

---

---

# 서울시 상권 & 고객 분석 서비스

---

---

정수연

---

---

# 목차

- 01 프로젝트 배경
- 02 프로젝트 일정
- 03 팀 구성 및 역할
- 04 수행결과 및 시연
- 05 기대효과
- 06 느낀점
- 07 향후과제

# 01. 프로젝트 배경

- 중국집을 새로 오픈하려는 사장님
  - 어느 상권에서 어떤 고객을 타겟팅하면 좋을까?
- 우리가 배운 내용으로 고민을 해결해보자
  - 서울시 우리동네 상권 데이터, 상권별 좌표 API 활용
  - Colab 환경에서 데이터 분석, 머신러닝, 시각화

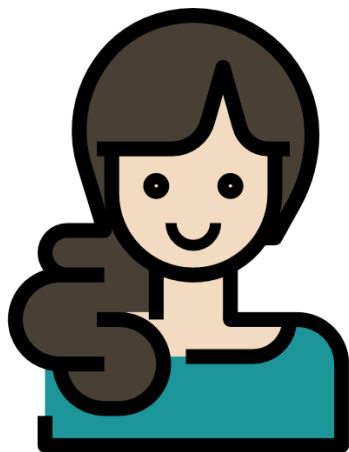


## 02. 프로젝트 일정

분류	항목	12월 2~3주차							12월 4주차				12월 5주차 ~ 1월 1주차						
		9	12	13	14	15	16	17	20	21	22	23	26	27	28	29	30	2	
기획	프로젝트 주제 선정																		
	마일스톤 (중간 목표) 확립																		
	세부 일정 수립																		
데이터 수집 & 전처리	공공데이터 수집																		
	데이터 통합																		
	데이터 전처리																		
멘토링	팀미팅 & 중간발표																		
	멘토링																		
데이터분석 & 시각화	업종별 매출액 분석																		
	상권 class별 매출 건수 분석																		
	상권별 고객 특성 분석																		
	지도 내 통합 시각화																		
머신러닝	예측 모델 컨셉 논의																		
	모델 성능 테스트																		
	개별 모델 세부사항 확인																		
	모델 최종 선정																		
테스트 & 서비스 공개	최종 결과물 검토																		
	발표 자료 작성																		
	발표 및 시연																		

공동  
휴가

### 03. 팀 구성 및 역할



정수연

# 데이터 분석  
# 시각화  
# 머신러닝



김준호

# 데이터 분석  
# 시각화

## 04. 수행 절차 및 방법 (1) 데이터 수집, 탐색

- Main 데이터

- 2021 서울시 상권분석 (CSV)



- 좌표 데이터

- 서울시 상권영역 좌표 (API)

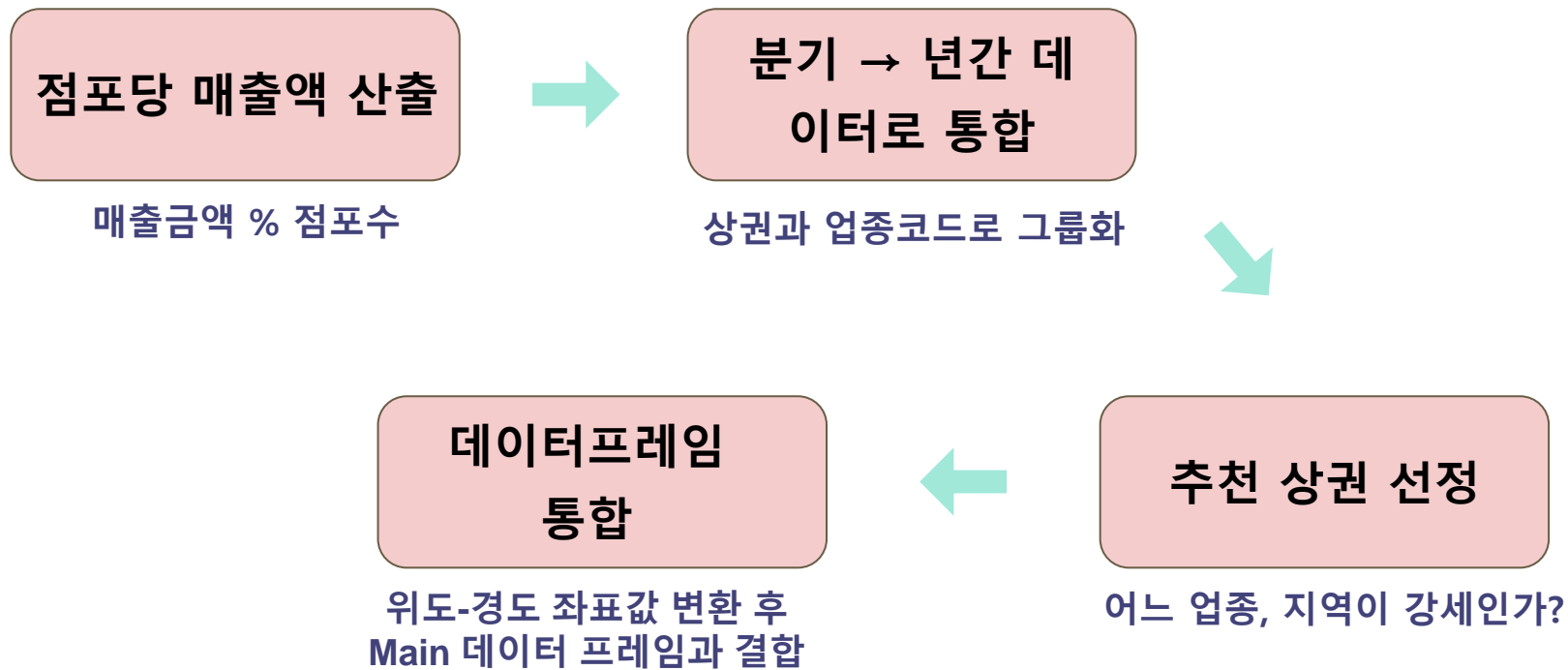
상권 코드	X좌표	Y좌표
-------	-----	-----

- 수도권 지하철역 좌표 (Excel)

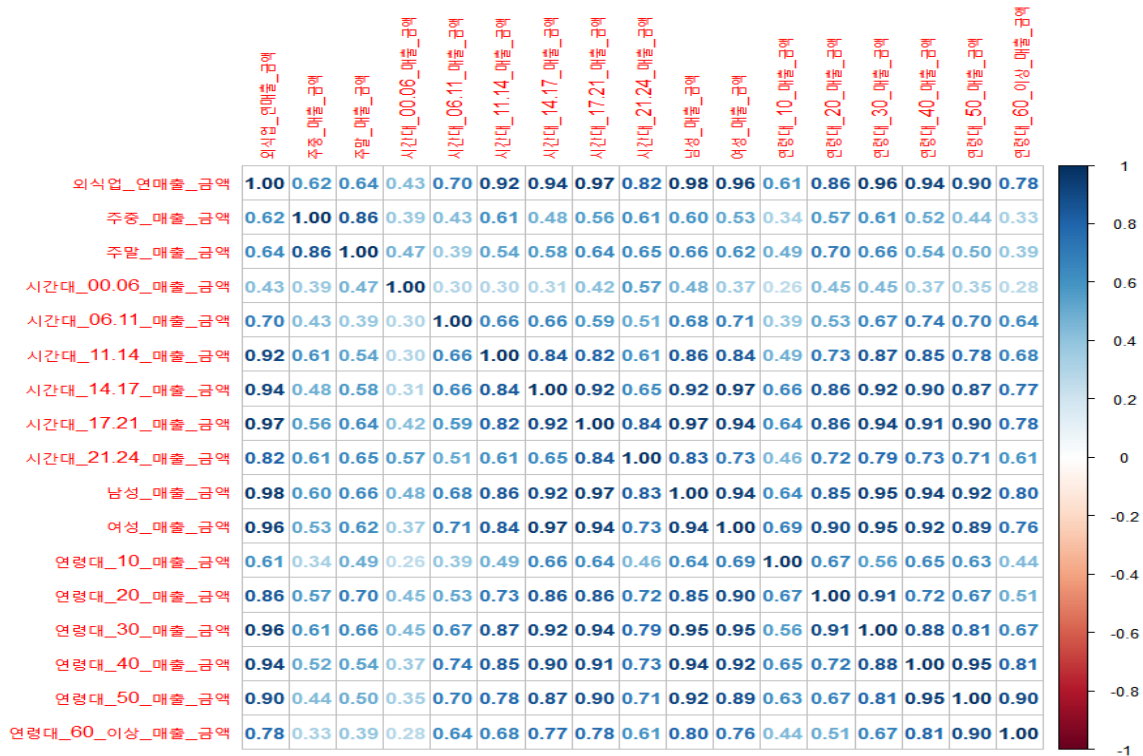
역명	경도	위도
----	----	----

기준 분기	1분기 ~ 4분기
상권 구분	골목상권, 발달상권, 전통시장, 관광특구
상권명, 코드	1606개 지역 (ex. 독립문역 1번)
업종 구분	외식업, 서비스업, 소매업
업종명	10개 업종 (ex. 한식음식점, 제과점...)
고객	성별, 연령, 요일, 시간대
매출	매출액, 매출건수

## 04. 수행 절차 및 방법 (2) 데이터 전처리



## 04. 수행 절차 및 방법 (3) 데이터 분석 <데이터 간 상관관계>

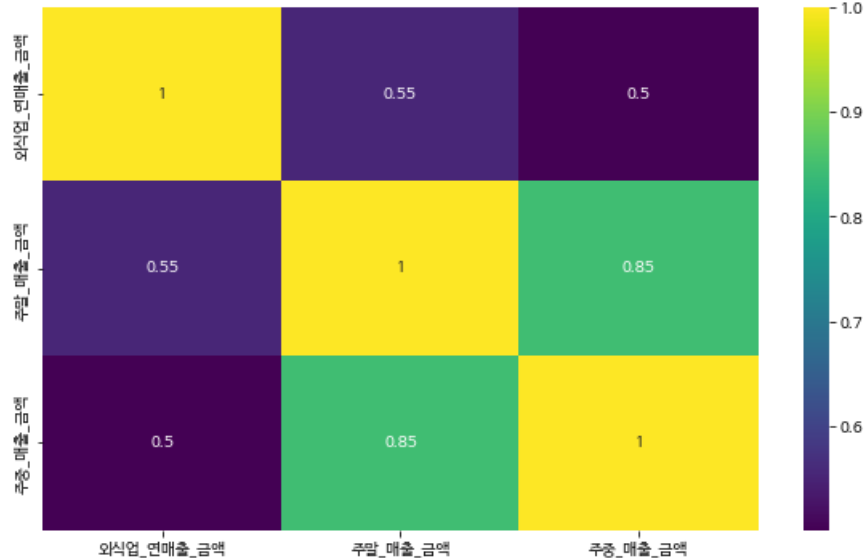




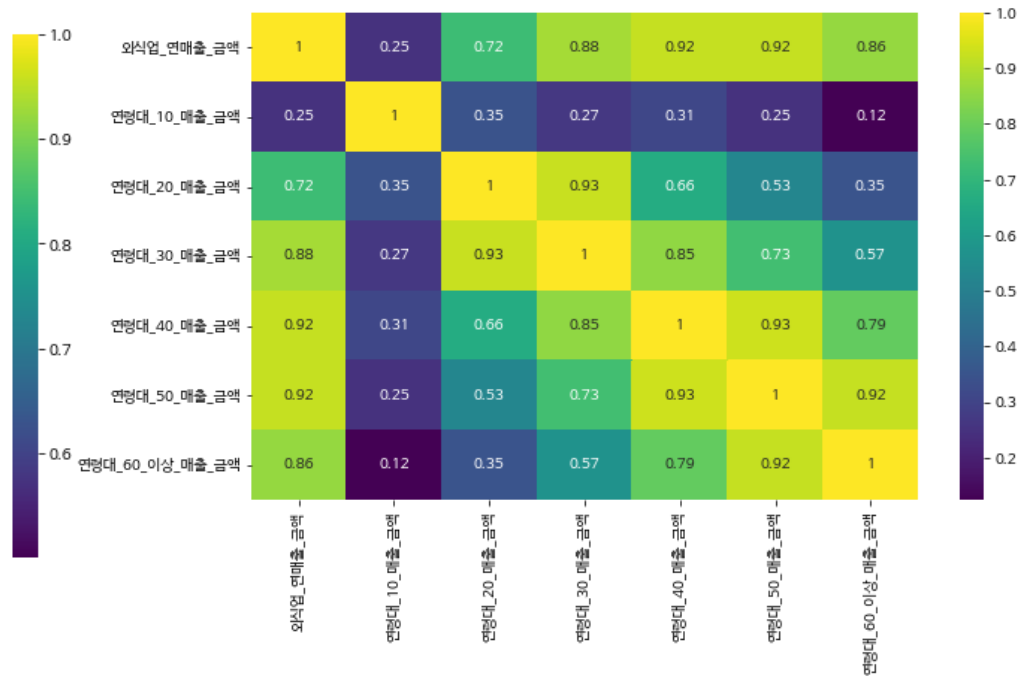
# 04. 수행 절차 및 방법

## (3) 데이터 분석 <업종별 고객특성>

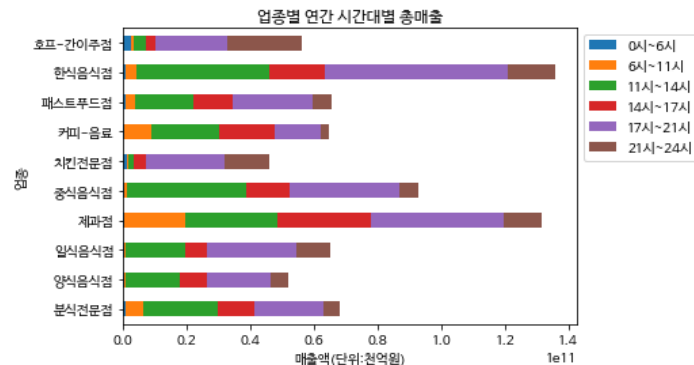
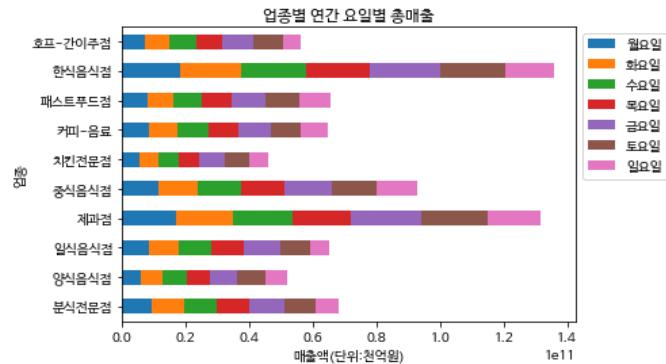
중식당의 총 매출 - 주말, 주중 매출간 상관관계



중식당의 총 매출 - 연령대별 매출간 상관관계

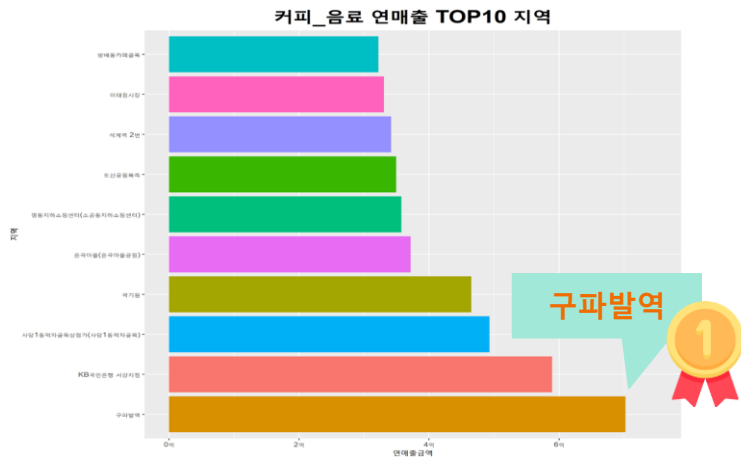
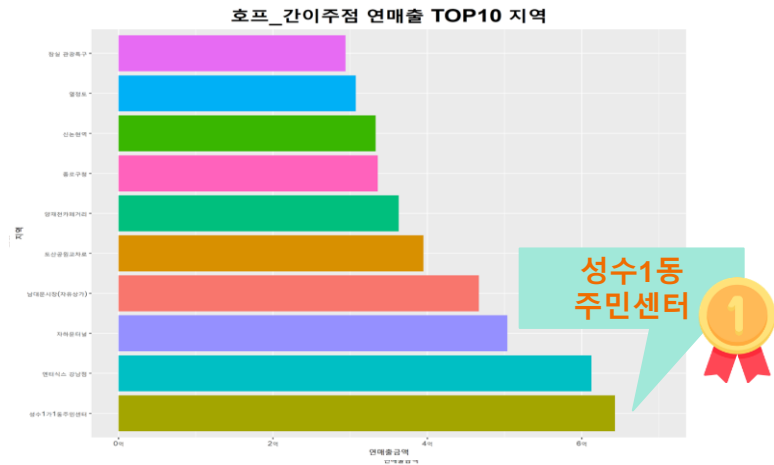
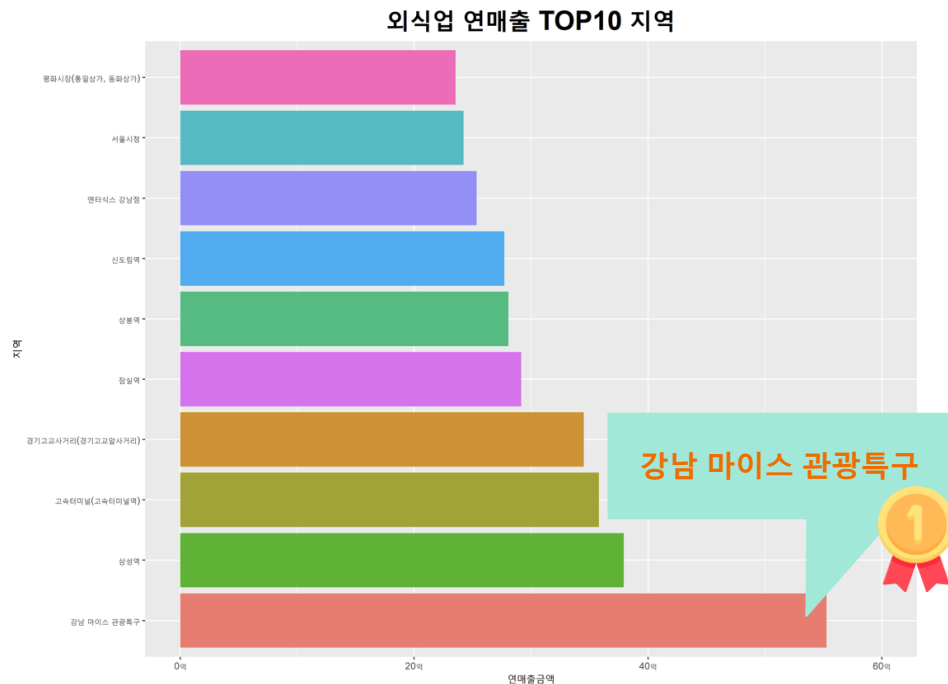


### (3) 데이터 분석 <업종별 매출>



# 04. 수행 절차 및 방법

## (3) 데이터 분석 <전체 외식업 지역별 TOP10>



# 04. 수행 절차 및 방법

## (3) 데이터 분석 <최종 서비스>

길음역 10번 한식당 주요 고객: ['#토요일', '#저녁', '#남성', '#60대 이상']



한식음식점



중식음식점



일식음식점



양식음식점



제과점



패스트푸드점



치킨전문점



분식전문점



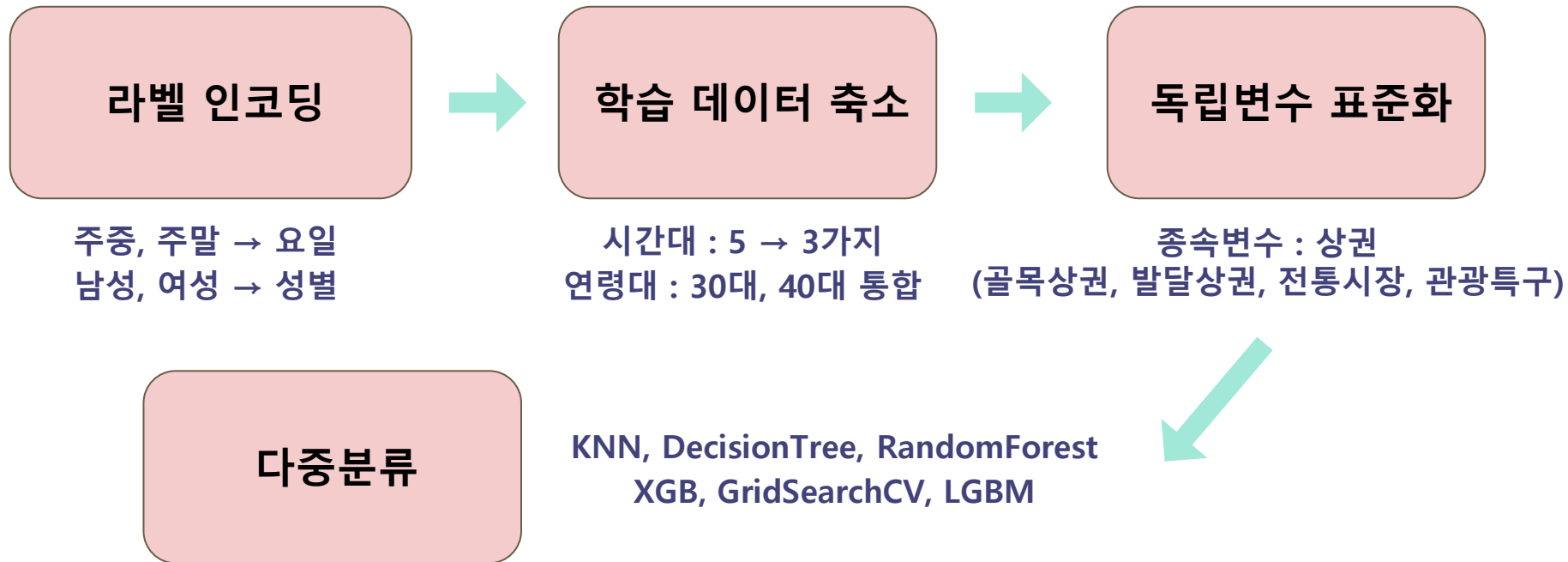
호프-간이주점



카페-음료

## 04. 수행 절차 및 방법

### (4) 머신러닝 <상권 특성 예측 모델링>



# 04. 수행 절차 및 방법

## (4) 머신러닝 <상관 특성 예측 모델링> 성능평가지표 : 정확도(accuracy\_score)

	하이퍼 파라미터	최고 예측 정확도
KNN	K = 7	0.673
DecisionTree	max_depth = 6	0.682
RandomForest	max_depth = 13	0.701
GridSearchCV (random forest 튜닝)	max_depth = 10 min_samples_split = 6 cv = 3	0.694
XGB	max_depth = 6 eval_metric='logloss'	0.705
LGBM	max_depth = 3 n_estimators = 400 eval_metric='logloss'	0.705

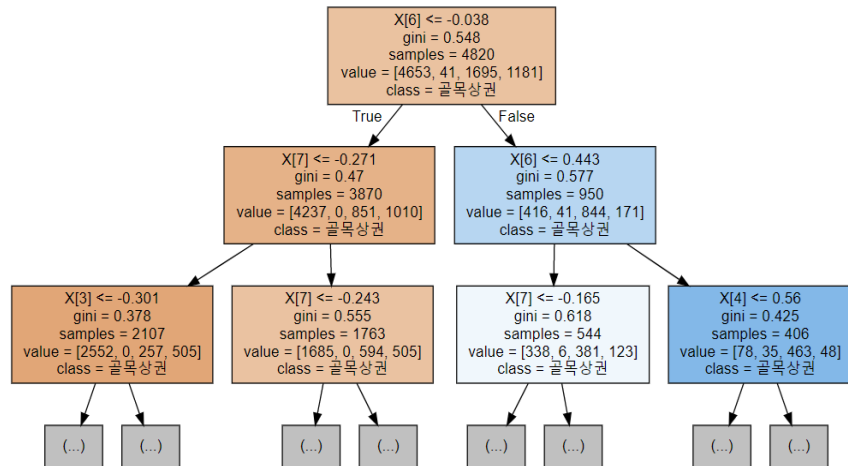
train score = 0.7941  
test score = 0.7052

train score = 0.7546  
test score = 0.7052

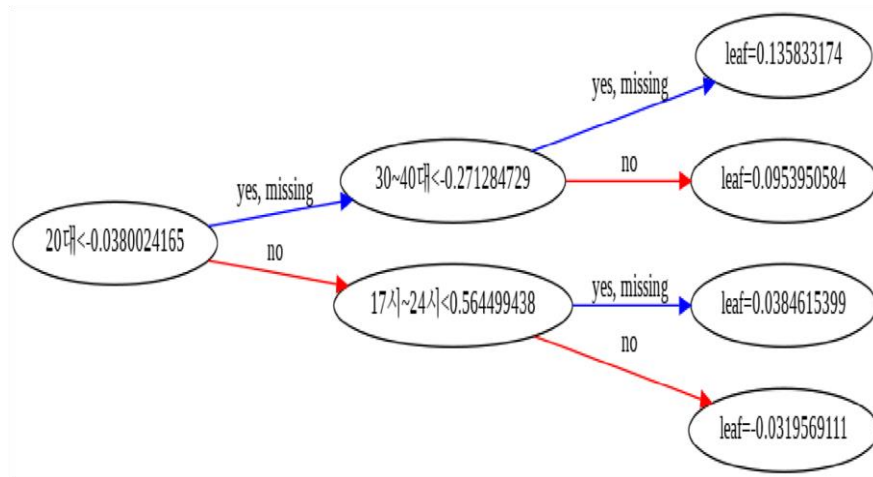
# 04. 수행 절차 및 방법

## (4) 머신러닝 <트리 시각화>

<Random Forest>



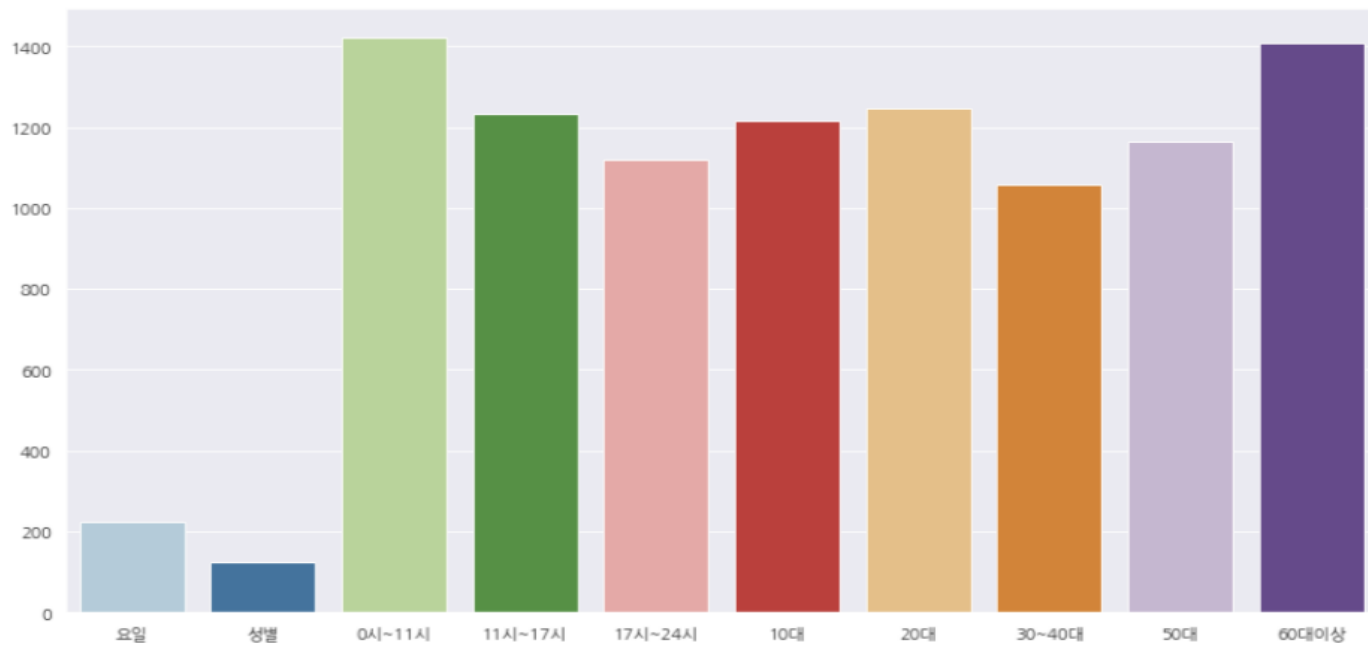
<XGB>



## 04. 수행 절차 및 방법

### (4) 머신러닝 <Feature별 중요도 시각화>

	feature	importance
0	요일	223
1	성별	125
2	0시~11시	1421
3	11시~17시	1232
4	17시~24시	1121
5	10대	1214
6	20대	1248
7	30~40대	1058
8	50대	1163
9	60대이상	1407





## 05. 기대효과

- 외식업종별 신규 자영업자에게 창업 지역 추천
- 고객 특성에 맞는 매장컨셉 or 마케팅 전략 수립

## 06. 느낀점

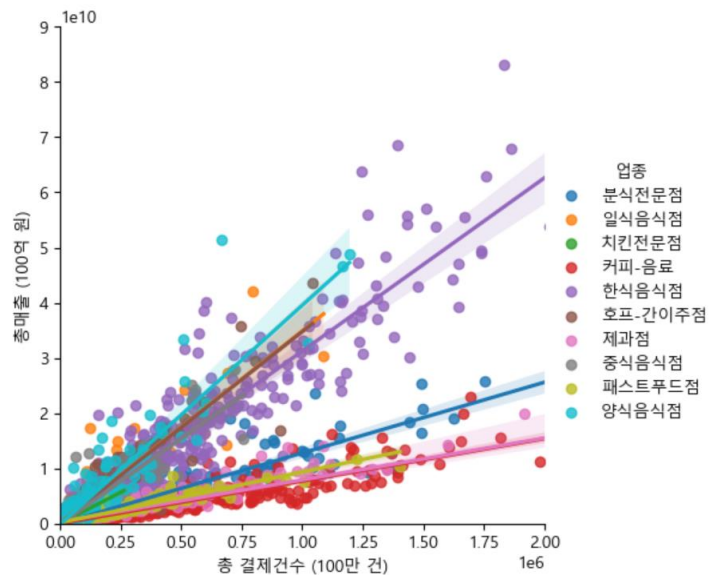
### 정수연

- 팀원 수가 적음에도 각자 맡은 일을 적극적으로 수행하여 최종목표인 머신러닝까지 완성하였다.
- 초점을 어디에 두느냐에 따라 해석이 바뀌어 여러 의견을 정리하는데 어려움이 있었다.

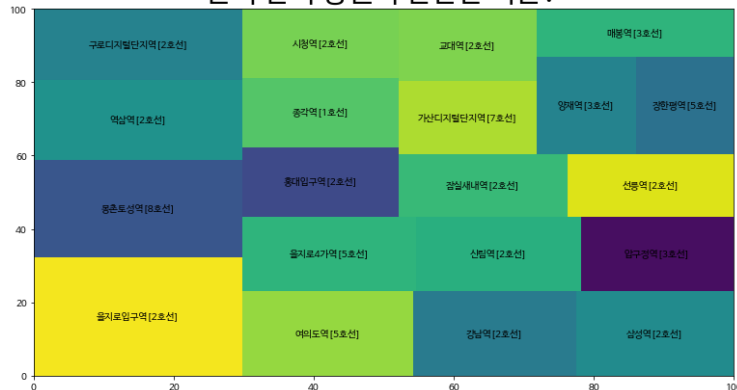
### 김준호

- 재미있었다

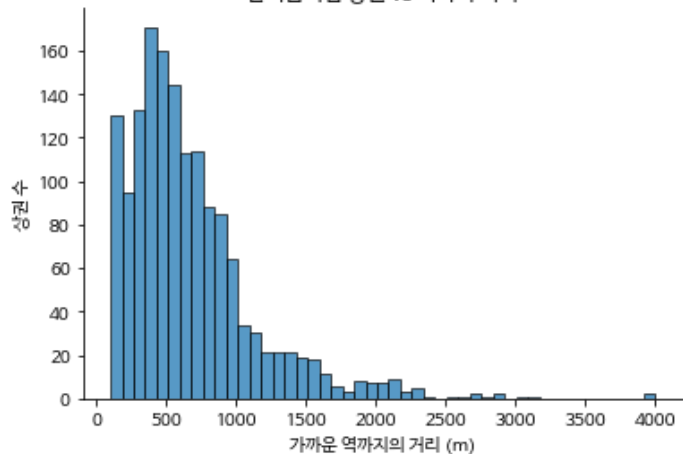
## 07. 향후 과제



근처 한식 상권이 발달한 역은?



한식음식점 상권 vs 역까지 거리



감사합니다😊