날씨에 따른 소비패턴 변화를 기반으로 한 상품 추천

 TEAM
 ELF

 박 주 호

 이 재 범

CONTENTS

01

02

03

04

배경

- 가설 설정
- 마케팅 대상
- 마케팅 목표

EDA

- 날씨 데이터 병합
- 이상치 제거
- 날씨별 차이
- 컬럼 선정

모델

- 사용 모델
- 결과

결론

- 활용 방안
- 개선 방안

" 날씨와 소비?"

한국은 뚜렷한 사계절을 통해 다양한 날씨가 존재 기사를 보면 날씨별로 엄연히 소비가 다르며, 이를 겨냥한 다양한 마케팅이 있음

이런 점을 근거로 날씨 별 소비 패턴을 분석, 개인에게 제품을 추천해주는 개인화 마케팅 전략을 세움

올해 들어 서울 가장 더웠다...덥고 습한 날씨에 '불티'나는 이 것

때이른 더위에 냉방제품 판매량↑ 유통업계, 냉방제품 마케팅도 활발

배윤경 기자 입력: 2022.04.26 16:35:12 🖭 0



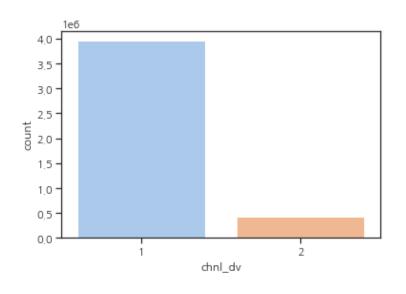
"장마철 필수템은?"···유통업계, 장마 마케팅 틈새시장 공략

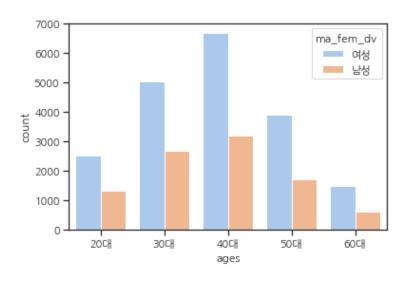
기사승인 2022-06-29 06:30:06



마케팅 대상 선정

해당 마케팅에 가장 적합한 대상을 선정, 그 데이터를 바탕으로 진행





- 1. 왼쪽 그래프를 보면 오프라인의 비율이 압도적으로 높은 것을 확인 할 수 있음
 - → 날씨는 오프라인에 영향을 주는 요소이기에 매우 적합
- 2. 오른쪽 그래프를 보면 40대의 lpay 가입이 가장 많음
- ⇒ 마케팅 대상으로 오프라인을 활동을 하는 40대가 가장 적합

마케팅 목표

해당 마케팅에 가장 적합한 대상을 선정, 그 데이터를 바탕으로 진행

Lpay 사용 촉진

추천 시스템을 이용한 마케팅으로 Ipay의 이용률을 높일 수 있음

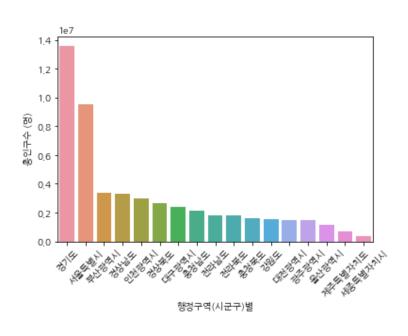
Lpay 가입 촉진

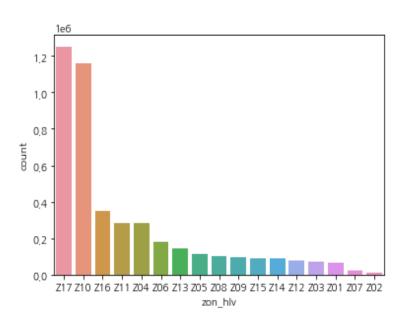
추천 시스템을 통해 비슷한 사용자를 찾고, 이를 기반으로 비슷한 사용자의 가입을 유도

마케팅 비용 감소

자사 앱을 통하여 홍보가 가능하기에 비용이 절감됨

날씨 데이터 병합





행정구역 별 인구 분포를 보면 오른쪽의 고객 데이터의 지역 분포와 상당히 흡사 한 것을 확인 할 수 있음

⇒ Z17과 Z10은 경기/서울로 추론 가능

날씨 데이터 병합

	-			L ATT	Б.,
Date	Day	ΑT	TT	MT	Rain
20210101	fri	-4.2	1.6	-9.8	0.0
20210102	sat	-5.0	-1.4	-8.4	0.0
20210103	sun	-5.6	-2.0	-9.1	0.0
20210104	mon	-3.5	0.3	-8.4	0.0
20210105	tue	-5.5	-2.1	-9.9	0.0
20210106	wed	-7.4	-1.9	-12.0	2.3
20210107	thu	-14.5	-8.4	-16.5	0.0
20210108	fri	-14.9	-10.7	-18.6	0.0
20210109	sat	-12.2	-7.5	-16.6	0.0
20210110	sun	-7.7	-2.7	-12.8	0.0
20210111	mon	-6.7	-4.0	-9.6	0.0
20210112	tue	-3.9	-0.5	-9.3	2.6
20210113	wed	2.0	8.3	-3.0	0.0
20210114	thu	1.7	7.0	-3.6	0.0
20210115	fri	4.5	9.9	-0.7	0.2
20210116	sat	-4.9	-0.7	-8.7	0.0
20210117	sun	-5.5	-1.3	-10.8	0.0
20210118	mon	-3.8	0.2	-9.1	0.3

		cust	rct_no	chnl_dv	cop_c	br_c	pd_c	de_dt	
weather	Rain_status								
cold	no_rain	1474711	1474711	1474711	1474711	1277901	1474711	1474711	1
	rain	116203	116203	116203	116203	100456	116203	116203	
hot	no_rain	509690	509690	509690	509690	489726	509690	509690	
	rain	79394	79394	79394	79394	76097	79394	79394	
warm	no_rain	1606194	1606194	1606194	1606194	1462093	1606194	1606194	1
	rain	595551	595551	595551	595551	546969	595551	595551	

1. 웹크롤링을 통한 2021년 서울/경기 날씨를 추출

자료: 기상청 날씨 누리 https://www.weather.go.kr

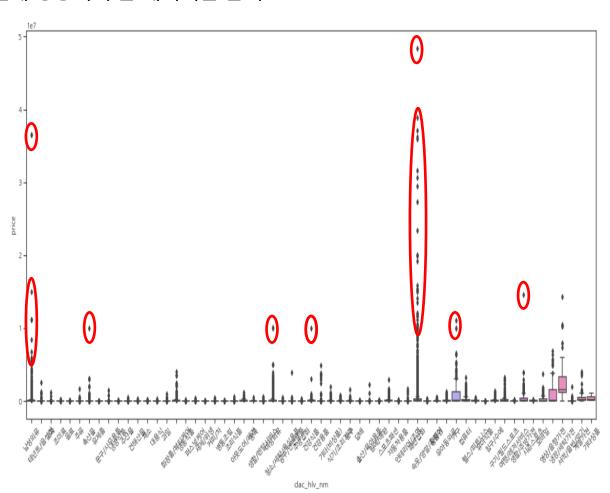
- 2. 평균기온이 10℃ 아래면 cold, 10℃에서 25℃ 사이면 warm, 25℃보다 높으면 hot
- 3. 해당 데이터를 일자에 맞게 병합

이상치 탐지

주제와 연관이 없지만 모델에 영향력이 큰 데이터를 탐지

	pd_nm	price
0	남성시계	48380000.0
1	반지	38870000.0
2	기타보석류	37070000.0
3	기타남성의류세트	36550000.0
4	기타패션잡화	30680000.0
1694	영유아레깅스	1000.0
1695	기타스키/보드용품	790.0
1696	봉투보증금	150.0
1697	영유아스커트	100.0
1698	여아한복	100.0

1699 rows × 2 columns

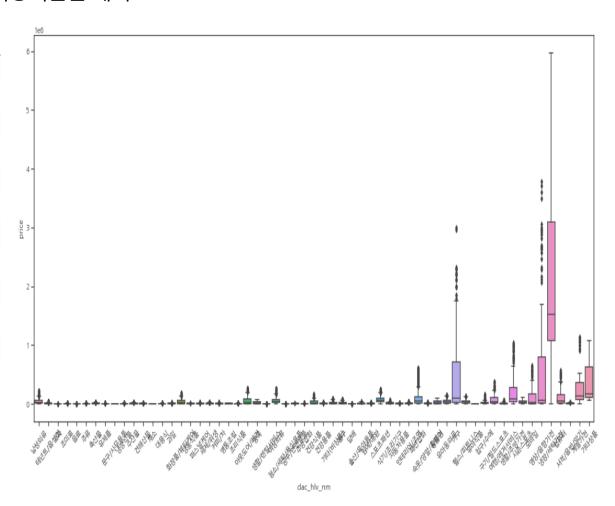


이상치 제거

일부 재화에서 발견 된 이상치들을 제거

	pd_nm	price
0	드럼세탁기	5972000.0
1	양문형냉장고	4822000.0
2	UHD	3785000.0
3	냉장/냉동가전소모품	3497000.0
4	스탠드형김치냉장고	3230000.0
1619	여아한복	100.0
1620	영유아스커트	100.0
1621	의류스타일러	100.0
1622	기타여아특수목적의류	100.0
1623	이사/청소/수선/세탁서비스	100.0

1624 rows × 2 columns



날씨별 차이

평균 기온

25℃ 이상 : 더움(hot)

10°C 에서 25°C 사이 : 따뜻함(warm) 2.5mm 이상 : 비가 옴(rain)

10℃ 이하 : 추움(cold)

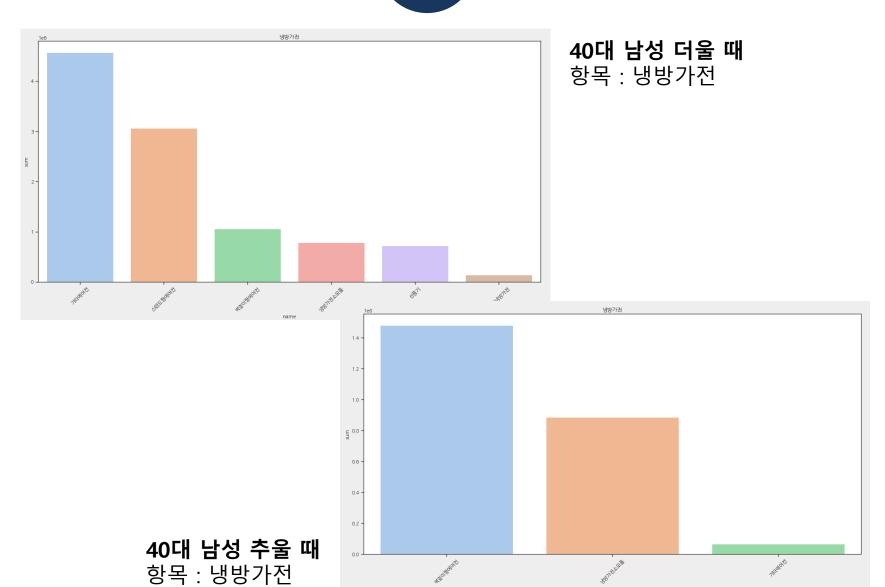
강수량

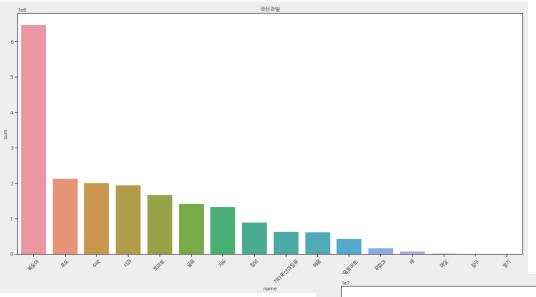
2.5mm 미만 : 비가 오지 않음(no_rain)

→ 성별, 날씨, 강수량에 따라

일부 재화에서 소비패턴의 차이가 보임

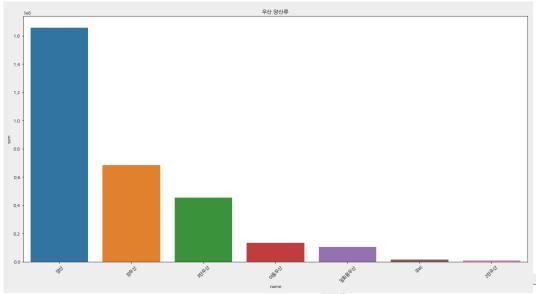
ex) 주로 덥고 습한 날에 살충/방충제 항목 재화 구매율이 증가함





40대 여성 더울 때 항목 : 국산과일

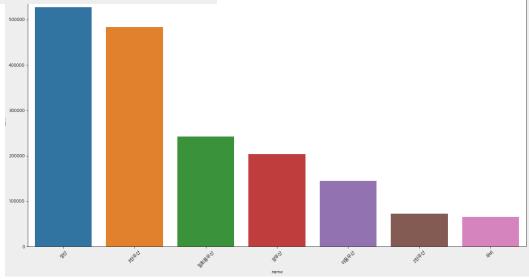
40대 여성 추울 때 항목 : 국산과일



40대 여성 따뜻하고 비가 안 올 때

항목 : 우산/양산류

40대 여성 따뜻하고 비가 올 때 항목 : 우산/양산류



성별, 날씨, 강수량에 따라 소비패턴에 차이가 보이는 컬럼들을 추출하고 비슷한 항목들을 그룹화함

음식: 국산과일, 냉동과일, 미국산소고기, 생선회, 아이스크림, 열매채소, 온장조리, 잎채소, 조개류, 프리믹스, 차류 (11개)

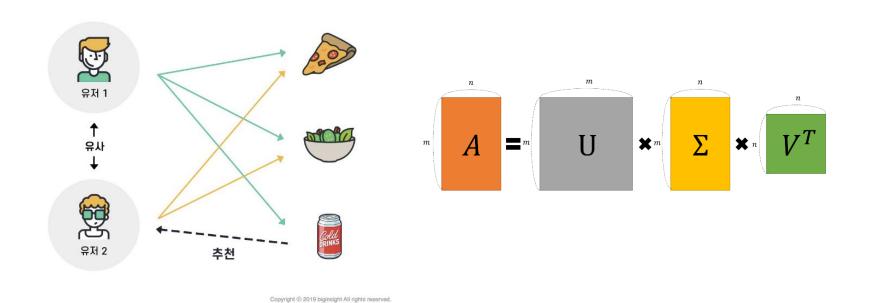
전자제품: TV, 공기청전/가습/제습, 난방가전, 냉방가전, 냉장/냉동고, 오디오 (6개)

취미: 수영/물놀이, 스키/보드, 안마/찜질용품, 여성등산/아웃도어의류 (4개)

생활: 거실가구, 방향/제습/탈취제, 살충/방충제, 성인침구, 수납가구, 우산/양산류, 카페트/러그/매트류, 선케어(8개)

패션: 남성의류아우터, 안경/선글라스, 여성속옷, 여성양말류, 장갑,패션액세서리, 핸드/풋케어,향수 (8개)

하이브리드 추천 시스템



SVD를 이용, User와 Item의 Latent Vector를 구하여 이를 바탕으로 유저 기반 추천 시스템을 활용하는 방식

40대 여성 더울 때 vs 추울 때 음식 추천 비교

Best Match: 0.9999999371509534, Best Match ID: 1431

	M016822877		M961815188	pd_nm		
pd_nm		pd_nm		0	애호박	
깻잎	1480.0	상추	3280.0			
상추	3680.0	애호박	990.0			

Best Match: 0.9999999703930463, Best Match ID: 1412

		M000705571		M652857427	92	pd_nm
р	d_nm		pd_nm		0	귤류
S.	귤류	7980.0	감	12480.0	1	참외
5	티마토	7480.0	귤류	28600.0	2	감
			석류	9960.0	3	석류
			쌈채소류	3280.0	4	토마토
			참외	19970.0	5	쌈채소류
			토마토	12900.0		

더울 때

왼쪽 고객과 가장 비슷한 구매를 한 오른 쪽 고객과 비교하여 새 야채인 애호박을 쪽 고객과 비교하여 더 다양한 과채류를 추천하는 것을 확인 할 수 있음

추울 때

왼쪽 고객과 가장 비슷한 구매를 한 오른 추천하는 것을확인 할 수 있음

40대 남성 더울 때 vs 추울 때 패션 추천 비교

Best Match: 1.0, Best Match ID: 13

M173329292 M008809647 pd_nm pd_nm pd_nm 남성캐주얼재킷 305900.0 0 남성캐주얼재킷 남성캐주얼재킷 70000.0

Best Match: 1.0000000000000002, Best Match ID: 9

MANAGANAAA

d	WIUU5392126				
pd_nm		pd_nm			pd_nm
립스틱/립라이너	48000.0	립스틱/립라이너	49000.0	0	립스틱/립라이너

MACTOCKETOS

더울 때

왼쪽 고객과 가장 비슷한 구매를 한 오른 왼쪽 고객과 가장 비슷한 구매를 한 오른 쪽 고객과 비교하여 남성캐주얼재킷을 조 쪽 고객과 비교하여 립스틱/립라이너를 금 더 구매하는 것을 추천하고 있음

추울 때

더 구매하는 것을 추천하고 있음

활용 방안

제품 추천

- 고객에게 필요한 제품을 예측하여 해당 제품을 추천
- 추천된 제품에 대하여 Ipay만의 혜택을 제공하여 Ipay 사용 촉진

사용자 추천

- 패턴이 비슷한 사람에게 Ipay 사용 추천
- 초대 코드처럼 패턴이 비슷한 사람에게 추천 시 추가 혜택 제공

개선 방안

시계열 성질의 추가

•조금 더 고급 모델을 활용하여 시계열 추천 시스템을 통한 시간별 제품 추천 으로 발전이 가능

더 많은 데이터

- •모델의 특성상 데이터를 세분화 할 수록 더 좋은 추천을 해줌
- •그 와 동시에 데이터의 수가 급격하게 감소됨
- •이를 해결하기 위하여 더 많은 데이터가 필요

다양한 종류의 데이터

•조금 더 세분화된 제품 분류/ 조금 더 자세한 제휴사 등 더 깊은 데이터를 통하여 더 구체적이고 정확한 추천을 할 수 있음(상품별 유사도를 추가하여 추천하주는 방식 등)

THANK YOU