
제주지역 음식물 쓰레기 배출량 감소를 위한 예측 모델 수립 및 주요요인 도출

2조 러닝머신

목차

1. 프로젝트 개요

기획 배경 및 목표

구성원 및 역할

2. 데이터 분석

데이터 수집

탐색적 데이터 분석

데이터 분석 모델

데이터 분석 및 결과

3. 데이터 분석 결과

인사이트 도출

문제 현황 분석

결론 및 제언

4. 후기 및 느낀점

1. 프로젝트 개요



1. 프로젝트 개요

1-1. 프로젝트 기획 배경 및 목표

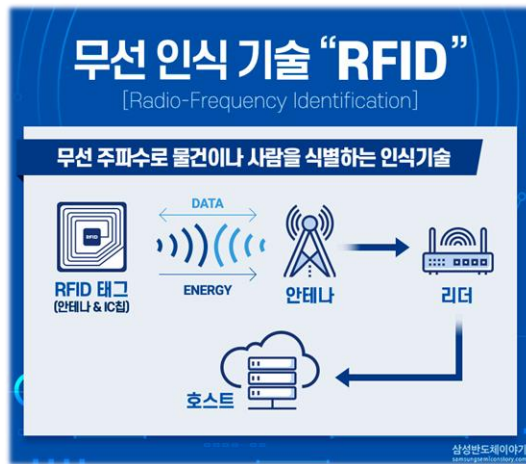


제주 전지역 RFID를 활용한
음식물 쓰레기 배출 시스템 도입 완료



음식물 쓰레기 배출 데이터 구축
일/행정동별 배출건수/배출량/결제금액

“음식물 쓰레기 배출 주요요인 도출”



1. 프로젝트 개요

1-2. 구성원 및 역할

이름	소개	담당 역할
윤보람	<ul style="list-style-type: none"># 융합 지성인# 탄탄한 배경지식# 꼼꼼함과 세심함 	<ul style="list-style-type: none">• 조장 - 프로젝트 운영• EDA 및 인사이트 도출• 데이터 분석 시각화
김형림	<ul style="list-style-type: none"># A to Z형# 만능 재주꾼# 팀 플레이어 	<ul style="list-style-type: none">• EDA 및 인사이트 도출• 데이터 분석• 분석 결과 시각화
박형준	<ul style="list-style-type: none"># 호기심왕# 타고난 진행 본능# 샘솟는 비즈니스 아이디어 	<ul style="list-style-type: none">• EDA• 데이터 분석• 결과 해석 및 인사이트 도출
장우창	<ul style="list-style-type: none"># 재능 원석# 거침없이 질주# 눈보다 빠른 손 	<ul style="list-style-type: none">• EDA• 데이터 전처리• 분석 및 결과 정리

2. 데이터 분석

1. 데이터

데이터 수집

데이터 소개

2. EDA

거주인구

유동인구

음식 관련 카드 소비

음식물 쓰레기

3. 데이터 분석 및 결과

데이터 전처리

데이터 분석

데이터 분석 결과

2. 데이터

1-1. 데이터 수집

데이터명	설명		기간	제공 기관
음식물 쓰레기	일/행정동별 음식물 쓰레기 배출건수/량 및 총 지불금액 ※ 음식점 음식물 쓰레기 제외		2018. 01 ~ 2021. 06	환경관리공단
유동인구	내국인	해당 시각 정각에 측정한 일/행정동별 유동인구 (ex. 30분은 0.5로 집계)		KT
	단기 체류 외국인			
	장기 체류 외국인			
거주인구	월/행정동별 거주인구			제주도
음식 관련 카드 소비	일/행정동/업종별 음식 관련 카드 결제 내역			신한카드
날씨	일/행정동별 날씨 데이터(평균기온, 일강수량 등)			기상청
공휴일	연도별 주말(토, 일) 및 공휴일			공공데이터

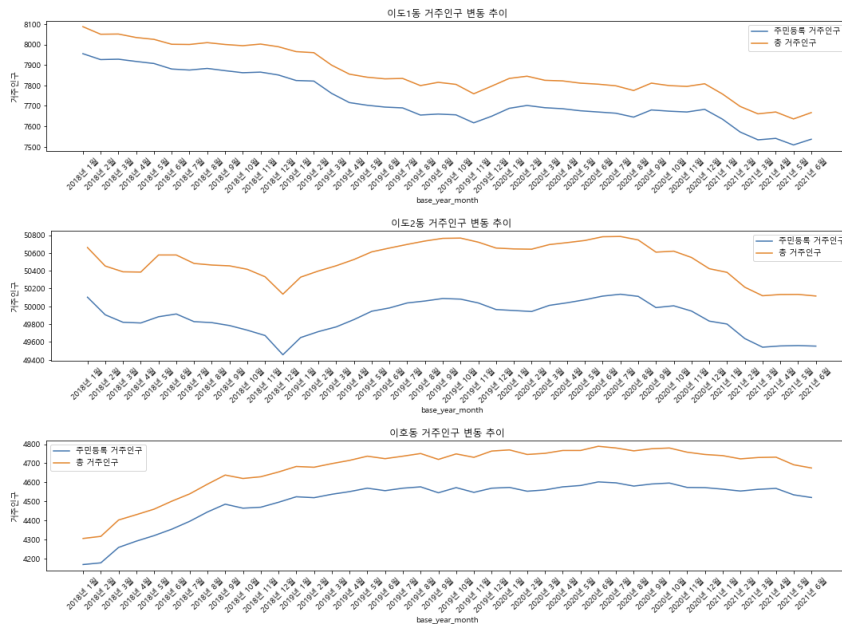
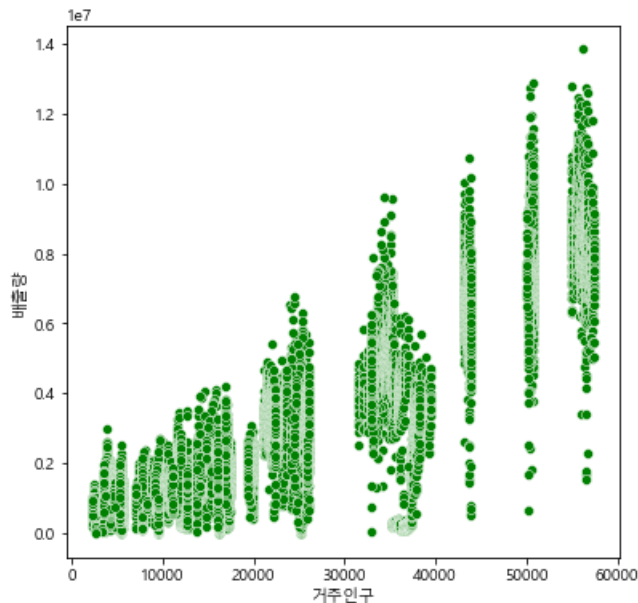
2. 데이터

1-2. 데이터 소개

음식물 쓰레기		유동인구		거주인구	
컬럼 ID	컬럼명	컬럼 ID	컬럼명	컬럼 ID	컬럼명
base_date	기준일자	time	집계 시각	resid_reg_pop	주민등록 거주인구
city	시 구분	sex	성별	foreign_reg_pop	외국인 거주인구
emd_cd	행정동 코드	age	연령대	total_pop	총 거주인구
emd_nm	행정동명	resd_pop_cnt	거주인구	음식 관련 카드 소비	
em_area_cd	배출거점지역 코드	work_pop_cnt	근무인구	mct_cat_cd	음식 관련 업종 코드
em_cnt	배출건수	visit_pop_cnt	방문인구	mct_cat_nm	음식 관련 업종명
em_g	배출량	resd	거주지역 행정구역	use_cnt	결제건수
pay_amt	총 지불금액	nationality	외국인 국적 코드	use_amt	결제금액

2. EDA

2-1. Feature ① : 거주인구

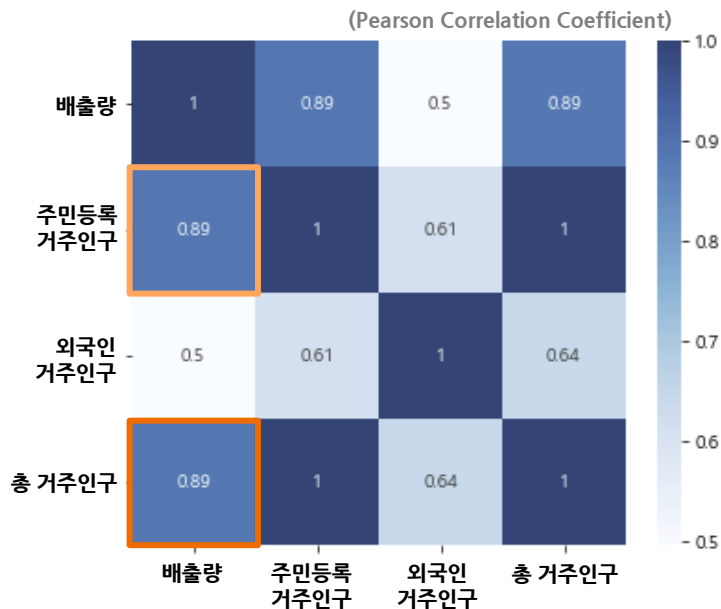
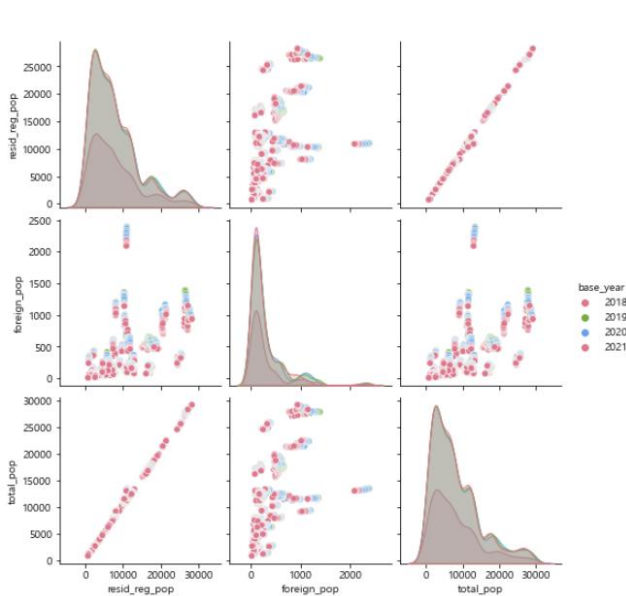


행정동별 거주인구는 1,000명 ~ 60,000명까지 다양하고,
거주인구가 증가하거나 감소, 혹은 등락이 반복되는 등 다양한 양상을 보임

2. EDA

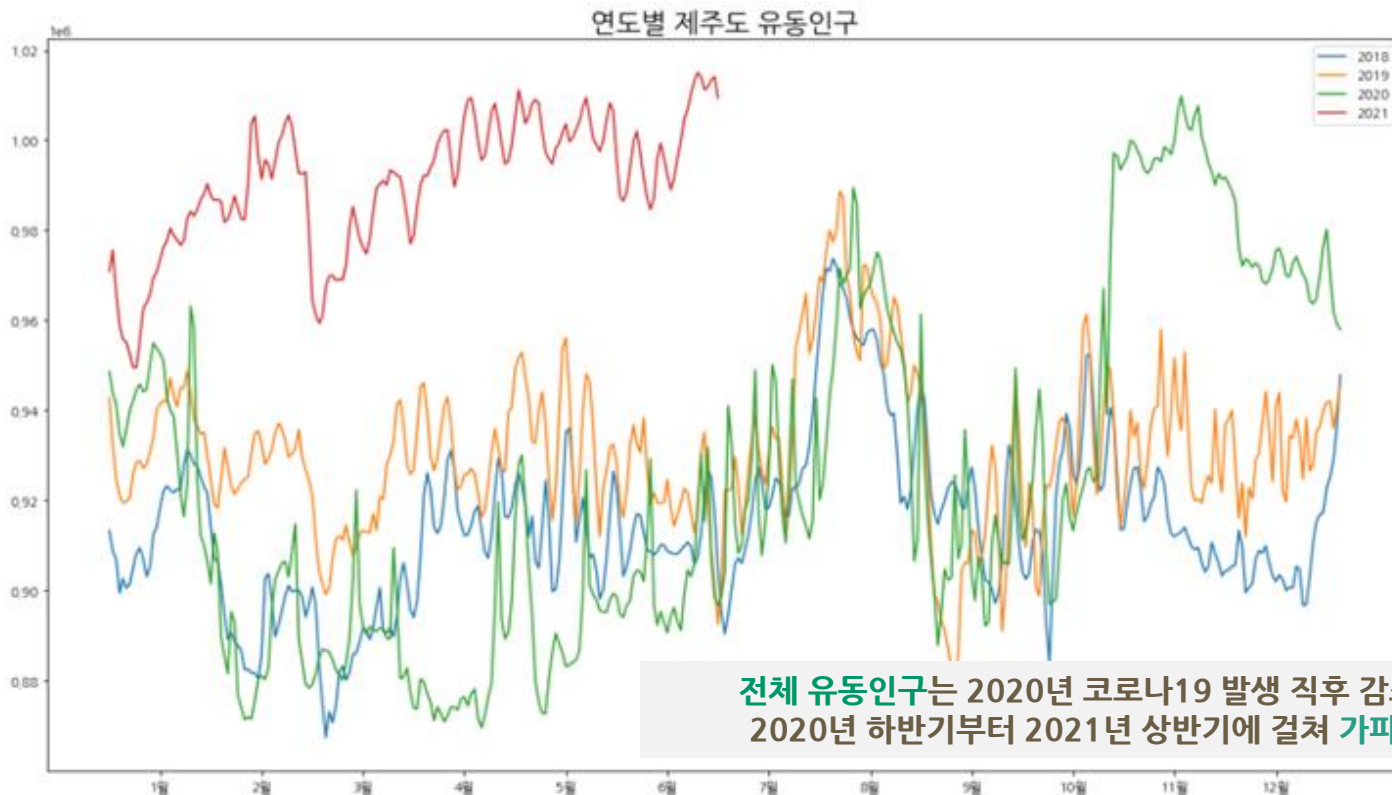
2-1. Feature ① : 거주인구

모든 행정구역에서 월별 거주인구 대부분은 **내국인 거주인구**
음식물 쓰레기 배출량과 **0.89의 높은 양(+)**의 상관관계를 보여 주요요인으로 고려



2. EDA

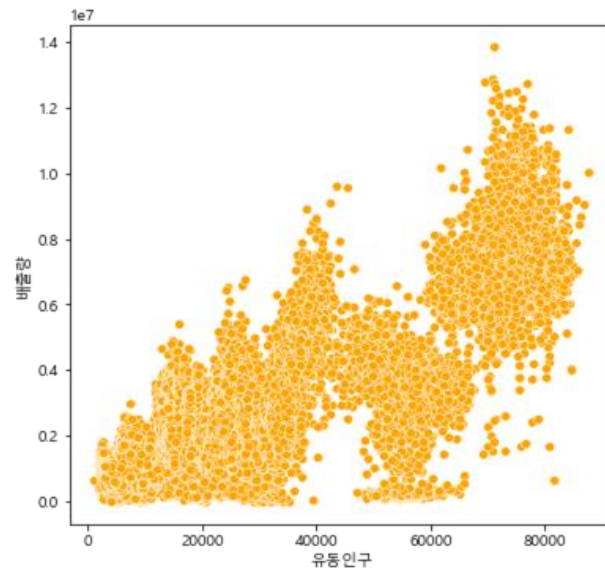
2-2. Feature ② : 유동인구



2. EDA

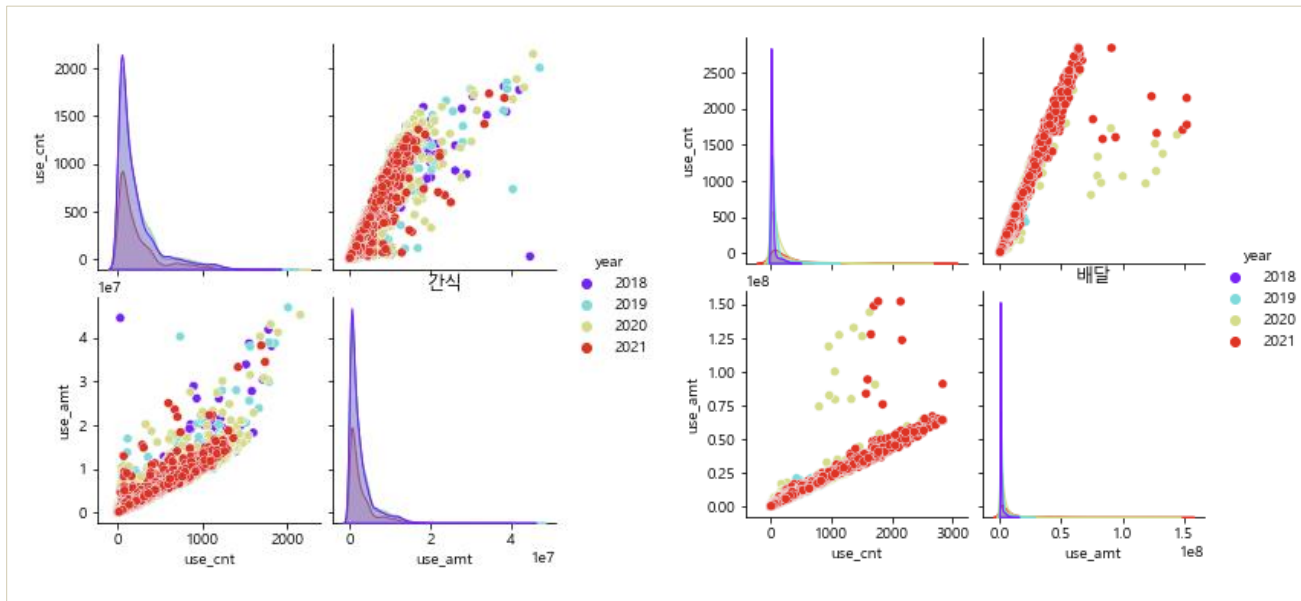
2-2. Feature ② : 유동인구

월 단위 유동인구와 내국인 거주인구 사이에 매우 높은 양(+)의 상관관계 존재
음식물 쓰레기 배출량은 유동인구에 따라 대체로 증가하는 것으로 보임



2. EDA

2-3. Feature ③ : 음식 관련 카드 소비

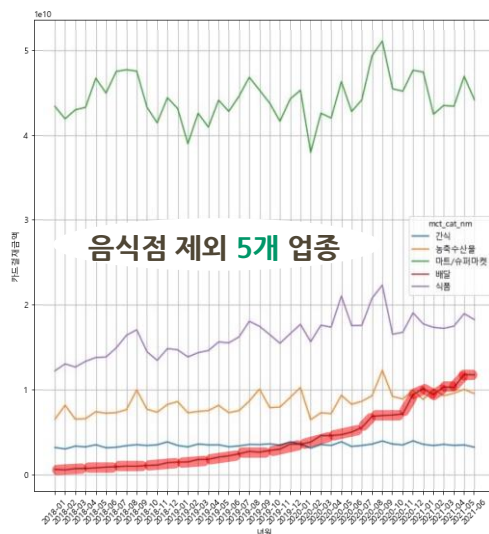
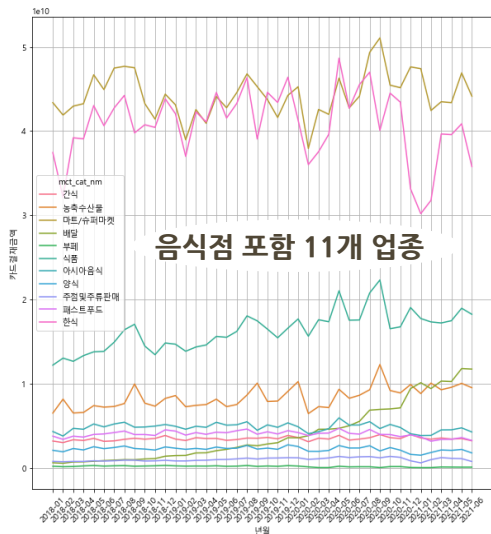


업종별 결제건수와 결제금액 사이의 양(+)의 상관성이 높게 나오므로
분석에는 결제금액만을 사용

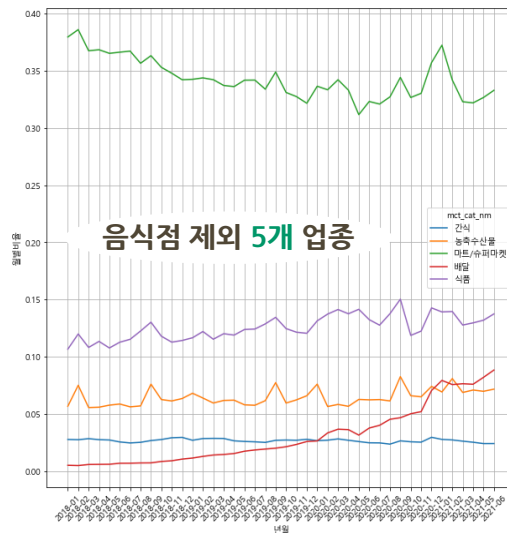
2. EDA

2-3. Feature ③ : 업종별 음식 관련 카드 소비

업종별 카드 결제금액



결제금액 비율

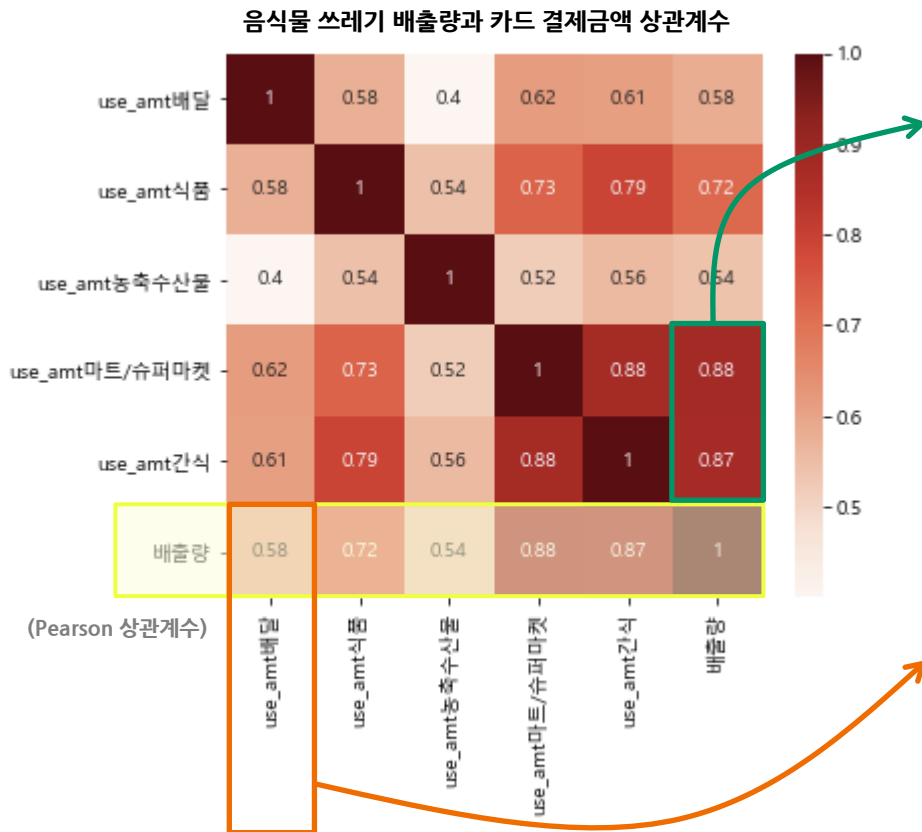


음식점을 제외한 카테고리 중 ‘마트/슈퍼마켓’의 결제금액이 가장 컸으며

시간이 흐름에 따라 ‘배달’ 결제금액의 상승세가 뚜렷함

2. EDA

2-3. Feature ③ : 음식 관련 카드 소비



음식물 쓰레기 배출량과
마트/슈퍼마켓 및 간식 결제금액
높은 상관관계



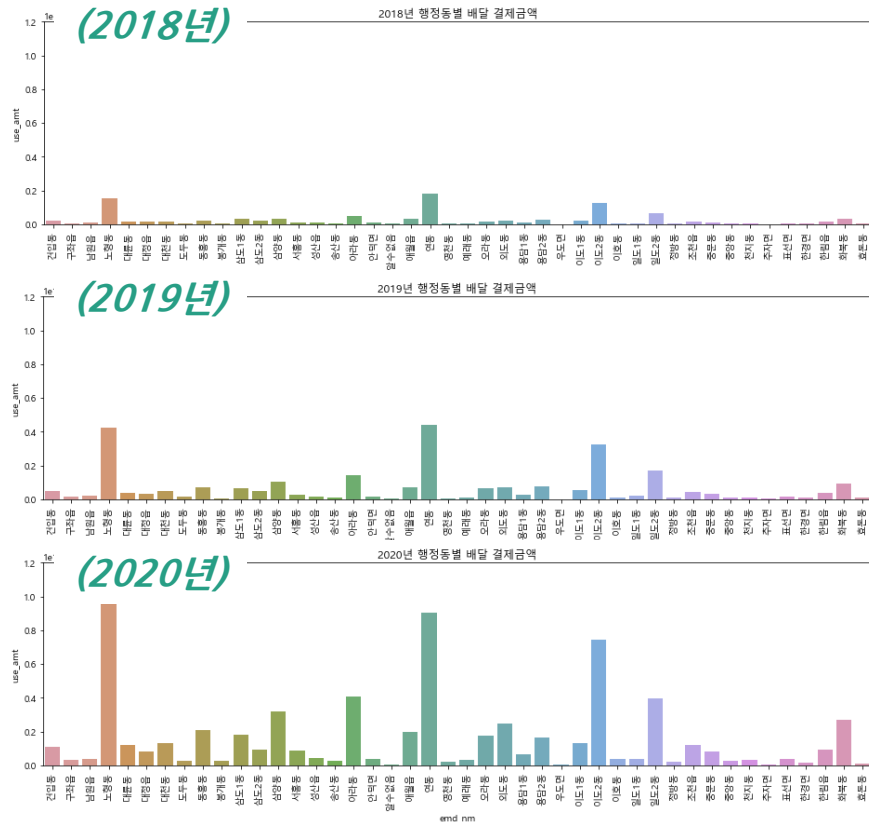
배출량 예측에
높은 순위 *Feature Importance* 예상

2018. 01~2021. 06 전체 기간 동안

음식물 쓰레기 배출량과 배달 결제금액
상관관계는 다른 특성들에 비해서 **미미**

2. EDA

2-3. Feature ③ : 음식 관련 카드 소비 - 배달



“배달 결제금액”

매년 약 2배 증가

2021년 상반기
→ 2020년 전체와 비슷한 수준

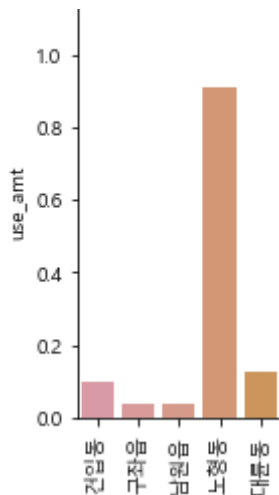


지속적인 배달 증가 추세



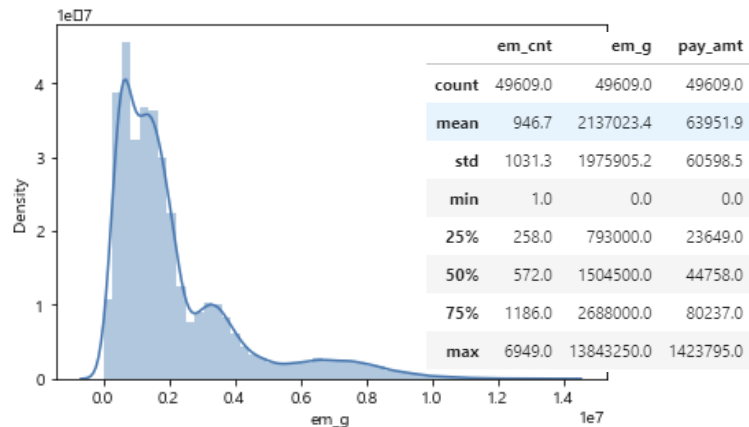
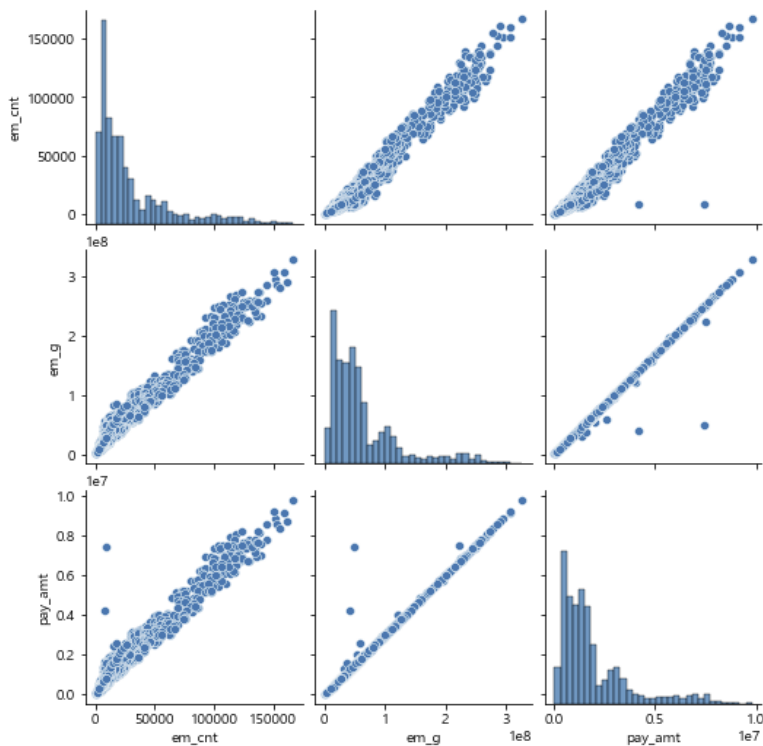
음식물 쓰레기
배출량 증가 예상

(2021년 상반기)



2. EDA

2-4. Target : 음식물 쓰레기 배출량



제주도 일별 읍면동별 배출량

1. 결측치 제거 - '알수없음' 지역 제거
2. 배출량(em_g)만을 feature로 사용
결제 금액(pay_amt) 상관 계수 = 0.99
3. 동일 행정구에 속하는 배출지점을 그룹화

2. EDA

2-4. Target : 제주도 전체 월별 배출량 증가 추세

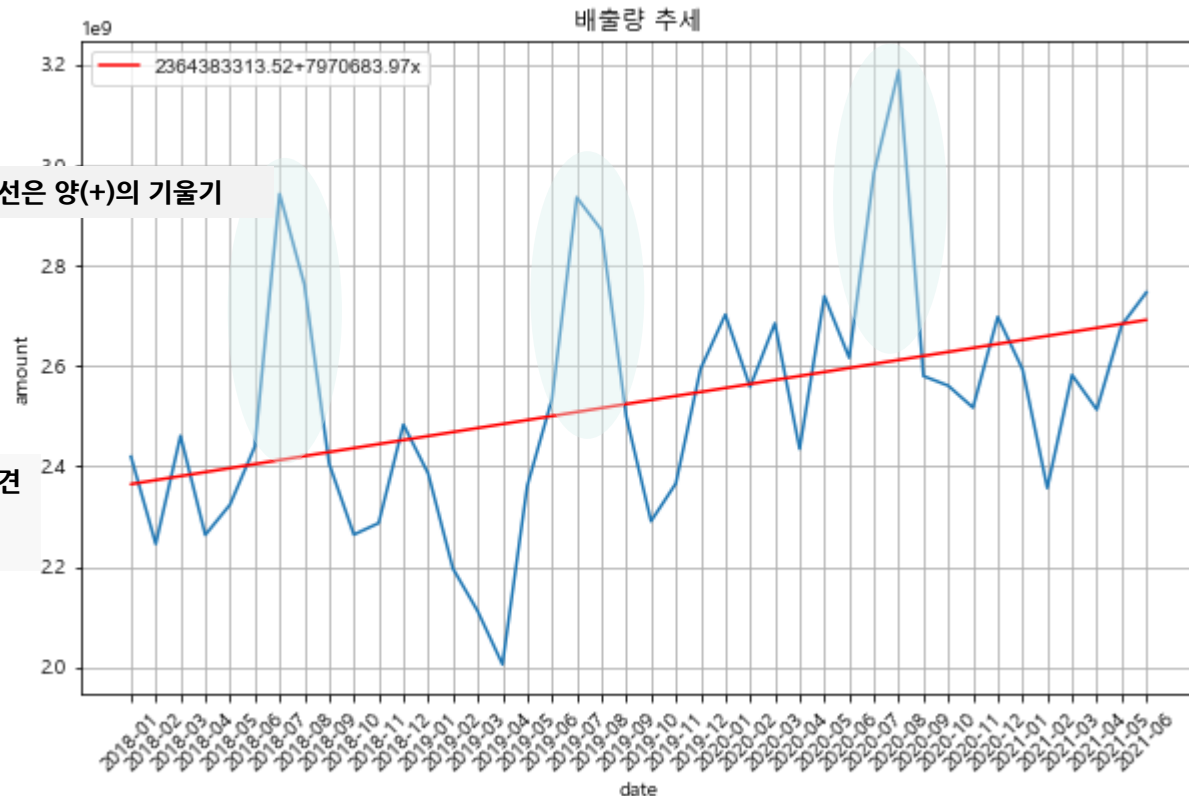
시간에 따른 월별 음식물 쓰레기 배출 회귀직선은 양(+)의 기울기



향후 음식물 쓰레기 배출량 증가 예상

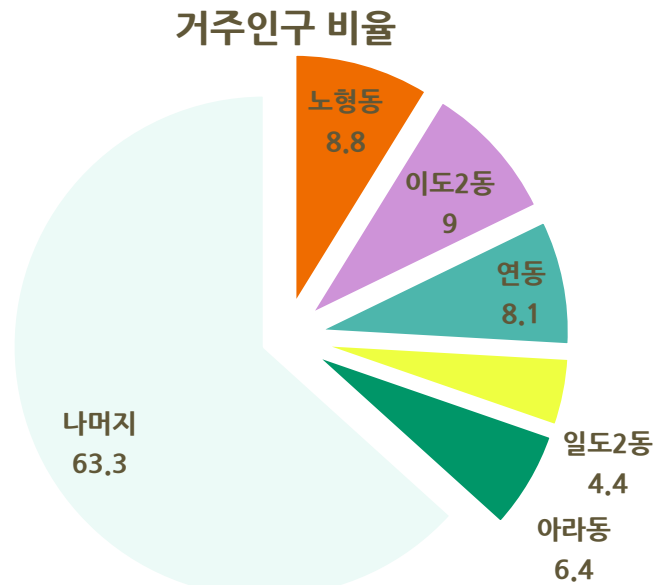
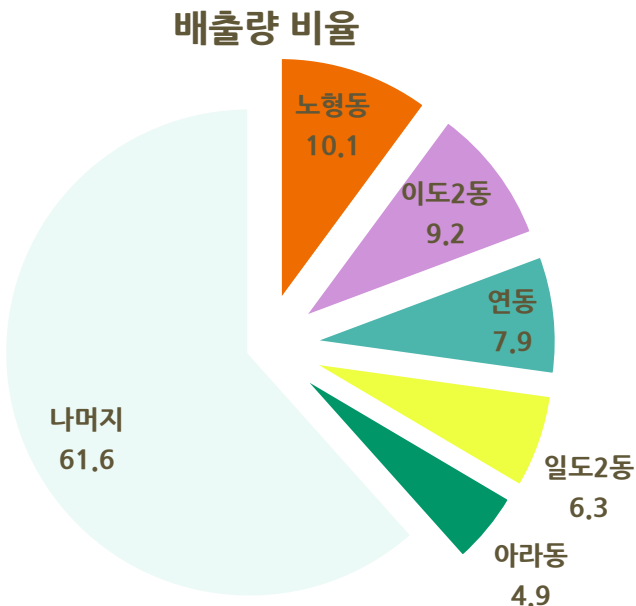
전체적인 배출량은 연도에 따라 상승

BUT, 매년 6-7월 배출량 급증하는 패턴을 발견
배출량 추세에 영향을 미치는 계절적 요인
(ex. 날씨 - 기압, 온도, 습도)



2. EDA

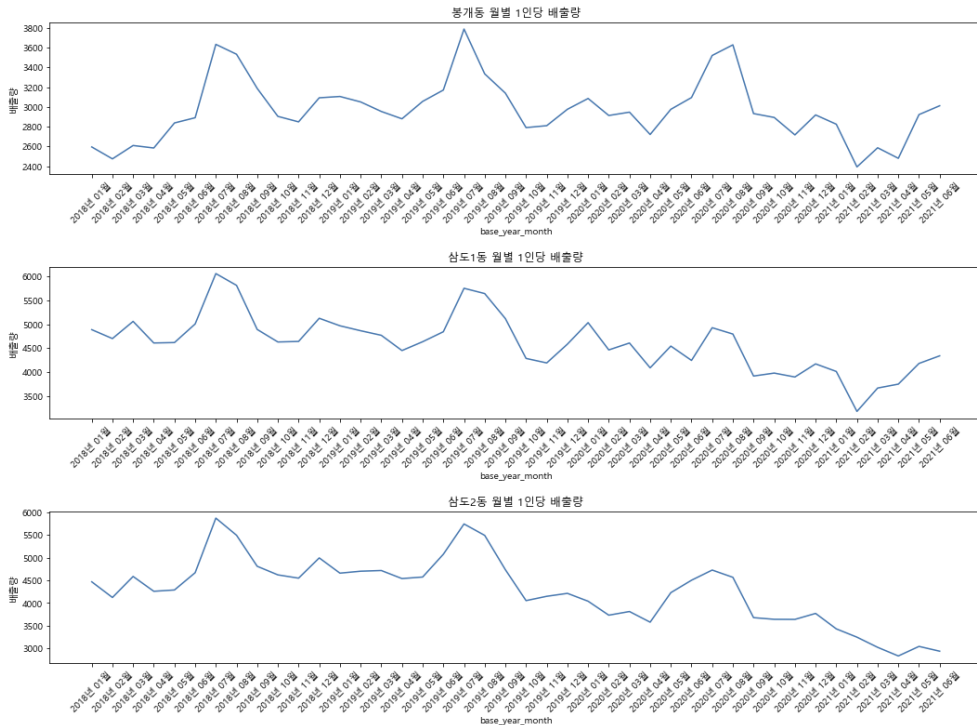
2-4. Target : 행정동별 음식물 쓰레기 배출 및 거주인구 비율



행정동별 음식물 쓰레기 배출량 비율은 거주인구 비율에 큰 영향을 받는 것으로 보임

2. EDA

2-4. Target : 거주 1인당 음식물 쓰레기 배출



월별 음식물 쓰레기 배출량을
행정동별 거주인구로 나누어,

행정동별 1인당 음식물 쓰레기
배출량 변동을 확인

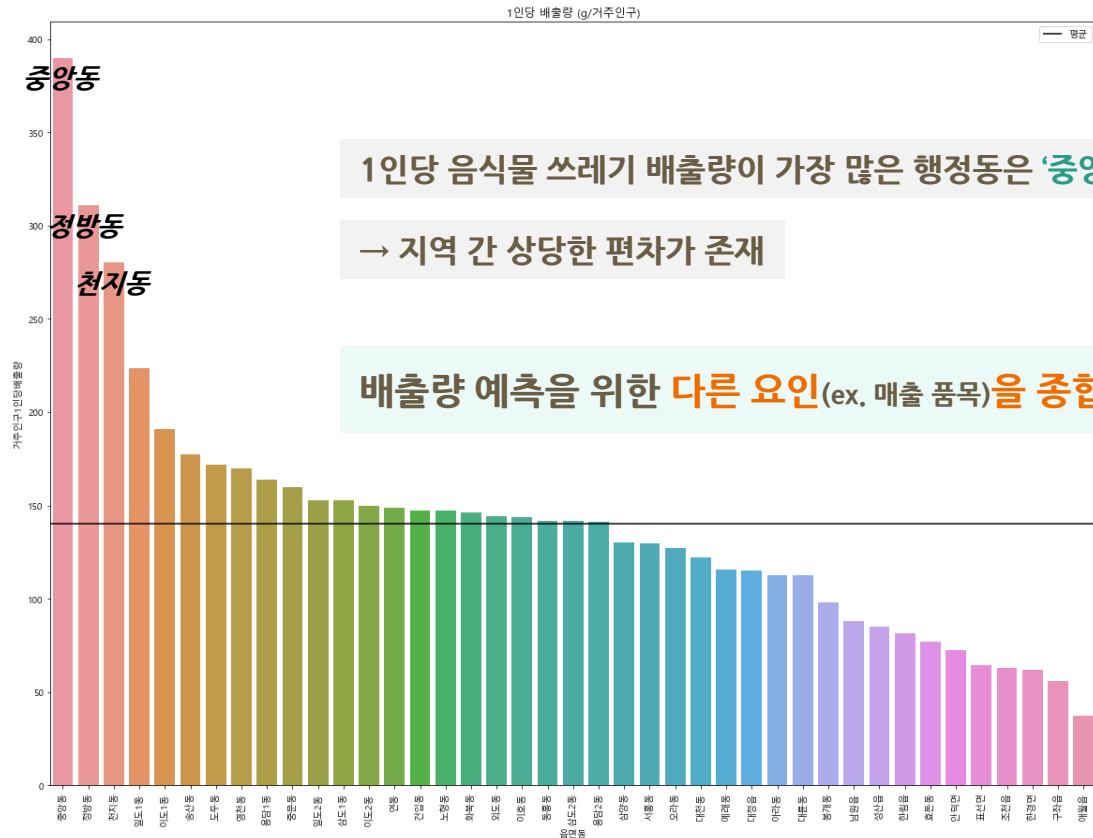
→ 대부분의 행정동에서 6, 7월
배출량이 높게 나타남



계절성을 반영하기 위해
날씨 데이터 추가

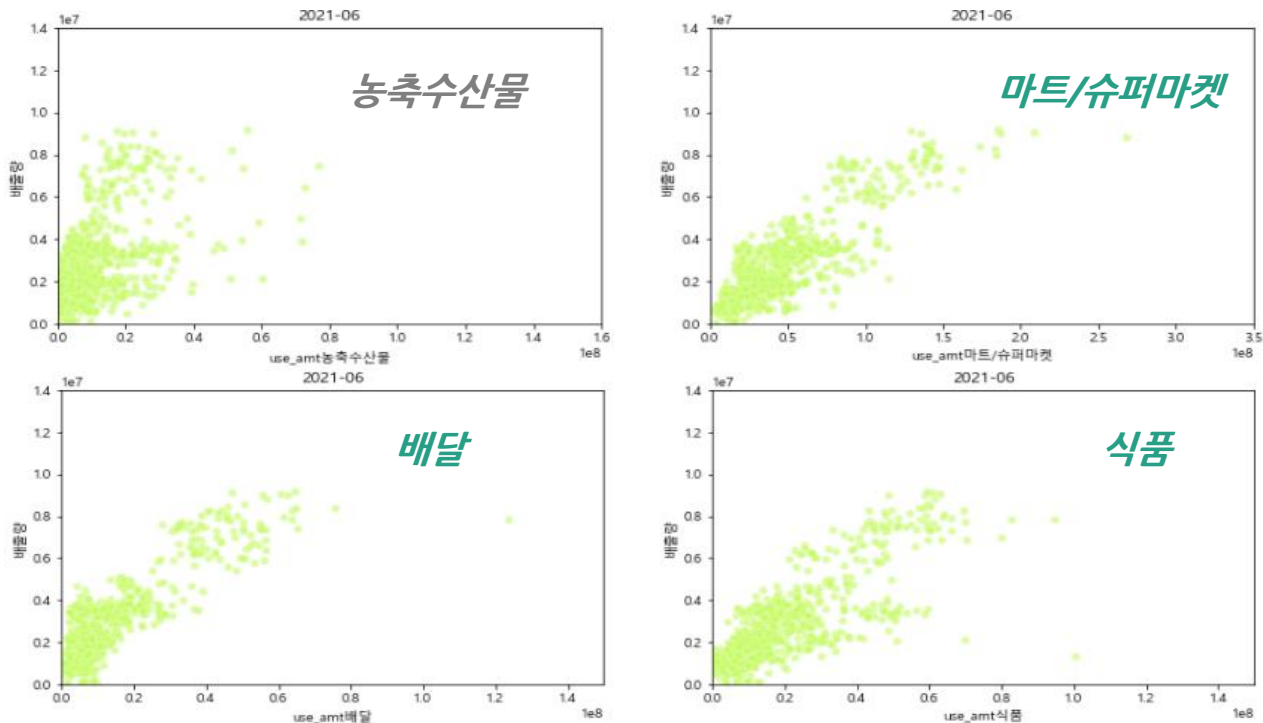
2. EDA

2-4. Target : 1인당 음식물 쓰레기 배출량 행정동 순위



2. EDA

2-4. Target : 음식 관련 소비 품목별 매출량 산포도



2021년 6월(최근 데이터) 기준 소비 품목에 따른 매출량 분포에 차이가 존재

데이터 분석 모델



2. 데이터 전처리

3-1. 데이터 통합

유동 인구

emd_nm	sex	age	resd_pop_cnt	work_pop_cnt	visit_pop_cnt
한림읍	여성	0	638.5509	15.5400	151.5149
한림읍	여성	0	0.0000	3.7249	171.3458
한림읍	여성	10	668.4367	17.0308	166.0500
한림읍	여성	10	0.0000	4.2577	195.8539
한림읍	여성	20	609.7710	5.9490	148.7246

카드 매출

emd_nm	mct_cat_cd	mct_cat_nm	use_cnt	use_amt
남원읍	C00400	간식	108	1565308
남원읍	C01200	농축수산물	54	2199680
남원읍	A00200	마트/슈퍼마켓	466	24828201
남원읍	C01400	배달	25	644192
남원읍	C00700	부패	5	294600

거주 인구

emd_nm	sex	resid_reg_pop	foreign_pop	total_pop
건일동	남성	5085	146	5231
건일동	여성	4715	82	4797
구좌읍	남성	7905	368	8333
구좌읍	여성	7609	184	7793
남원읍	남성	9806	428	10234

음식물 쓰레기

emd_nm	em_area_cd	em_cnt	em_g	pay_amt
외도동	W6X062	2	15500	464
아라동	W6XA87	25	59800	1787
아라동	W6XA96	8	25350	758
아라동	W6XA95	12	30000	898
아라동	W6XA94	14	44050	1317

	base_date	city	emd_nm	0여성 notvisit	0여성visit	0남성 notvisit
0	2018-01-01	서귀포시	남원읍	11030.2331	9791.3235	11498.2617
1	2018-01-01	서귀포시	대륜동	17647.5285	10160.5833	15216.9486
2	2018-01-01	서귀포시	대정읍	12538.2881	7792.5662	14360.4369
3	2018-01-01	서귀포시	대천동	10539.1926	7307.2766	10214.4234
4	2018-01-01	서귀포시	충흥동	20141.3526	7741.1595	22060.7555

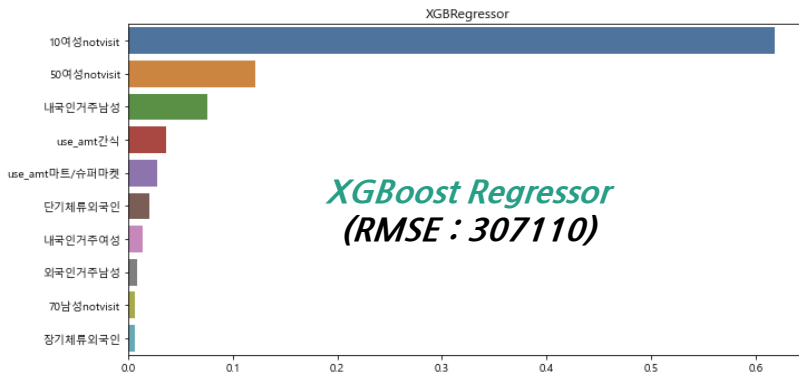
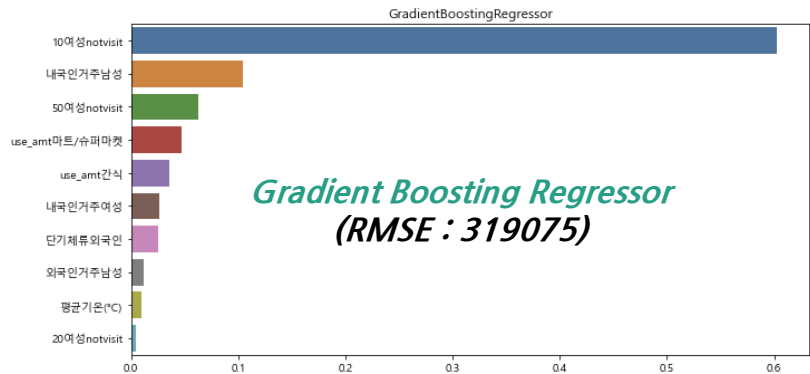
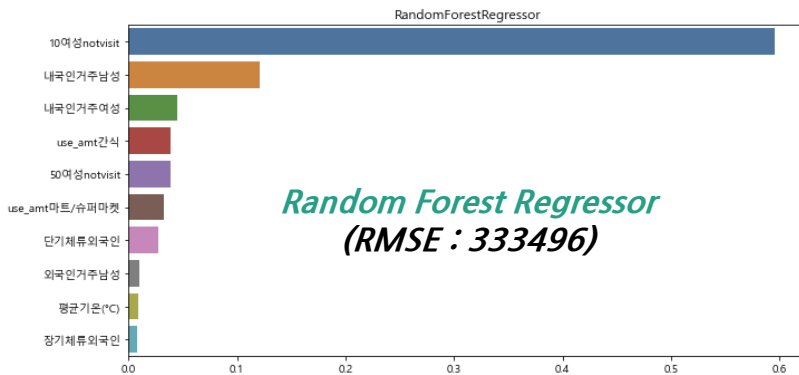
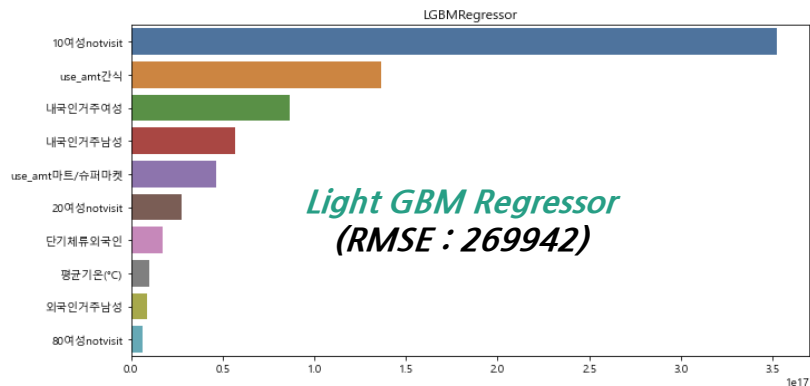
use_amt	식품	농축수산물	마트/슈퍼마켓
16392477		2199680	24828201
3178292		354993	37971731
1413442		2730942	37142481
2611973		687498	39231391
2994364		2572840	69274797

평균 전운량 (1/10)	내국인거 주남성	외국인거 주남성	내국인거 주여성	외국인거 주여성	배출량
0.0	9806	428	9306	200	1239600
0.0	6836	83	6637	95	1717700
0.0	10360	841	10725	677	1264950
0.0	6685	107	6475	137	1765600
0.0	11124	624	11569	642	4165050

일/행정동별 데이터로 통합 총 48,421 rows × 59 columns (Features : 55 columns)

2. 데이터 분석

3-2. 데이터 분석 : 전체 Features 모델 간 성능 비교



* 하루 평균 행정동별 음식물 쓰레기 배출량 : 약 2톤

2. 데이터 분석

3-2. 데이터 분석 : LGBM Feature 조합 성능 비교

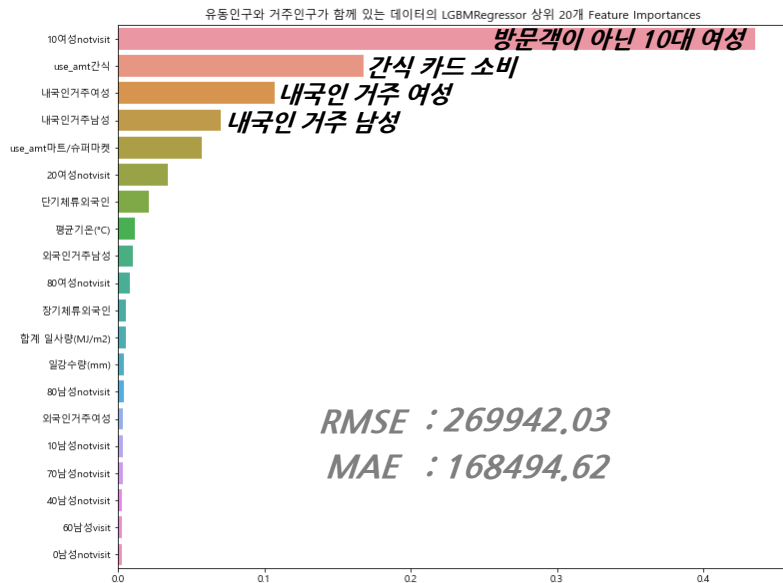
	① 전체	② 인구만	② 인구 제외	③ 유동인구 제외	④ 거주인구 제외	⑤ 상호작용 (배달, 간식)
RMSE	269942	380231	520627	255081	282677	255469
MAE	168494	221241	355562	163309	179068	163379
Feature 1st	방문객이 아닌 10대 여성	방문객이 아닌 10대 여성	간식 매출	간식 매출	방문객이 아닌 10대 여성	간식 매출
Feature 2nd	간식 매출	방문객이 아닌 20대 여성	마트/슈퍼마켓 매출	내국인 거주 남성	간식 매출	내국인 거주 여성
Feature 3rd	내국인 거주 여성	내국인 거주 여성	식품 매출	내국인 거주 여성	마트/슈퍼마켓 매출	내국인 거주 남성

* 상위 1-3위 Features : Information Gain 값이 높은 순서

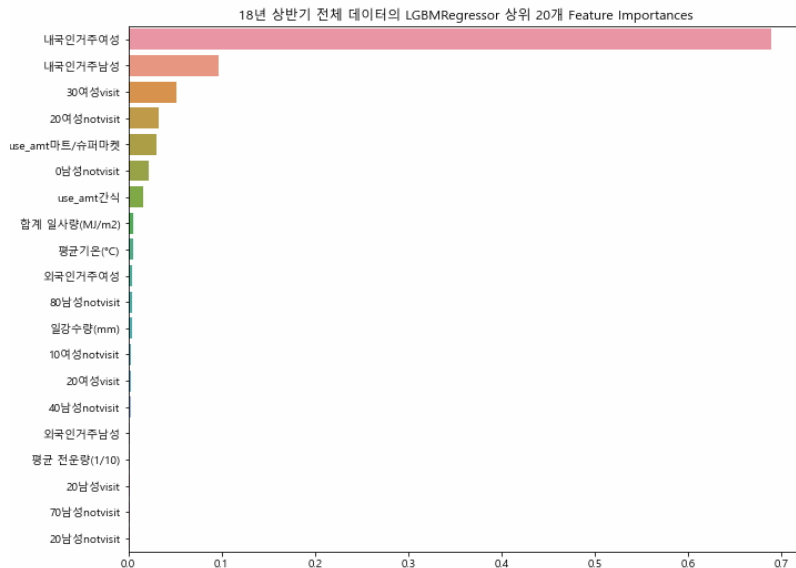
2. 데이터 분석

3-2. 데이터 분석 ① : LGBM 전체 Features

전체 Features



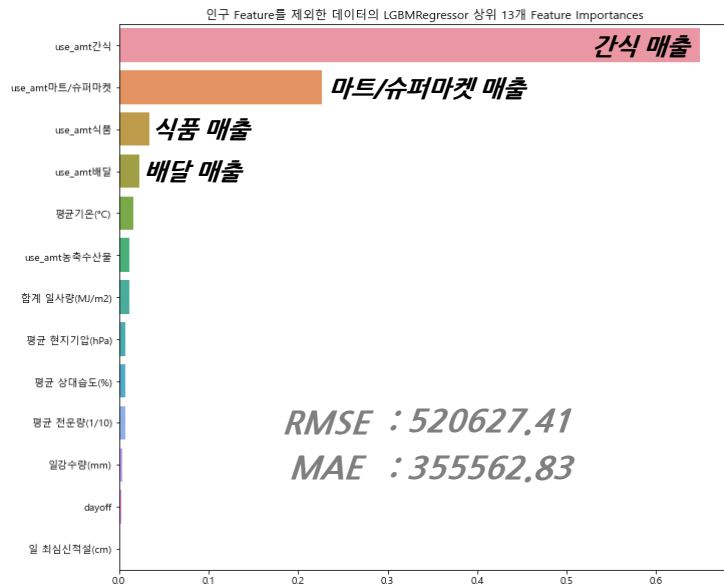
2018 - 2021년 상·하반기



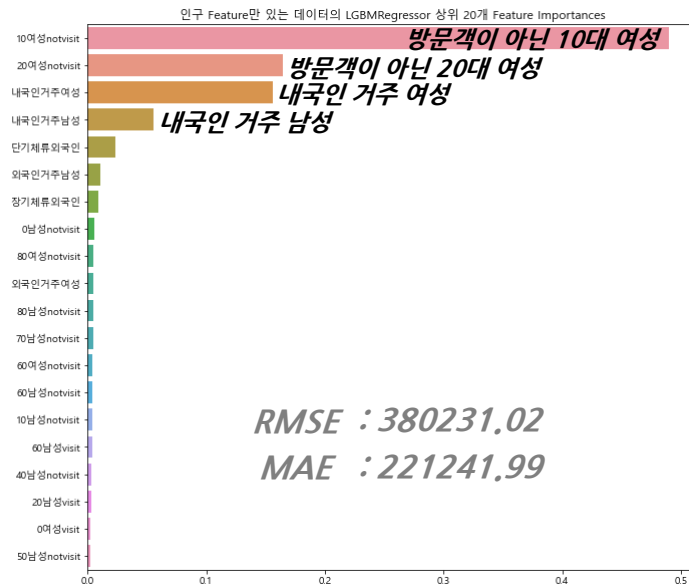
2. 데이터 분석

3-2. 데이터 분석 ② : LGBM 인구 Features

인구 요소를 제외한 데이터



인구 요소만으로 이루어진 데이터



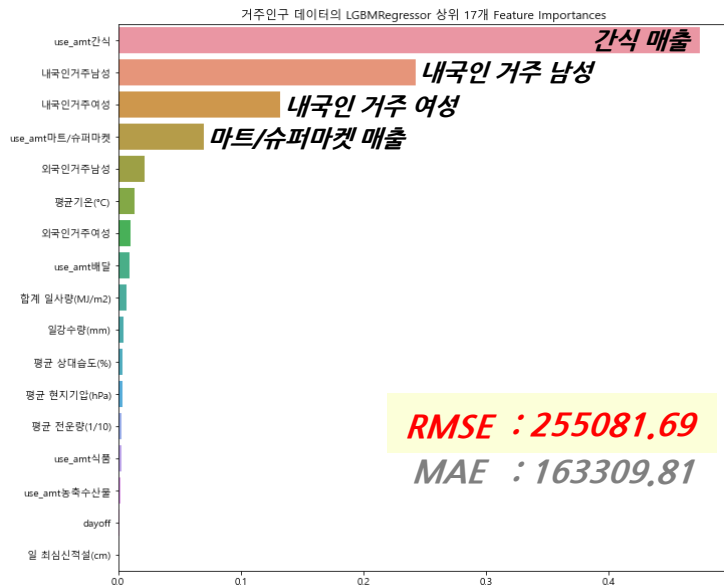
전체 데이터

RMSE : 269942.03
MAE : 168494.62

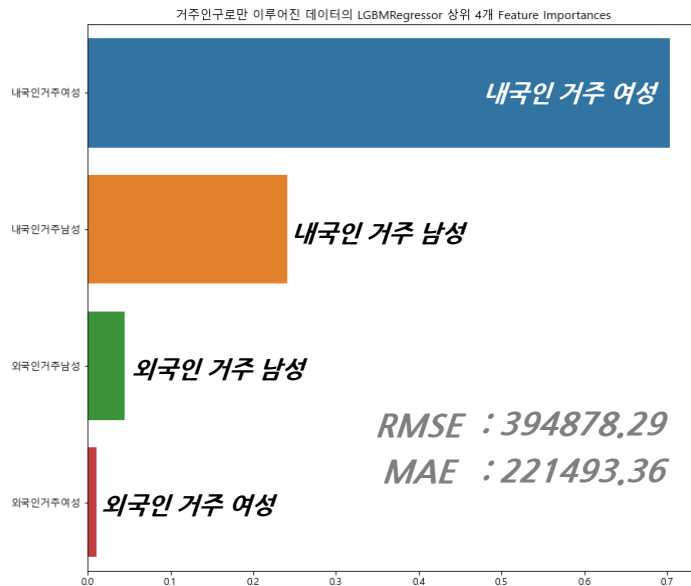
2. 데이터 분석

3-2. 데이터 분석 ③ : LGBM 유동인구 제외한 Features

유동인구를 제외한 데이터



거주인구만으로 이루어진 데이터



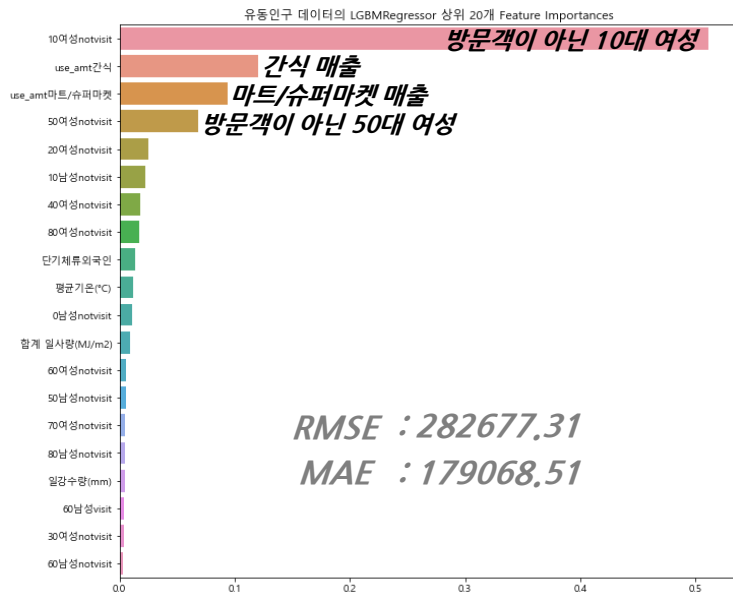
전체 데이터

RMSE : 269942.03
MAE : 168494.62

2. 데이터 분석

3-2. 데이터 분석 ④ : LGBM 거주인구 제외한 Features

거주인구를 제외한 데이터



유동인구만으로 이루어진 데이터



전체 데이터

RMSE : 269942.03
MAE : 168494.62

2. 데이터 분석

3-2. 데이터 분석 ⑤ : LGBM Polynomial Features • 거주인구 조합

상호작용 효과

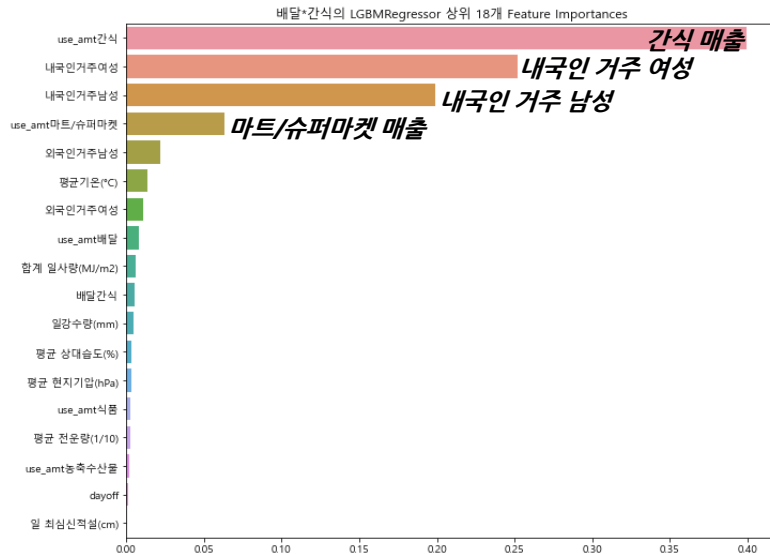
	배달 * 식품	배달 * 간식	배달 * 마트
RMSE	259256	255469	257168
MAE	164965	163379	164121

	식품 * 간식	식품 * 마트	마트 * 간식
RMSE	256934	261425	259725
MAE	164315	166673	164959

전체 데이터
RMSE : 269942.03
MAE : 168494.62

거주인구 데이터
RMSE : 255081.69
MAE : 168494.62

배달 • 간식 LGBM



3. 데이터 분석 결과

1. 인사이트 도출

2. 문제 현황 분석

3. 결론 및 제언

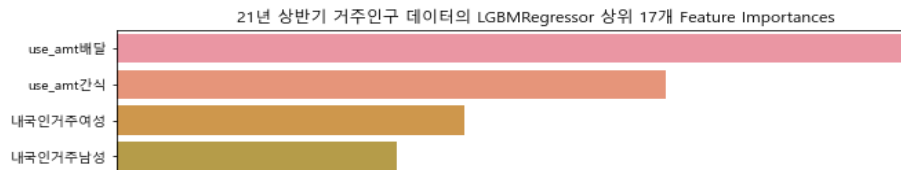
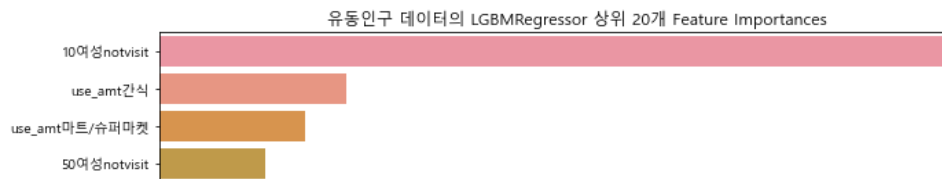
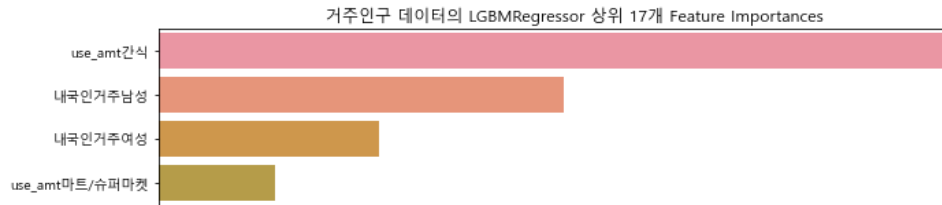
3. 인사이트 도출

1-1. 전략 목표 대상 추출을 위한 주요 요인 분석

(1) 거주인구를 기준으로 분석	
FI	간식 매출, 내국인 거주 남성
Corr	0.893
간식을 즐겨 먹는 내국인 남성	

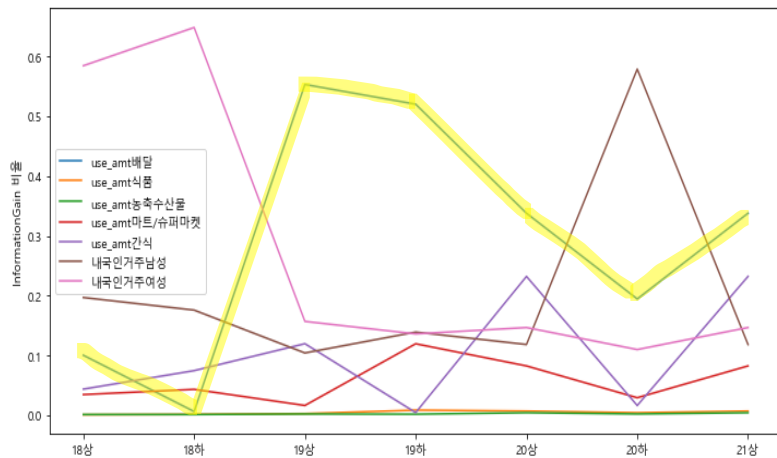
(2) 유동인구를 기준으로 분석	
FI	방문객이 아닌 10대 여성, 간식 매출
Corr	0.891
간식을 즐겨 먹는 내국인 10대 여성	

(3) 21년 상반기를 기준으로 분석	
FI	배달 매출, 간식 매출
Corr	0.612
모바일 앱으로 배달을 시키는 내국인	

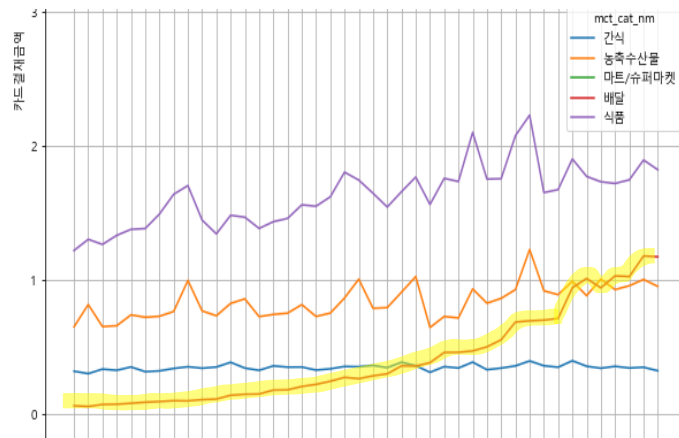


3. 인사이트 도출

1-2. 제주도 음식물 쓰레기 증가에 배달이 미치는 영향



소비 카테고리 중 **배달**이 가장 큰 영향



매월 **배달 지출 증가** 추세

배달의 영향력이 증대될 것으로 예상

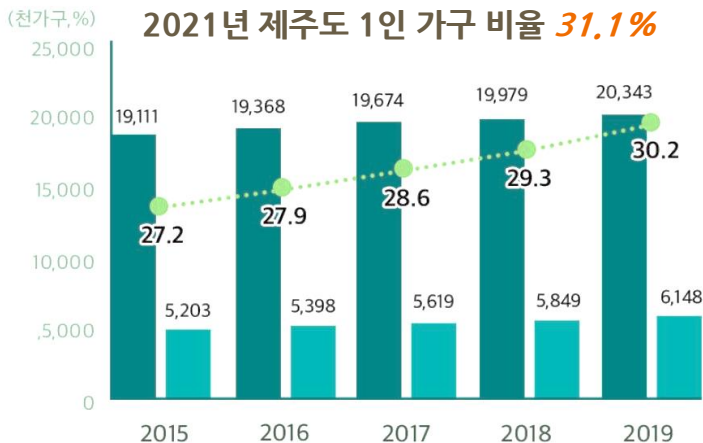
배달 업종 전반적으로 음식물 쓰레기 배출을 감소시킬 수 있는 방안 논의 필요

3. 문제 현황 분석

2-1. 1인 가구의 증가 및 외부 요인으로 인한 배달 증가

1인 가구 증가 추세

배달 : 1인 가구 63.6% > 3인 57.8%



코로나19 영향 · 배달 인프라 확대

코로나 이후 **배달음식 소비 33.4% ↑** ...음식물 쓰레기도 증가



등록 2020-12-19 05:00:00

제주, 도서·산간에 **드론 배달 인프라 구축**

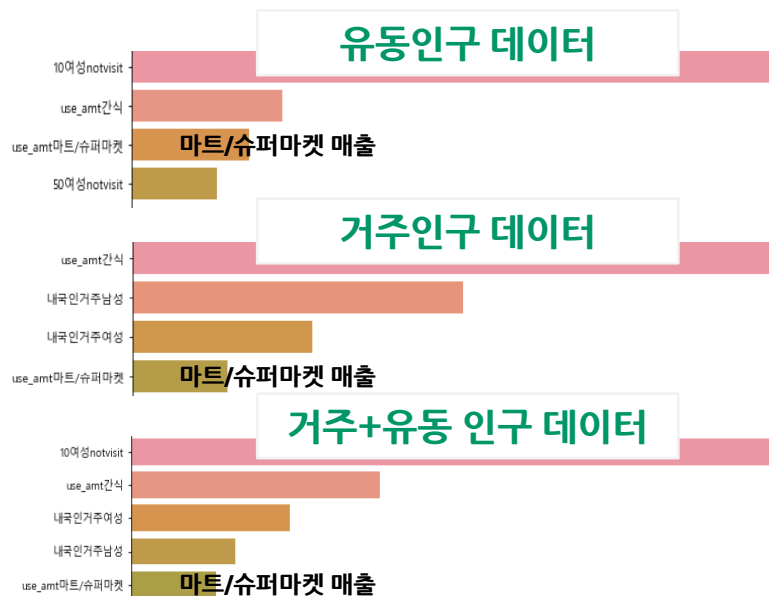
박미라 기자 입력 : 2021.05.26 22:02 | 수정 : 2021.05.26 22:02



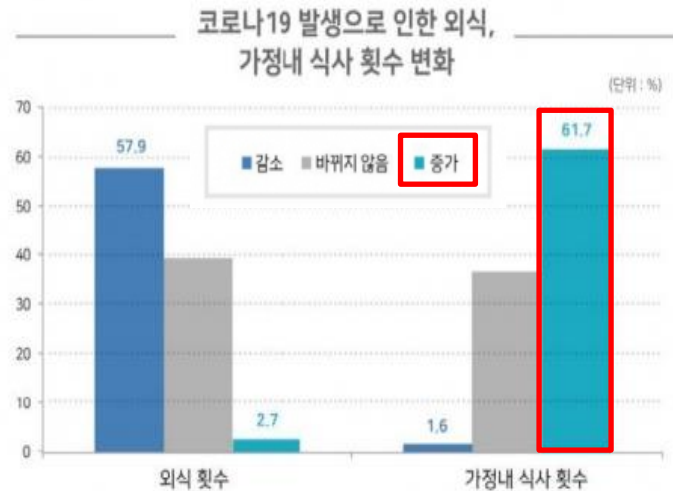
급증하는 배달량으로 인한 제주지역의
“**음식물 쓰레기 배출량 증가 예상**”

3. 문제 현황 분석

2-2. 코로나19 이후 마트/슈퍼마켓 이용으로 인한 음식물 쓰레기 증가



코로나19와 식품소비행태 변화



구현한 모든 모델에서 **마트/슈퍼마켓 매출**은 음식물 쓰레기 배출 요인 상위
코로나19 이후 가정식사의 비중이 높아진 상황

3. 결론 및 제언

3-1. 마트 소포장 - 1인 가구와 50대 여성을 중심으로



가정계 폐기물 발생 실태조사 연구(전태완 외, 2012) - 국내 100가구 모니터링 결과

**“폐기되는 음식물을 줄이는 효과적인 방법은
1~2인 가구를 위한 소포장 제품 판매가
38%로 가장 많았다.”**

“냉장고 내·외 보관량으로는 평균 55.0 kg의 음식물을 보관하고 있으며,
이 중에서 6.0%가 버려지는 것으로 나타났다.”

유동인구 기준 Feature Importance 3위 - 마트, 4위- 50대 여성 페르소나를 가정 :

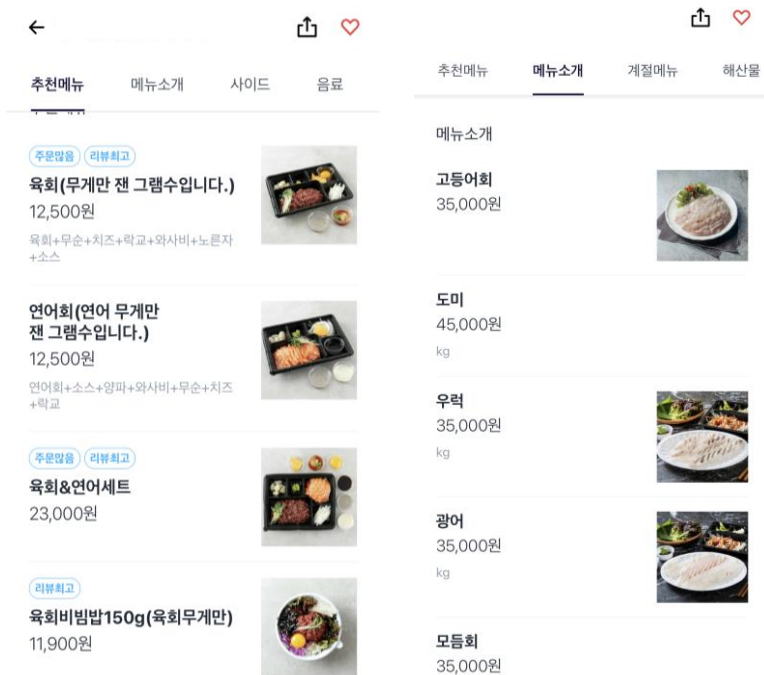
“제주도 노형동에 거주하는 50대 가정주부 김하르방씨는 코로나로 인해 다같이 함께하는 가정 식사를 준비하면서 10대 여성 딸과 함께 마트에 방문하였다.

외식을 대체할 평소 집에서 자주 해 먹지 않는 식자재를 고르다보니 다양한 종류의 재료를
구비하려 한다. 대부분 대용량의 식자재였지만, 나중에 먹을 것이라는 막연한 확신을 가지고
모두 구매하였다.”

- 남은 음식물은 냉장고에 오래 보관되었다가 버려졌다.

3. 결론 및 제언

3-2. 배달 앱 1인분 서비스 추가 확대 - 노형동을 중심으로



서울시 신림동

노형동

(제주도 배달량 & 음식물 쓰레기 배출량 1위)



그러나 '1인분 서비스'의 최소 주문금액은 1만 원 이하
이며 먹고 싶은 만큼만 주문할 수 있습니다.

1인분 주문 수요가 늘어나면서 가맹업체들의 수익 증가를
기대해 볼만 합니다. 실제로 서비스를 도입한 가맹업체들의
주문 건수는 도입 이전보다 **평균 40% 이상 증가**한 것으로
나타났습니다.

“추가적인 비즈니스 가치 창출”

4. 후기 및 느낀점

윤보람	데이터를 가공하고, EDA와 머신러닝 분석을 통해 인사이트를 도출해보는 첫 프로젝트라서 굉장히 뜻깊었던 시간이었다. 부족한 부분도 많았지만 팀원들과 서로 도와가며 어려움을 헤쳐나갈 수 있었고, 모두들 프로젝트에 적극적으로 참여하는 모습에서 힘을 얻을 수 있었다. 누가 더 열심히 했다고 할 수 없을 정도로 팀워크가 너무 좋아서 팀원들에게 고마웠고, 아마 가장 기억에 남는 프로젝트가 될 것 같다. 분석을 하면서 실력이 조금씩 느는 것을 느꼈고, 앞으로도 꾸준히 시간을 들여서 배우고 익히는 시간을 가져야 할 것 같다.
김형림	다양한 걸 시도해볼 수 있어서 좋았다. 2주 동안 팀원들끼리 서로 모르는 것을 알려주고, 배울 수 있었다. 프로젝트를 할 때마다 항상 아직도 많이 부족하다는 생각을 하는데, 이번 프로젝트를 진행하면서 기초적인 부분부터 하나씩 홀로 또는 팀원들과 함께 도전하다보니 스스로도 실력이 조금은 향상된 것 같아서 뿌듯하다. 한편으로는 수업을 듣는 게 전부가 아니라 복습의 중요성과 다양한 경험을 쌓아야겠다는 것을 뼈저리게 느끼기도 했다. 이를 계기로 앞으로도 열심히 공부해야지!
박형준	모두가 공감 할 수 있는 음식물 쓰레기 문제에 대해서 실전적으로 분석을 해 볼 수 있는 기회가 되어 큰 동기부여를 가지고 진행할 수 있었다. 또한, 기술적으로는 EDA를 하면서 데이터프레임을 유연하게 다루는 연습을 해볼 수 있었다. 한편, 인사이트를 발굴해보는 과정에서, 직관과는 다른 결과를 도출하며 전략 구상에 있어 데이터 분석의 필요성을 더욱 체감하게 된 계기가 되었다
장우창	Feature를 넣고 뺄에 따라 성능이 달라진다는 것을 이론적으로 알고 있었으나 이번 프로젝트를 통해 직접 겪어보니 feature 선택의 중요성을 확실히 알게 되어 더 공부해야할 필요성을 느꼈습니다.

Q & A

— 경청해 주셔서 감사합니다. —
