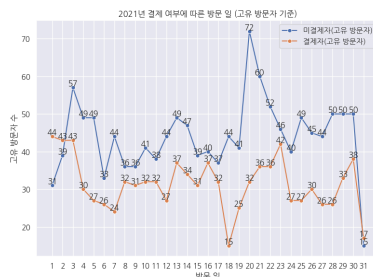
결제 여부에 따른 방문 일 (고유 방문자 기준)



인사이트

- 전체 기간의 일자별 패턴을 살펴볼 때, 요인을 판단하는데 어려움이 있다.
- 따라서 1차로 연도별로 나누어 패턴을 살펴본 후 2차로 월별/지점별로 구분하여 요인을 탐색하였다.

- 2021년 : 20일 미결제자 급증



2021년 결제 여부에 따른 방문 일 (고유 방문자 기준)

- 7월 20일 미결제 증가

해당 지점	2
가설	기업 고객 단체 방문으로 예측

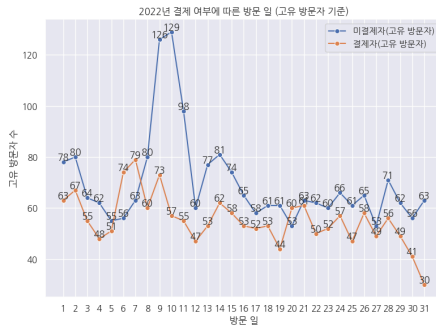
- 9월 20일 결제/미결제 동시 증가

해당 지점	1, 3, 4, 5
가설	1. 추석 (9/20 ~ 9/22) 2. 코로나19로 사회적 거리두기에 의한 활동 제한 예상

- 10월 20일 미결제 증가 (수)

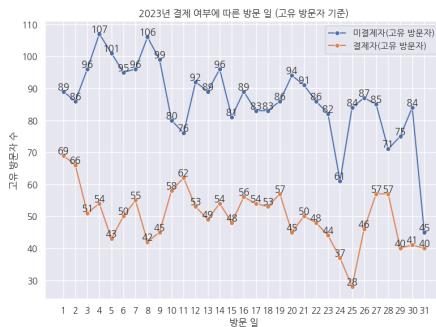
해당 지점	1
가설	기업 고객 단체 방문으로 예측

- 2022년 : 9~10일 미결제자 급증



2022년 결제 여부 따른 방문 일 (고유 방문자 기준)

- 2023년 : 24~25일 결제/미결제자 급감



2023년 결제 여부 따른 방문 일 (고유 방문자 기준)

- 4월 9~12 미결제 증가

해당 지점	3, 6
가설	기업 고객 단체 방문으로 예측

- 6월 8~10 미결제 증가

해당 지점	6
가설	기업 고객 단체 방문으로 예측

- 8월 10일 미결제 증가

해당 지점	2, 3, 5, 17
가설	지점 17 오픈 영향 (8/8 오픈)

- 12월 9~10 결제/미결제 동시 증가

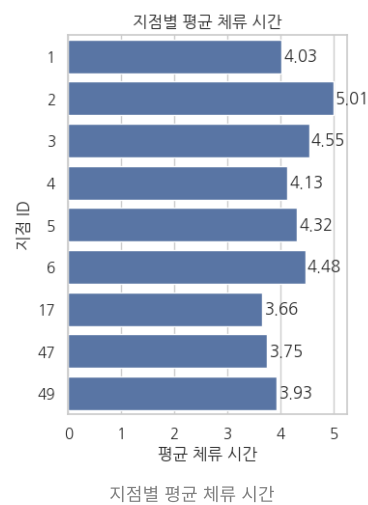
해당 지점	47
가설	지점 오픈 영향 (12/9 오픈)

경향	- 월별 뚜렷한 경향 X - 전체적으로 24~25일에는 방문자 저조
----	--

stay_time_second

- 지점별 하루 평균 체류 시간

지점ID	평균 체류시간
1	4.026237
2	5.005946
3	4.545646
4	4.134654
5	4.318335
6	4.476938
17	3.660766
47	3.749892
49	3.928464





지점별로 평균 체류 시간이 다르게 나타나는 요인에 대한 가설

- **지점별 평균 체류 시간 차이**
 - 지점별 평균 체류 시간이 다르게 나타나며, 이는 지점의 규모, 구성, 이용 고객층 등의 영향을 받을 수 있다.
- **지점 크기와 체류 시간 관계**
 - 면적이 클수록 쾌적한 환경을 제공하며, 이용자들이 더 오래 머무르는 경향이 있다.
- **기업 고객 비중과 체류 시간**
 - 면적이 큰 지점은 오피스 전용 공간이 많아 기업 고객이 주로 이용한다고 가정하였고, 이들은 일반 고객보다 더 긴 시간 체류하는 경향이 있다.

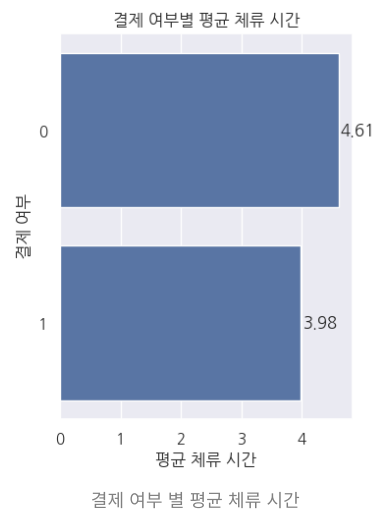
• 결제 여부에 따른 하루 평균 체류 시간

결제 여부	평균 체류 시간
미결제(0)	4.610773
결제(1)	3.983125



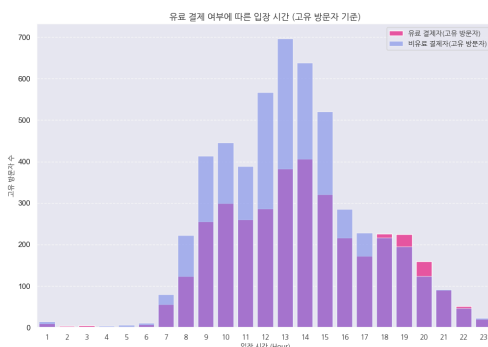
결제 여부별로 평균 체류 시간이 다르게 나타나는 요인에 대한 가설

- **미결제 이용자** : 체험 기간만 충분히 이용한 후 결제 없이 이탈하는 경향이 있다.
- **결제 이용자** : 보통 이미 결제할 의향이 있고, 서비스를 경험한 후 그 가치만 확인하면 즉시 결정을 내리기 때문에 비교적 짧은 체류 시간을 보인다.



first_enter_hour : 결제 여부에 따른 first_enter_hour 파악

• 결제 여부에 따른 (고유) 방문자 입장 시간(0시 입장 제외)



결제 여부에 따른 입장 시간(고유 방문자 기준)



결제 여부에 따른 입장 시간(고유 방문자 기준)

(전날 입장하여 다음 날 퇴장한 경우, 데이터가 0으로 기록되어 분석에서 제외함)



출근 시간대 (07:00~10:00)

- 방문자 수 증가 (미결제자가 조금 더 많음)

점심 시간대 (12:00~14:00)

- 방문자 수는 증가하나, 미결제 이용자수도 증가

오후 시간대 (15:00~18:00)

- 방문자 수 전체적으로 감소

저녁, 야간 시간대(19:00~)

- 방문자 수 급감, 결제 전환율이 상대적으로 높음

→ 1시 이전에는 기업 고객 비중이 높으며, 1시 이후부터는 프리랜서 및 일반 고객의 비중이 증가할 가능성이 크다.

4. 제안

4-1. 신규 요금제 기획

현재 이용자들은 방문 시간, 체류 시간, 방문 요일 등의 패턴이 서로 다르게 나타나고 있어, 일률적인 요금제보다는 개인 맞춤형 요금제의 필요성이 증가하는 추세다. 이를 반영하여 다양한 고객층을 유치하고 이용률을 극대화 할 수 있는 새로운 요금제를 제안한다.

| 이용 패턴별 요금제

- 월별/ 요일별/ 일별 수요에 따라 차별화 된 가격 및 프로모션 적용
- 특정 요일 방문 시 혜택 제공

| 주말 평일 전용 요금제

- 주말 요금제 : 주말에만 이용할 수 있도록 구성(재택 근무자, 주부, 프리랜서 등)
- 평일 요금제 : 평일 기업 근무자들을 위한 요금제

| 이용 횟수 기반 요금제

- 주 3회 / 5회 요금제 : 정해진 횟수만큼 방문 가능한 요금제, 유동적인 근무 일정을 가진 고객을 대상으로 운영
- 비정기적인 오피스 이용자 또는 하이브리드 근무자를 위한 적합한 요금제

| 횟수권 요금제

- 필라테스, PT 센터에서 운영하는 방식과 유사한 이용권
- 종일권 : 하루동안 자유롭게 이용 가능 (원거리 방문객, 장시간 체류 고객 대상)
- 5시간권 : 짧은 시간 집중 이용객(회의 장소 등)을 위한 합리적 선택지 제공

4-2. 3일 무료 체험 타당성 검토

| 분석 개요

- 3일 무료 체험 및 월 회원권 금액 산정

- 무제한 회원권 월 385,000원 (전 지점, 지정석)
- 하루에 약 12,400원, 3일 무료 체험 약 37,000원
- 연도별 서비스 비용 및 수익 분석

연도	미결제인원	결제인원	전체인원	결제비율(%)	3days 비용	수익	투자
2021	1001	703	1704	41.3	63,388,800원	270,655,000원	23.4
2022	1376	963	2339	41.2	87,010,800원	370,755,000원	23.5
2023	1571	821	2392	34.3	88,982,400원	316,085,000원	28.2

- 방문일 수 비교

결제 여부	방문일 수 (평균)	방문일 수 (중앙값)
결제	1.818571	2.0
미결제	1.637490	1.0

타당성 및 제안

- 무료 체험의 타당성
 - 결제 전환 비율은 약 40% 정도로 높으며 결제 전환 인원들의 지불 금액을 최소로(1달) 잡았을 때의 수익으로 계산을 했을 때 3일 무료 체험의 투자 비용은 약 25%정도이다.
 - 평균 방문 일수는 결제 인원 1.8일, 미결제 인원 1.6일로 전체 방문 인원의 평균이 2일이 넘지 않는다.
- 체험일 수 제안
 - 투자 비용을 줄이고 효율적인 운영을 하기 위해 3일이 아닌 **2일**로 무료 체험일 조정을 제안한다.

5. 제언 및 시사점

5-1. 시사점

신규 요금제 도입을 통한 공유 오피스 적자 해결

분석 배경에서 살펴보았듯이, 최근 공유 오피스의 사용률 감소와 신규 출점 둔화가 지속되고 있다. 실제 분석 결과에서도 연도별 방문자 수 증가율이 둔화되고 있으며, 신규 고객 유입 역시 감소하는 경향을 보였다.

이러한 흐름 속에서 공유 오피스가 다시 성장 궤도에 오르려면, 변화하는 고객 니즈를 반영하여 차별화된 요금제 전략이 필요하다. 특히, 유연한 업무 환경을 원하는 고객층을 적극적으로 공략할 수 있는 맞춤형 요금제를 도입한다면, 공유 오피스는 새로운 성장 기회를 확보할 수 있을 것이다.

5-2. 한계점

현재 데이터로 기본적인 방문자 흐름과 결제 여부 분석은 가능하지만, 더 깊이 있는 인사이트 도출하기 위해서는 행동 패턴 분석, 지점별 최적화, 결제 데이터 확장, 방문 경험 개선 등을 위한 추가 데이터 수집이 필요하다.

3일 체험 데이터

- 방문 일자별 정보를 담고 있지만 사용자가 어떤 활동을 했는지에 대한 세부 정보는 포함되지 않았다.
- 3일 체험 신청 정보만 포함하기 때문에 사용자들의 개별적인 행동 패턴 분석(재방문 여부, 지속적인 이용 여부 등)에 한계가 존재함.

데이터 형식

- 대부분의 데이터가 Datetime 데이터로 존재하고, 일자 외에 속성의 종류가 다양하지 않아 결과 변수와의 영향을 세부 분석하는데 한계가 존재함.

지점별 정보의 한계

- 지점별 데이터는 면적 정보만 포함하고 있으며, 지점의 위치 및 수용 인원 등의 정보가 포함되어 있지 않아 교통 정보, 날씨 등의 외부 데이터를 활용하는데 한계가 존재함.

결제 정보의 한계

- 결제 데이터는 결제 여부만 포함하고 결제 일자, 실 결제 금액, 결제 수단, 환불 여부 등의 정보가 포함되어 있지 않아 세부적인 분석에 한계가 존재함.

보완 방안

- 추가적인 로그 데이터(예: 활동 로그, 사용자 피드백, 세부 결제 정보 등)를 수집하거나 외부 데이터(예: 금융, 날씨, 이벤트 일정 등)와 결합하는 방법을 고려할 수 있다.

5-3. 추가 분석 제안

혼잡도 분석

- 특정 공간(회의실, 핫 데스크, 지정석 등)의 사용 빈도 분석을 통한 좌석 수 및 비용 산정

출퇴근 패턴 분석

- 지하철/버스 정류장 근처 지점의 방문율
- 출퇴근 시간대 이용 패턴과 연계하여 멤버십 상품 최적화 방안 분석

6. 참고자료

- 건물주 우위 시장서 공유 오피스 매력 급감...올해 출점 2곳뿐 / 디지털타임스 / 2024-09-18

https://www.dt.co.kr/contents.html?article_no=2024091802109963066003&ref=naver

- 위워크 파산 4개월, 토종 공유 오피스는 어떻게 살아남았나 / 비즈한국 / 2024-03-08

<https://www.bizhankook.com/bk/article/27252>

- 2024 서울 공유 오피스 현황 및 전망 / 2024 Seoul Flexible Office Market / 2024-09-11

<https://www.cbrekorea.com/insights/reports/2024-서울-공유오피스-현황-및-전망>

- 공유 오피스 '패스트파이브', 엑시트 시계 짹짹 / Dealsite / 2024-12-05

<https://dealsite.co.kr/articles/132622/075033>

- 패스트파이브, 올해 공유 오피스 트렌드 발표 / Byline Network / 2025-02-11

<https://byline.network/2025/02/11-418/>