

Hackerthon

나드리콜 개선 해커톤

외곽 지역 맞춤형 운행 서비스 제공

CONTENTS

1

분석 배경

· 분석 배경

2

분석 과정

· 분석 과정

3

전략 수립

- 아이디어 및 전략

4

비즈니스 모델 캔버스

· BM Canverse

5

결론

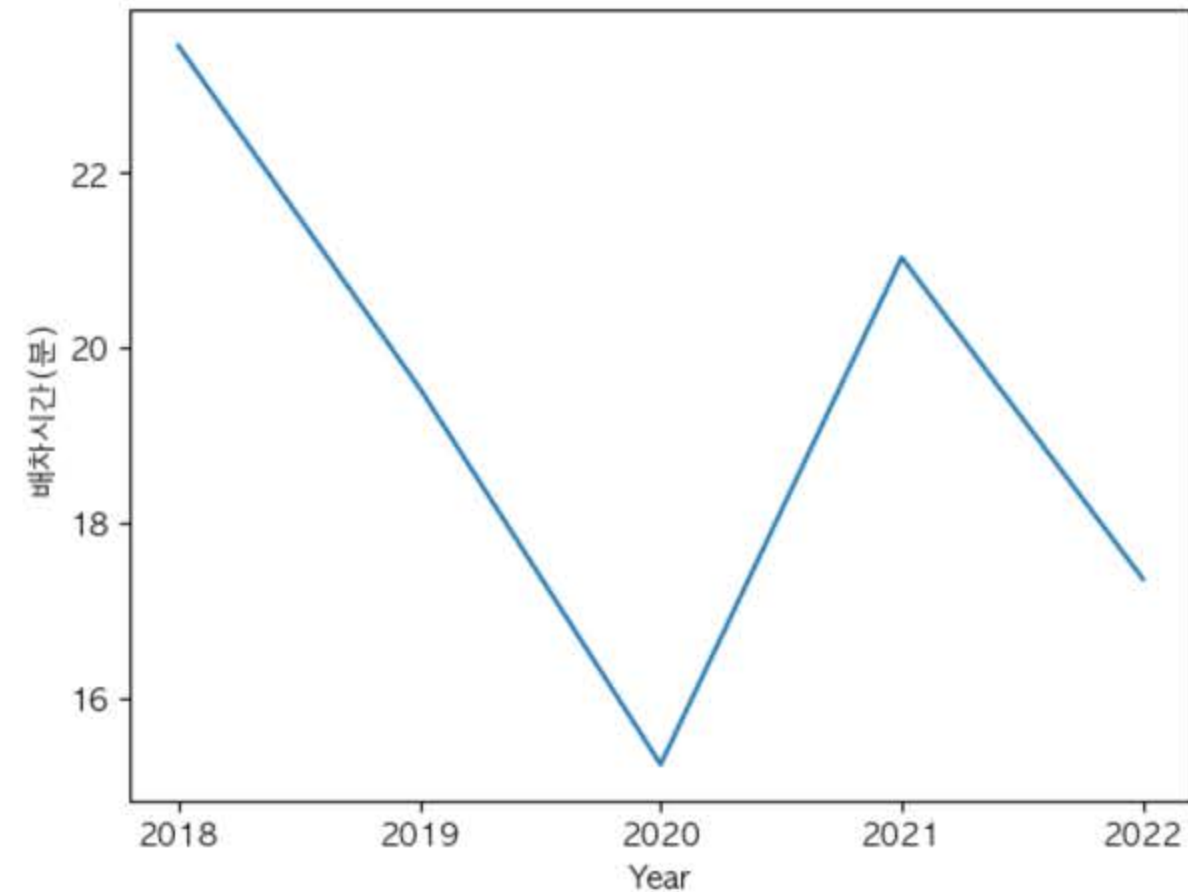
· 기대효과
· 향후 계획

분석 배경

배차시간 선도표

2018	2019	2020	2021	2022
23분 45초	19분 54초	15분 24초	21분 03초	17분 36초

대구공공시설관리공단 나드리콜 운영 종합현황 (2023. 11월 기준) 보고서 내



나드리콜 배차시간을 줄일 수 있는 방안을 마련하여 승객에게 편의성을 제공하고자 함.

분석 과정

STEP 1 데이터 수집 및 전처리

데이터 수집

접수 데이터

나드리콜 접수 데이터
(17~21년도)

DATA 공공데이터포털
GO . KR

평점 데이터

나드리콜 접수 데이터
(21~23년도)
고객 평점 데이터
(21~23년도)

나드리콜



데이터 전처리

시간

- 시간 관련 변수 Datetime 변환
- 하차 시간이 승차 시간보다 빠른 경우 행 삭제
- 접수시간보다 이른 승차 경우 행 삭제
- 이용시간과 배차시간 초(s) → 분(m) 변환

이용시간

- 이용시간 0분 삭제
- 이용 시간 200분 이상 이용하는 경우 많지 않아 삭제

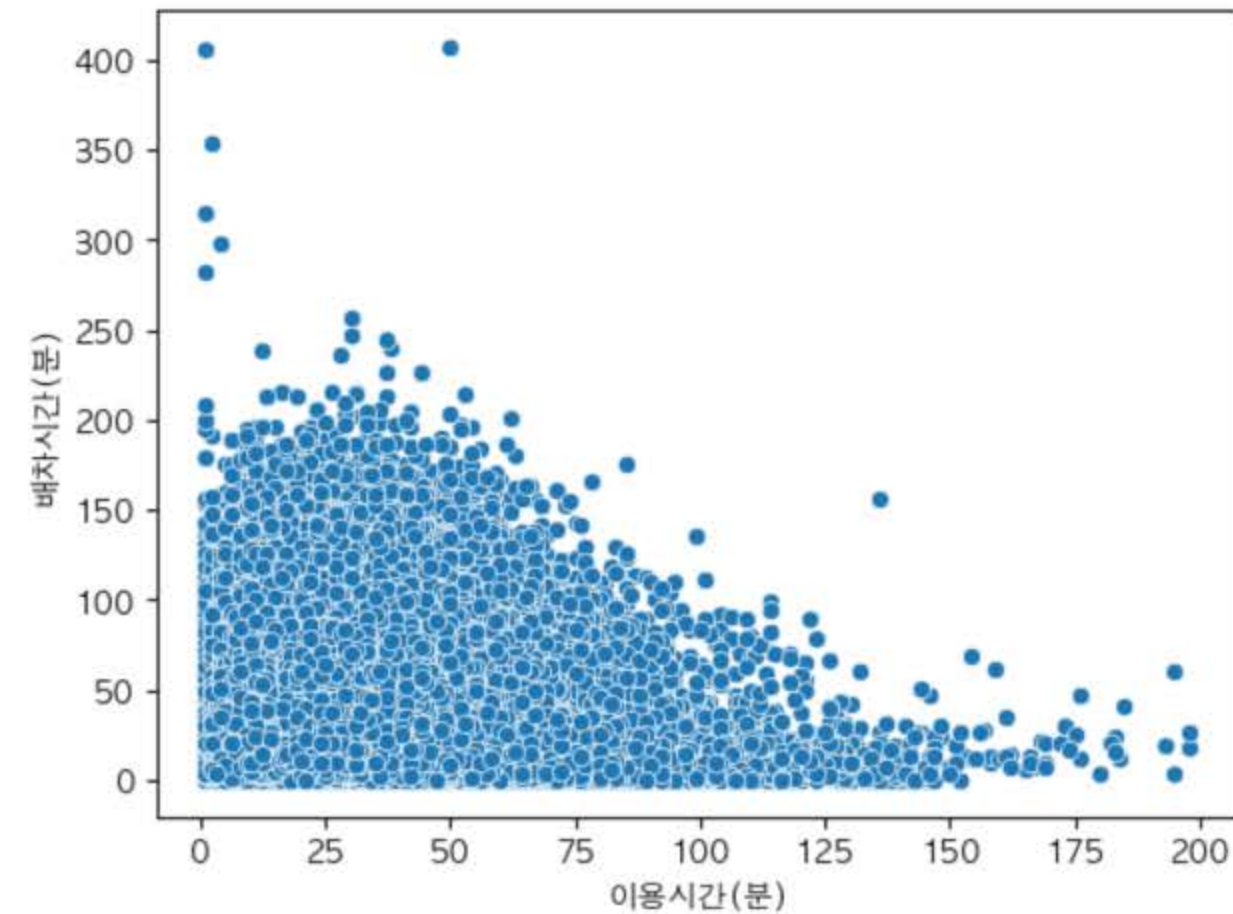
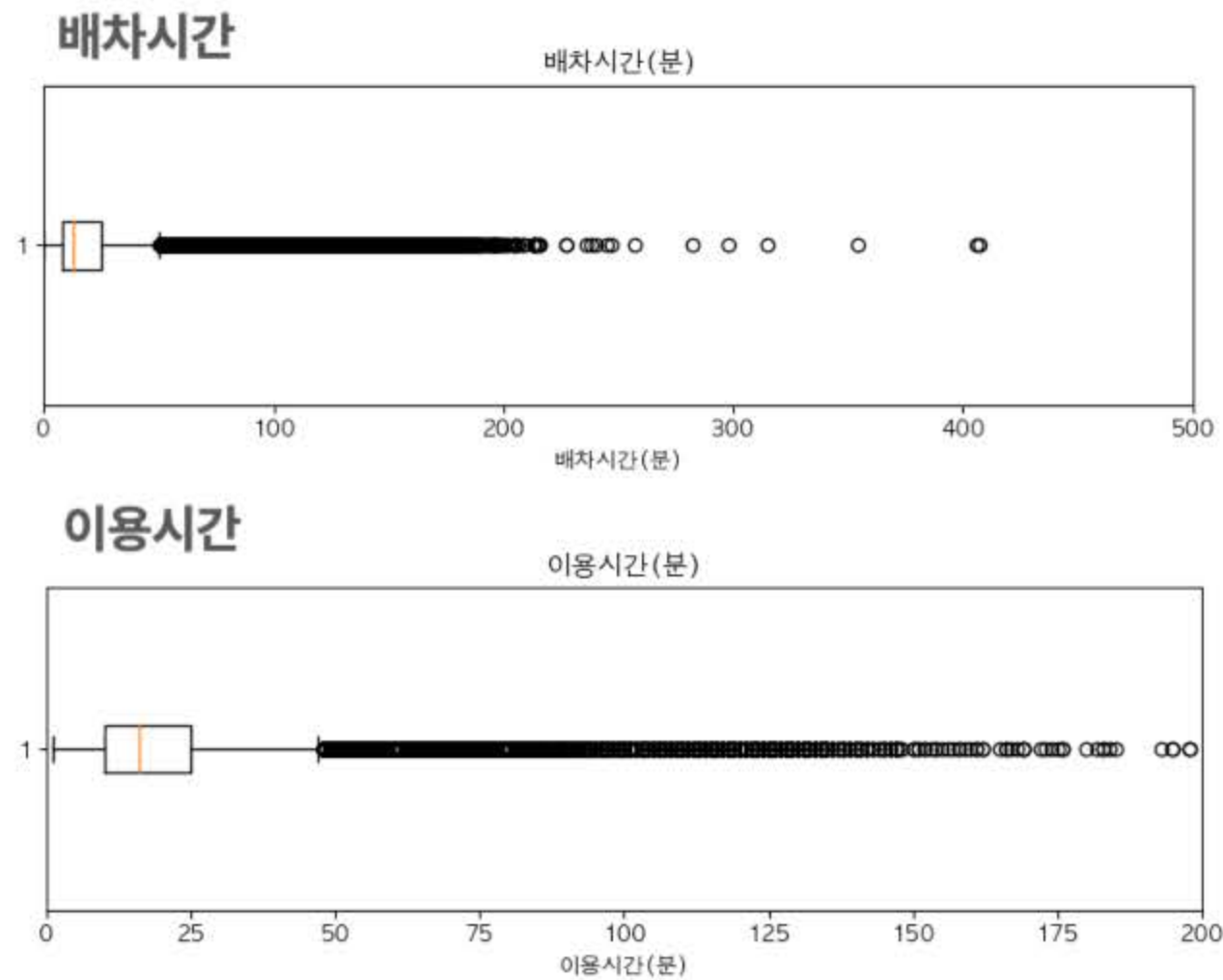
배차시간

- 배차시간이 2500분 이상인 경우 삭제

접수 데이터에는 접수시간, 승하차시간, 이용 및 배차시간
출발지 행정동, 휠체어 유무 등 변수 존재

STEP 2 데이터 분석 (가설 1)

가설 1. 목적지가 가까운 경우 콜을 거절할 것이다.



이용시간과 배차시간은 선형적인 관계를 보이지 않아, 본 가설과 거리가 멀다.

STEP 2 데이터 분석 (가설 2)



ANOVA는 다수의 집단을 서로 비교하고자 할 때 사용하는 통계분석

가설 2. 시간과 요일에 따라 배차시간은 다를 것이다.

시간별

```
groups_hour = [df[df['HOUR'] == dong]['배차시간(분)'] for dong in df['HOUR'].unique()]

# 분산분석 (ANOVA) 수행
f_value, p_value = f_oneway(*groups_hour)

print("F-value:", f_value)
print("P-value:", p_value)

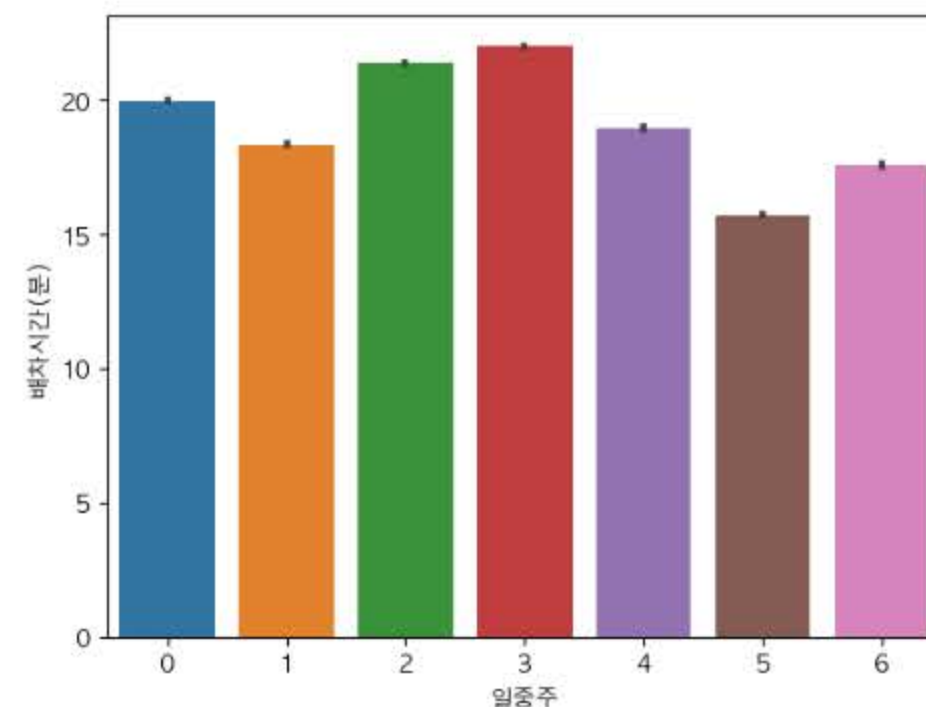
# 유의미한 차이가 있는지 결론 도출
if p_value < 0.05:
    print("시간에 따른 배차 시간에는 통계적으로 유의미한 차이가 있습니다.")
else:
    print("시간에 따른 배차 시간에는 통계적으로 유의미한 차이가 없습니다.")
```

F-value: 2066.1388175576612

P-value: 0.0

시간에 따른 배차 시간에는 통계적으로 유의미한 차이가 있습니다.

요일별



분산분석(ANOVA) 결과, 시간에 따른 배차시간은 평균적으로 차이가 있다.

평일(0~4)이 주말(5~6)보다 배차시간이 길다.

STEP 2 데이터 분석 (가설 3)

가설 3. 행정동에 따라 배차시간은 다를 것이다.

```
from scipy.stats import f_oneway
groups = [df[df['출발지 행정동'] == dong]['배차시간(분)'] for dong in df['출발지 행정동'].unique()]

# 분산분석 (ANOVA) 수행
f_value, p_value = f_oneway(*groups)

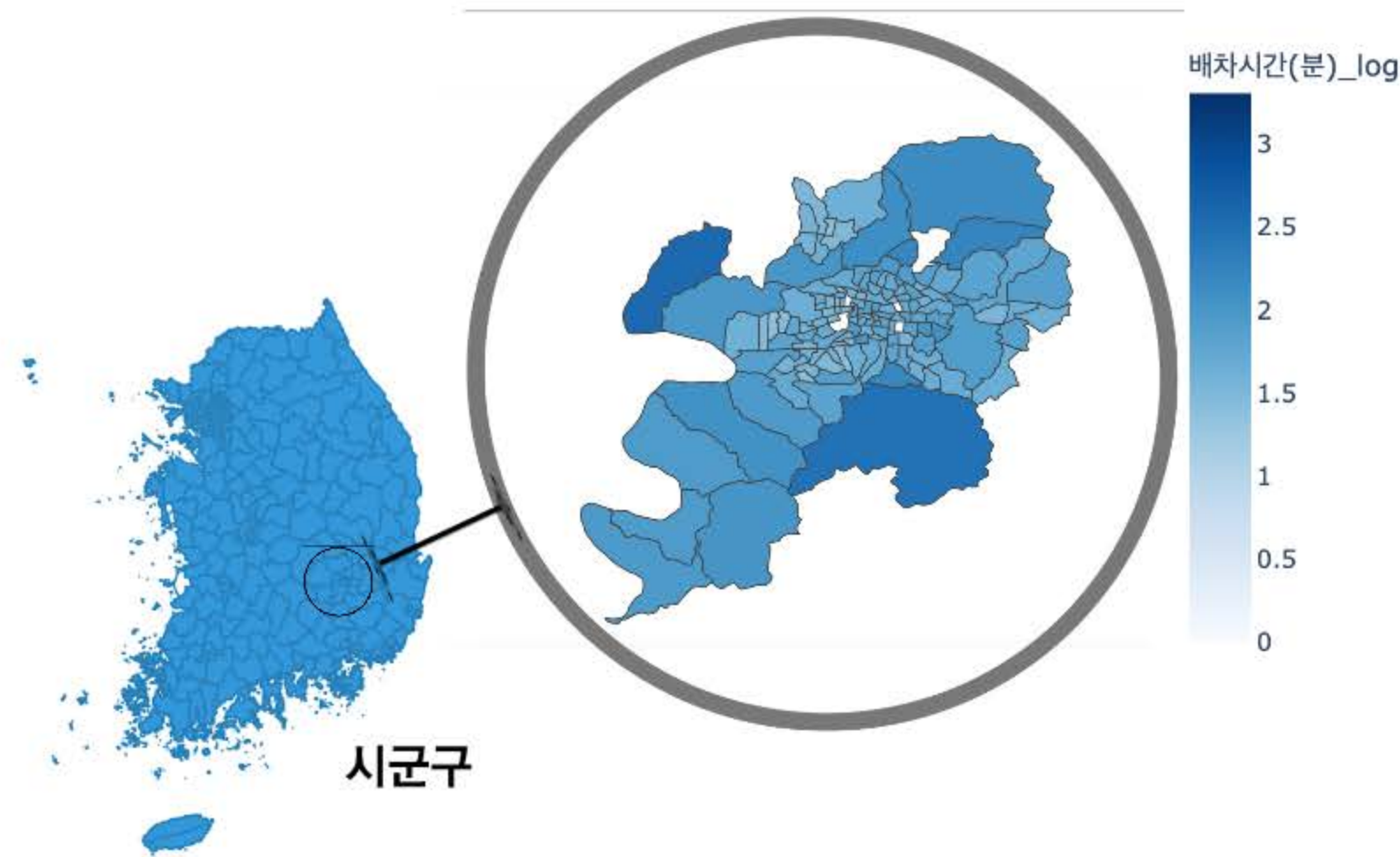
print("F-value:", f_value)
print("P-value:", p_value)

# 유의미한 차이가 있는지 결론 도출
if p_value < 0.05:
    print("행정동에 따른 배차 시간에는 통계적으로 유의미한 차이가 있습니다.")
else:
    print("행정동에 따른 배차 시간에는 통계적으로 유의미한 차이가 없습니다.")
```

F-value: 632.5675834095289

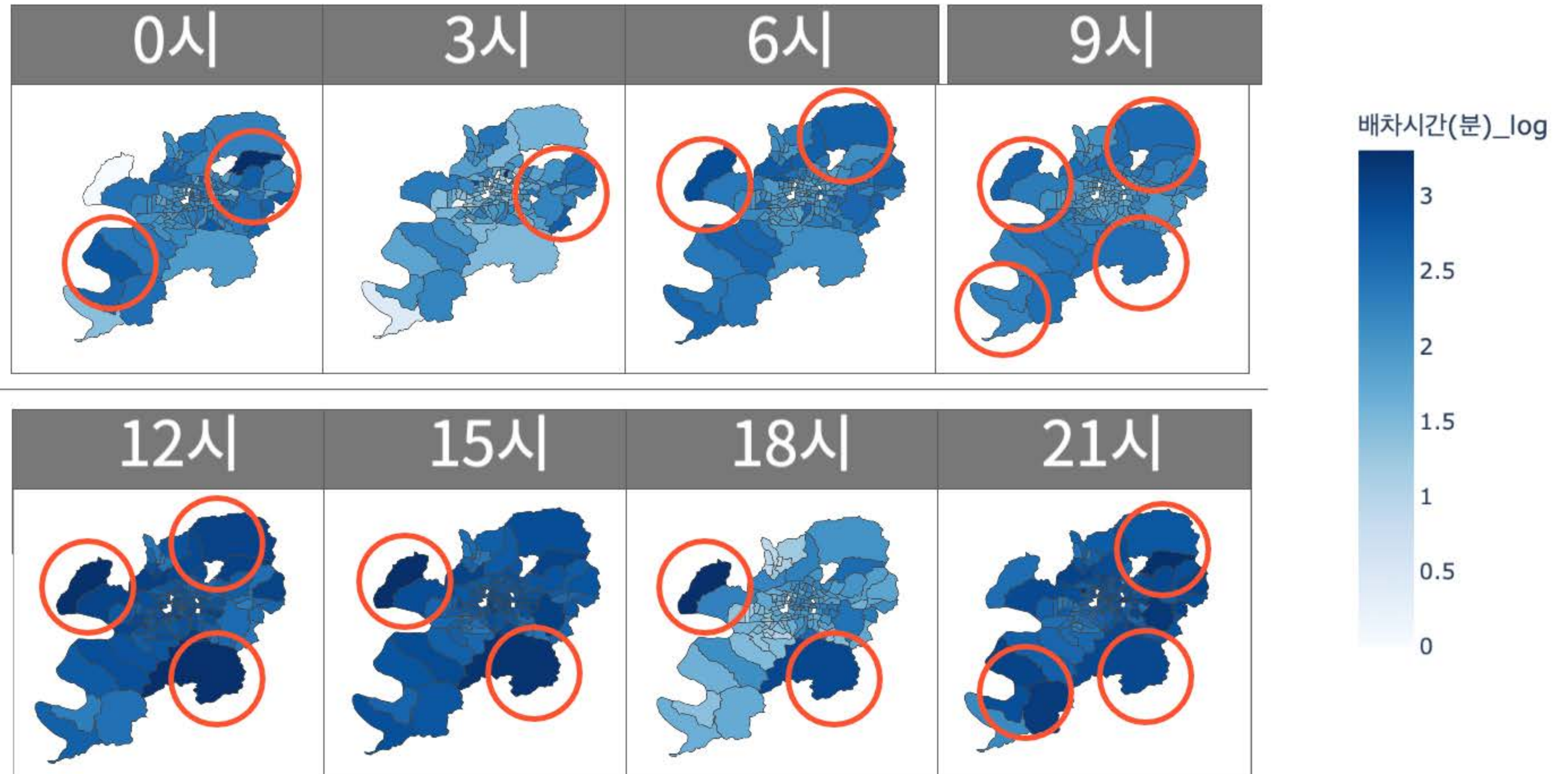
P-value: 0.0

행정동에 따른 배차 시간에는 통계적으로 유의미한 차이가 있습니다.



분산분석(ANOVA) 결과, 행정동에 따른 배차시간은 평균적으로 차이가 있다.

행정동별 배차시간 시각화



시간대와 행정동을 고려한 결과, 외곽 지역의 배차시간은 모든 시간대에 도심지에 비해 길며 지역도 다름.

전략 수립

STEP 3 전략 수립

한정된 예산과 법정대수

시간대와 행정동에 따른 배차시간 차이

도심지 연쇄 운행으로 인한
외곽지역 배차지연

외곽 지역의 시간대별 접수건수 예측으로
맞춤형 운행 서비스를 제공하여
나드리콜 시스템과 결합

STEP 3 전략 수립

접수건수 예측 모델 구축

변수 정의

행정동별로 그룹화한 후

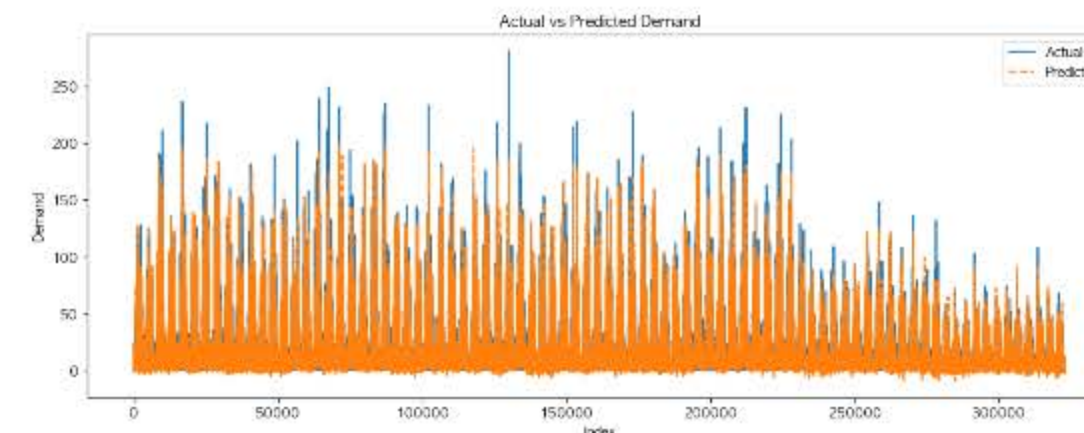
요일, 일, 월, 행정동을 독립변수로

접수건수를 종속변수로 하여 회귀 모델 구축

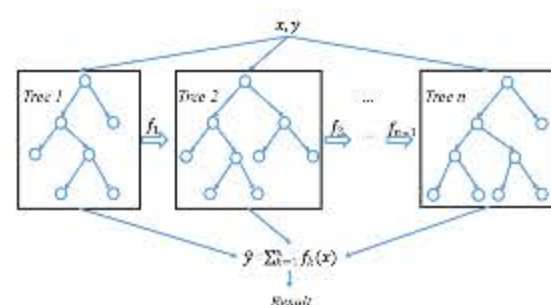
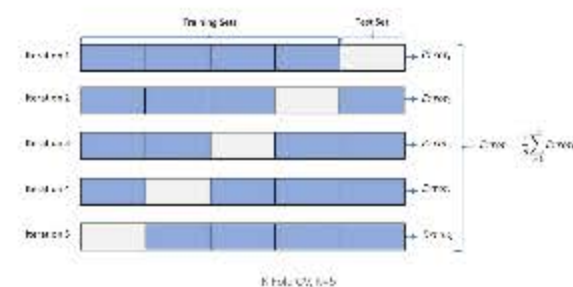
파라미터

- eta (learning rate) : 0.3
- max_depth : 6
- num_estimators : 500
- random_state : 42

예측 결과물



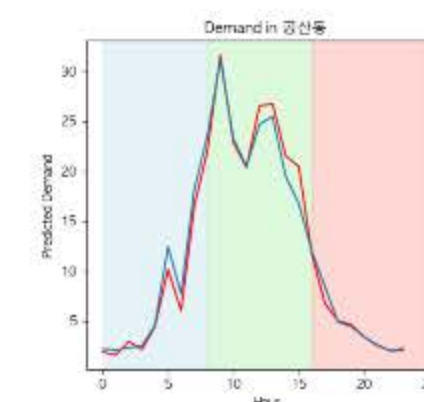
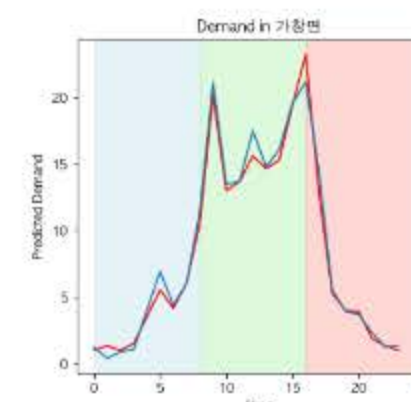
5 Fold Cross Validation XGBoost Model



성능 평가

- MSE : 5.125
- R2 : 0.921

평균제곱오차 합은 5.12이며
92.1%의 설명력을 가짐.



정규화항을 사용한 일반화 성능 보장과 큰 데이터 처리에 능숙한 XGBoost를 사용함.

공휴일과 같은 잠재요인으로 Peak한 구간이 존재

STEP 3 전략 수립

스팟존 구축 개념

배차시간이 긴 외곽 지역에서 가장 많은 이용자가 있는 시간을 분석 및 예측하여 지도에 스팟존을 표시



STEP 3 전략 수립

초기 화면



버튼 클릭 시



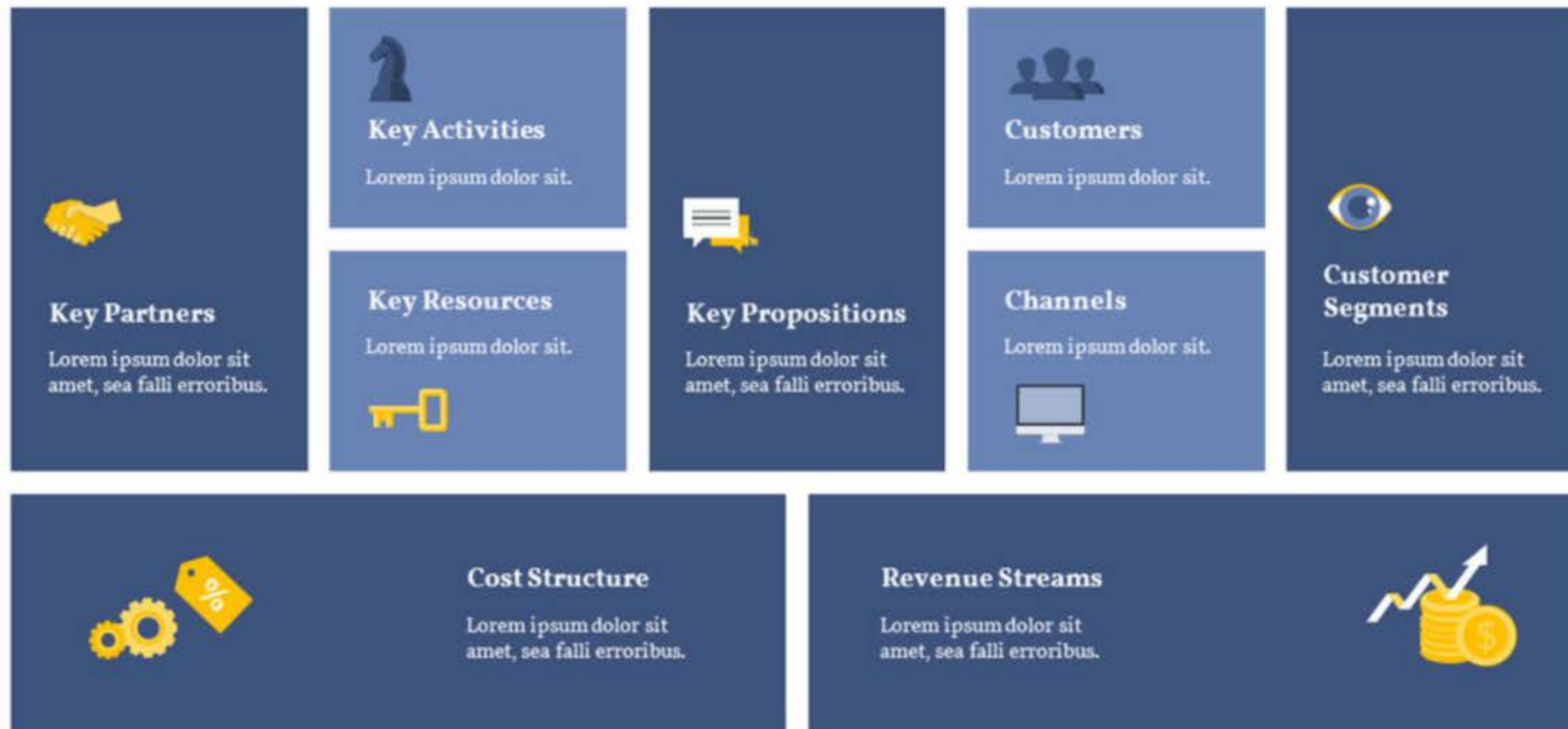
스팟존이 생성될 때



비즈니스 모델 캔버스

비즈니스 모델 캔버스 개념

비즈니스 모델 캔버스는 비즈니스에 포함되어야 하는 9개의 주요 사업 요소를 한눈에 볼 수 있도록 만든 그래픽 템플릿





비즈니스모델 캔버스(Business Model Canvas)

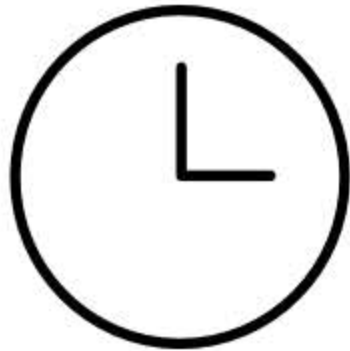
7.핵심 파트너 (Key Partners) <ul style="list-style-type: none">● 나드리콜 운전기사● 개인택시 운전기사	8.핵심 활동 (Key Activities) <ul style="list-style-type: none">● 배차가 긴 외곽지역의 콜을 시계열 분석을 통해 스팟존 생성	2.가치 제안 (Value Proposition) <ul style="list-style-type: none">● 배차시간 감소● 콜 예측으로 휴차시간 감소	4.고객 관계 (Customer Relationship) <ul style="list-style-type: none">● 배차시간 단축으로 승객 만족도 향상● 콜 예측으로 기사 수입 증가	1.고객분류 (Customer Segment) <ul style="list-style-type: none">● 노약자 및 장애인 등● 운전 기사
	6.핵심 자원 (Key Resource) <ul style="list-style-type: none">● 접수 데이터● 스팟존 근처 차량 선택지 창 제공● 스팟존 위치 제공		3.채널 (Distribution Channel) 나드리콜 어플	
9.비용 구조 (Cost Structure) <ul style="list-style-type: none">● 딥러닝 모델 개발 비용● 개발 하드웨어 구입 비용			5.수익 흐름 (Revenue Stream) <ul style="list-style-type: none">● 택시 이용료● 나드리콜 어플 내 광고 등록	

기대효과 및 향후계획

5

기대효과

승객



평균 배차시간 감소

기사의 휴차시간 감축으로 평균 배차시간 감소

나드리콜



고객 만족도 향상

승객과 기사에게 효율적인 서비스를 제공해
고객 만족도 향상과 가치 제안

기사



수익 증가

회귀모형을 통한 시간대별 예상 콜 수 제공해
휴차시간 감소와 수익 증가

향후 계획

A



외곽과 도심지의 배차시간 조율

스팟존 운영시 외곽에서 줄어드는 배차시간과
도심지에서 늘어나는 배차시간을 계산하여
Trade-off 조율

B



구체적인 콜 좌표 제공

스팟존 위경도접수콜 위경도 데이터를 활용하여
세밀하게 분석해 구체적인 예측 콜 좌표 제공



Thank You