



고객 라이프 스타일 분석을 통한
맞춤형 상품 추천 시스템

오프투온.

(오프라인 고객을 온라인으로)

중앙대학교 응용통계학과 박재우
중앙대학교 소프트웨어학부 이승진
중앙대학교 응용통계학과 이영현



Contents

1

분석 개요 및 데이터 탐색 03

1. 분석 목표 및 전략 2. 데이터 탐색 및 시각화 3. 데이터 전처리 및 파생 변수 생성

2

고객 및 상품군 분석 11

1. 온라인 행동 기반 고객 세분화 2. 상품 소비 기반 고객 군집화 3. L-Palette

3

추천 알고리즘 개발 19

1. 고객 분류 별 맞춤 상품 추천 시스템 2. 추천 시스템 결과

4

신규 비즈니스 모델 제안 03

1. 비즈니스 모델 (1) : 오프라인을 온라인으로 2. 비즈니스 모델 (2) :



- 분석 목표 및 전략
- 데이터 탐색 및 시각화
- 데이터 전처리 및 파생 변수 생성



분석 목표

온라인과 오프라인을 동시에 이용한
고객들의 라이프스타일 및 행동 양식 파악

라이프스타일 및 행동 양식을 기반으로
‘상품 소비’ & ‘행동 기반’ 고객 군집화 실행

오프라인 매장만 이용하는 고객들에게
차별화된 온라인 상품 추천 서비스 제공

Only 오프라인 이용 고객을 온라인으로
유입시켜 ‘옴니 채널’ 마케팅 BM 강화

분석 전략

(1) 데이터 전처리

- 시각화(Boxplot 등) 통해 이상치(Outlier) 제거
- EDA를 통해 도출한 결론을 기반으로 파생 변수 생성
- 날씨 (평균 기온, 강수량 등), 공휴일 등의 외부 데이터 수집

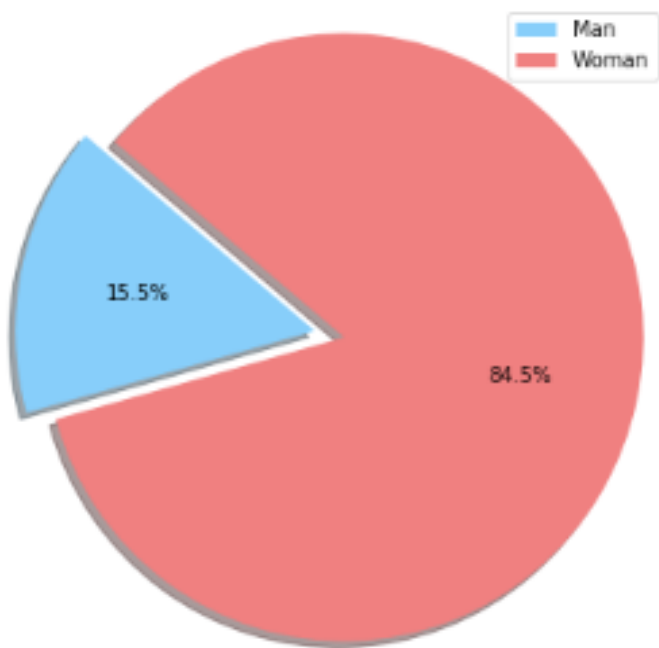
(2) 세부 과제 별 이용 기법

- ‘행동 기반’ 고객 군집화 : K-means 클러스터링, SVD
- ‘상품 소비’ 고객 군집화 : K-means 클러스터링, 랜덤 포레스트
- 행동 기반 추천 시스템 : 잠재 요인 기반, 아이템 기반 협업 필터링
- 비즈니스 모델 제안
 - 협업 필터링 모형을 이용해 ‘오프온투’ 신 BM 모델 제안
 - RFM 모형을 이용해 고객 맞춤 마케팅 전략 제안



고객 성별 분포

▶ 03.고객 Demographic 정보 기준

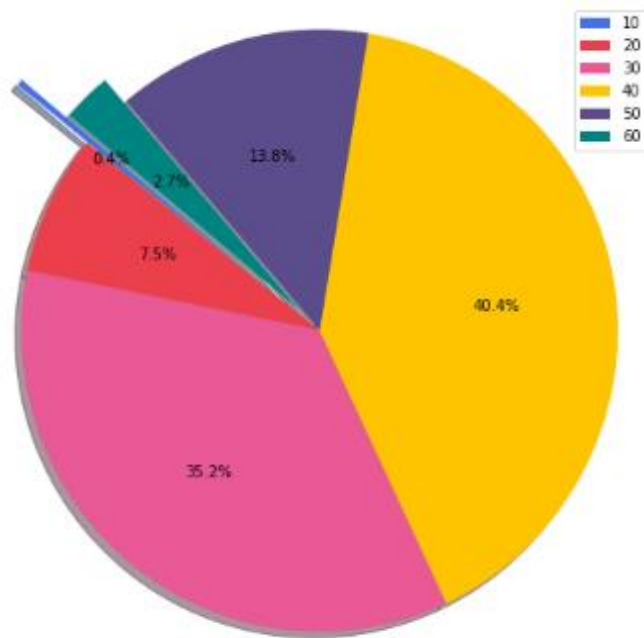


여성이 전체의 약 85% 차지

남자의 비율이 여성에 비해 매우 낮음

고객 구매 연령 분포

▶ 03.고객 Demographic 정보 기준



40대의 비율이 전체의 40% 차지

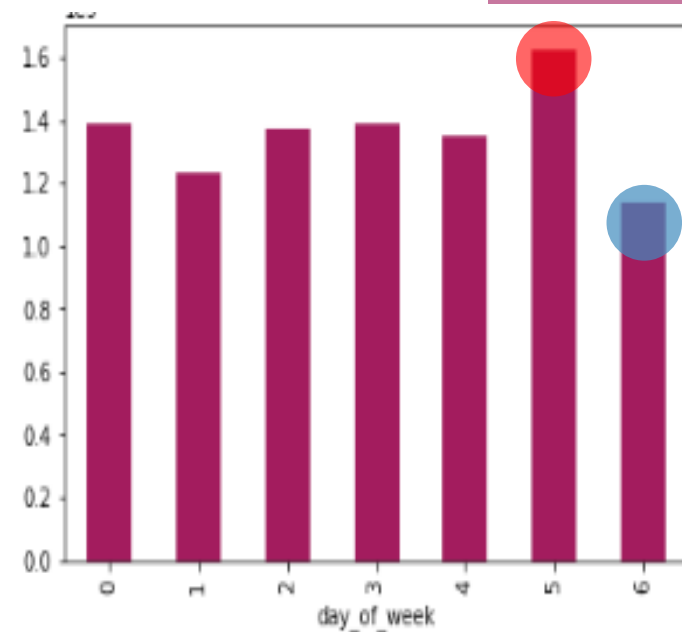
젊은층의 비율이 중장년층에 비해 매우 낮음

▶ 젊은층 (10대~20대), 중장년층 (30~50대)

요일 별 매출액

05

단위: 십억원



토요일의 매출이 가장 높고, 일요일이 최저

월요일 ~ 금요일 매출 차이 미미

데이터 탐색 및 시각화

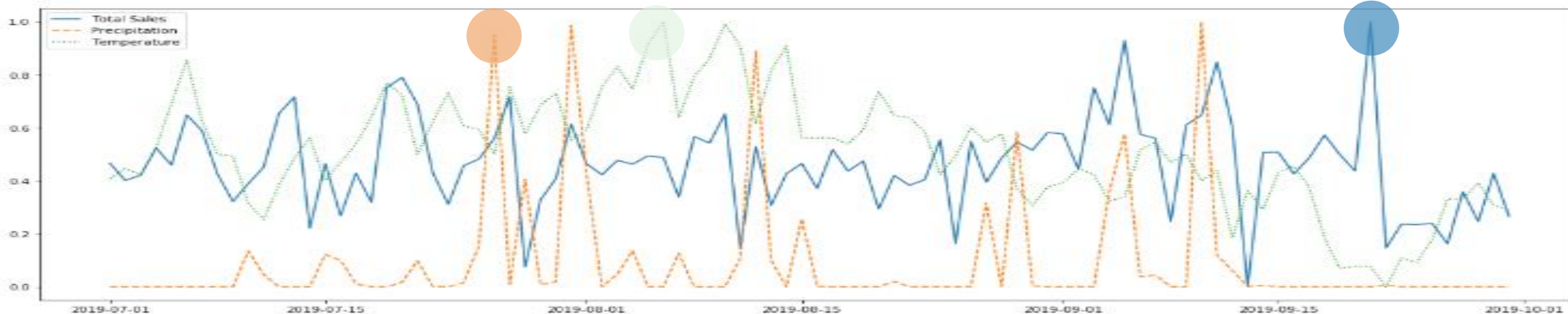


일간 매출액, 강수량, 평균기온 차트

외부 데이터 사용 : 기상청 오픈데이터 이용

MinMaxScaler를 통해 분포 조정

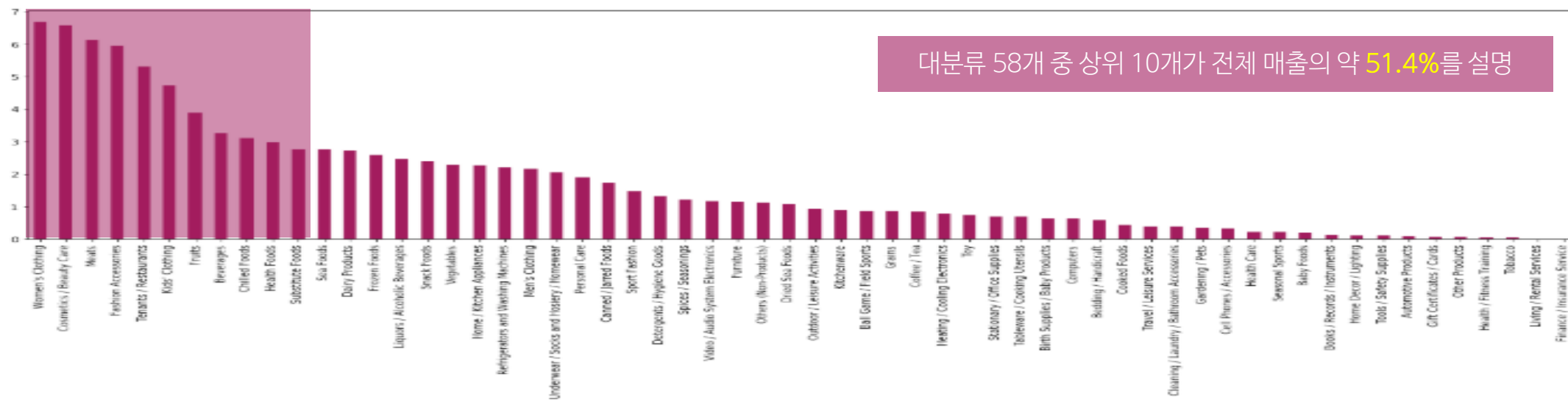
평균기온, 강수량, 매출액 간 상관관계 부족



상품군별 매출 비율

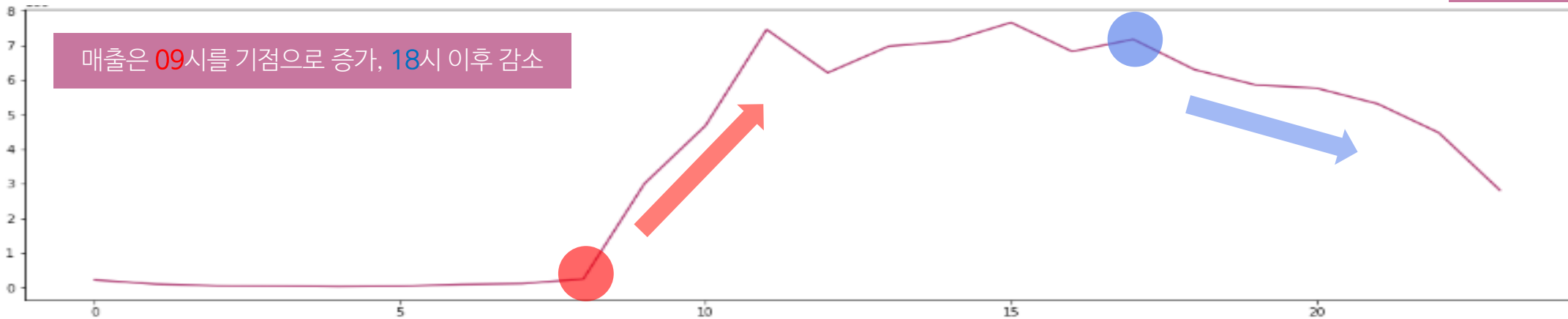
대분류('clac_nm1') 이용

단위 : %



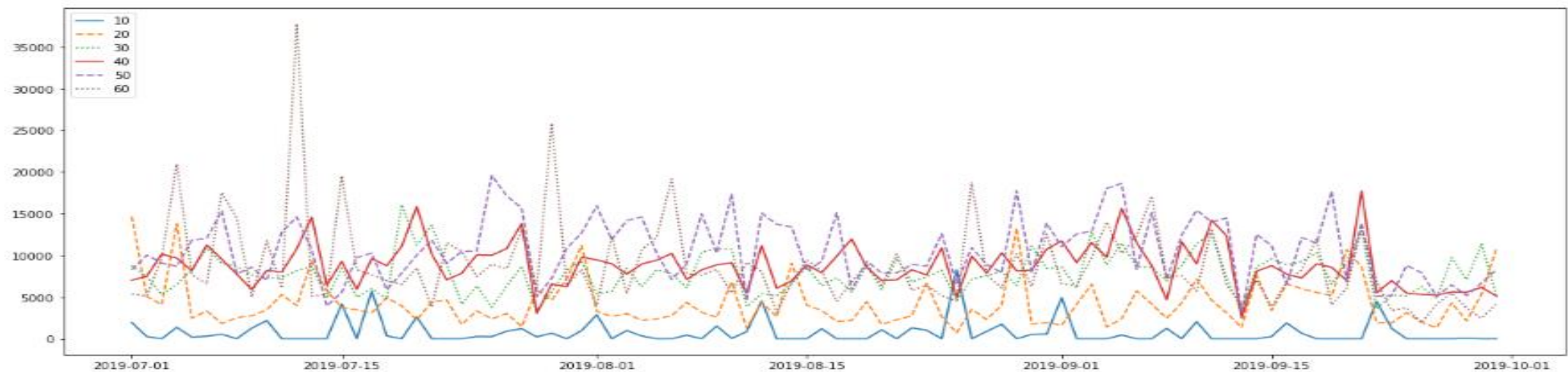


시간대별 매출액 추이



연령대 별 1인당 매출액 추이

연령대 별 매출액 순위는 60대 > 50대 > 40대 > 30대 > 20대 > 10대

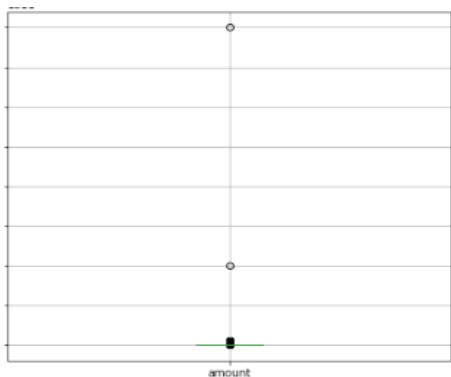




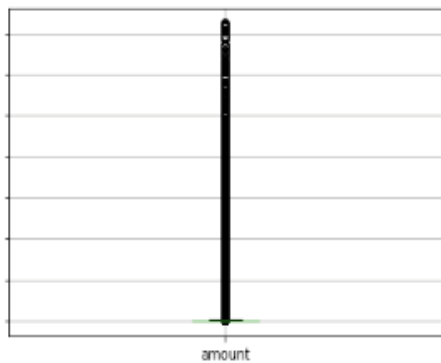
What

데이터 전처리

▶ 이상치 (Outlier) 제거



이상치 제거 전



이상치 제거 후

- 물품 구매 총액(물품 가격 × 구매 개수)의 분포를 boxplot으로 시각화
- $(-2\sigma, +2\sigma)$ 를 만족하는 데이터 외에는 이상치로 판단하여 제거
- 구매 총액 ('amount') = 상품 구매 금액 × 상품 구매 수량 = 0인 기록 제거

What

외부 데이터 수집

▶ 날씨 데이터 수집

수집 출처	기상청 날씨누리 (https://web.kma.go.kr/weather/main.jsp)
수집 데이터	2019. 7~9월 전국 평균 기온 및 강수량
데이터 사용	매출과의 상관관계 분석 (EDA)

▶ 공휴일 데이터 수집

수집 출처	대한민국 공공데이터포털 (https://www.data.go.kr/dataset/15012690/openapi.do)
수집 데이터	2019. 7~9월 대한민국 공휴일
데이터 사용	온, 오프라인 행동 기반 고객 군집화 분석



What

파생 변수 생성

▶ 시간 변수 생성 (1)

변수 명	설명
R_weekday	고객 별 평일 구매 비율
R_weekend	고객 별 휴일 구매 비율 (공휴일 포함)

Why

- 월요일~금요일 간 매출 차이가 명확하지 않음.
- 휴일과 평일의 차이가 중요함.

Plus

* 외부 데이터 사용

- 공휴일 : 8/15 (광복절, 목), 9/12~13 (추석, 목~금)

▶ 시간 변수 생성 (2)

변수 명	설명
R_Dawn	01 ~ 08시
R_Afternoon	09시 ~ 18시
R_Dinner	19시 ~ 24시

Why

- 시간대 별 매출액을 살펴본 결과 09시를 기점으로 매출액이 급상승하고, 18시를 기점으로 매출액이 하락함.

▶ 대대분류 변수 생성

Why

- 58개의 대분류를 13개로 줄임으로써 고객의 성향을 더욱 효과적으로 반영할 수 있도록 하기 위함.

Plus

- 대대분류 설정 시 롯데 홈쇼핑 사이트의 분류 이용



▶ 대대분류 변수 생성

대대분류 표기	대대분류 설명	해당 대분류
Food	식품	'Vegetables', 'Meats', 'Frozen Foods' etc
Beverage	음료	'Beverages', 'Coffee / Tea', 'Liquors / Alcoholic Beverages'
Kitchen	주방, 가정용품/여가	'Detergents / Hygiene Goods', 'Book/...' etc
Furniture	가구/인테리어	'Home Decor / Lighting' etc
Baby	유아품목	'Birth Supplies / Baby Products', 'Toy' etc
Fashion	패션	'Women's Clothing', 'Fashion Accessories' etc
Beauty	화장	'Personal Care', 'Cosmetics / Beauty Care'
Electronic	전자기기	'Computers', 'Video / Audio System Electronics' etc
Pet	애완용품	'Gardening / Pets'
Travel	여행/금융	'Gift Certificates / Cards', 'Travel / Leisure Services' etc
Sports	스포츠	'Sport Fashion', 'Seasonal Sports' etc
Service	서비스	'Others (Non-Products)', 'Living / Rental Services' etc
Other Products	기타 상품들	'Other Products', 'Tobacco'



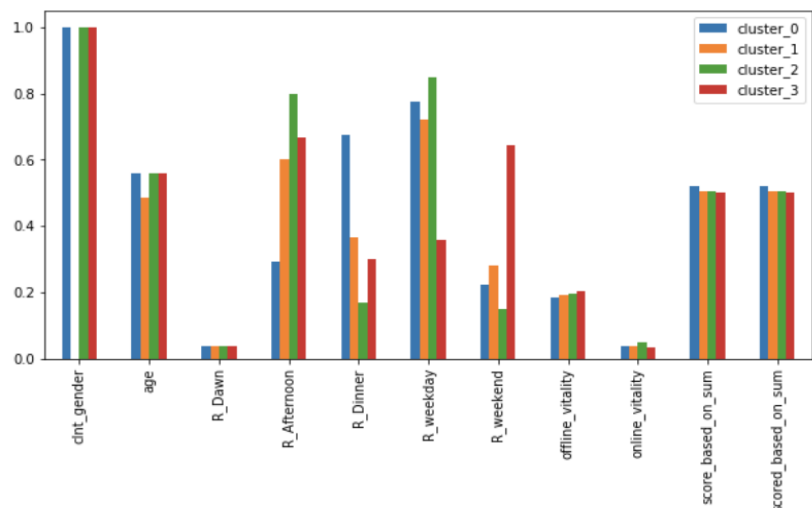
- 군집화 전략
- 온라인 오프라인 행동 기반 고객 군집화
- 상품 소비 기반 고객 군집화
- 최종 고객 군집화

2 고객 및 상품군 분석 군집화 전략

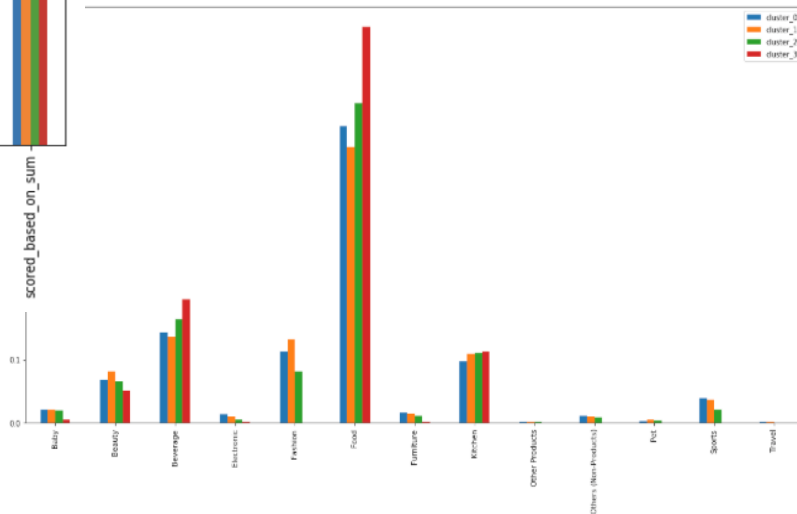


Issue

행동 양식과 상품 선호도는 관계가 없다.



행동 양식에 따른 분류



행동 양식에 따른 대대분류 선호도 점수

행동 양식을 군집화 한 뒤, 군집 별
11개의 대대분류 선호도 점수를 조사

군집 별 선호도 점수 차이가 명확하지 않아
행동 양식과 선호도를 합쳐 군집화 불가



행동 양식 기반, 상품 소비 기반
군집화를 나눠 고객을 최종 군집화

2 고객 및 상품군 분석 온라인 및 오프라인 행동 기반 고객 군집화



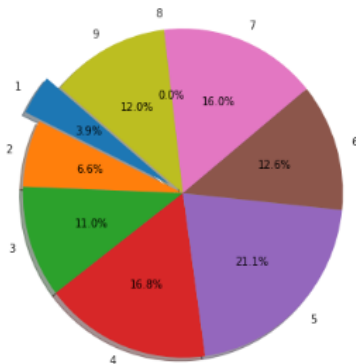
How

온라인 및 오프라인 기반 고객 군집화 방법

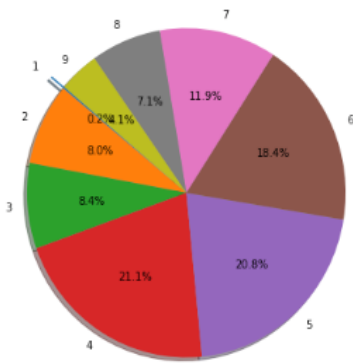
RFM 모형 도입

What is RFM?

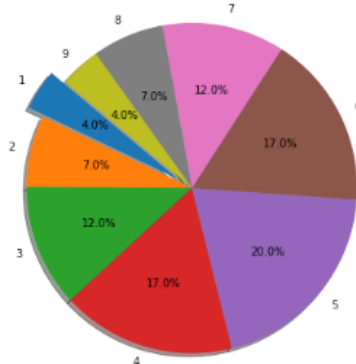
- CRM을 위한 고객 가치 분석의 대표적 개념
- 고객의 최근성 (Recency), 구매 빈도 (Frequency), 구매 금액 (Monetary) 데이터를 바탕으로 개별 고객이 기업의 수익에 얼마나 기여하는 가를 나타내는 지표



R (최근성) 구성비율



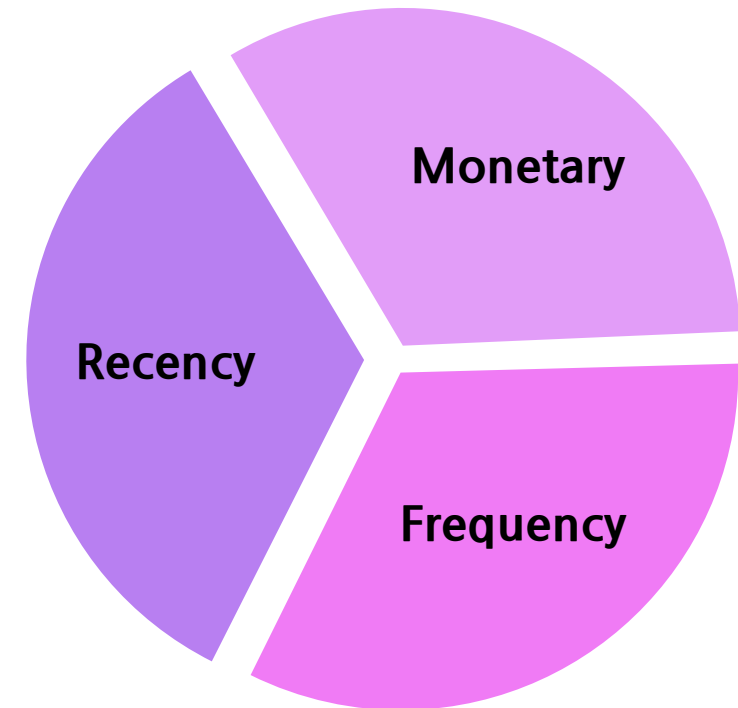
F (구매 빈도) 구성비율



M (구매금액) 구성비율

RFM 모형을 통한 제품의 가치 분석

고객의 가치 판단 -> 고객의 제품의 가치 판단



2 고객 및 상품군 분석 온라인 및 오프라인 행동 기반 고객 군집화



How

온라인 및 오프라인 기반 고객 군집화 방법

예시 (지수 계산 방법)

	index	count	total_buy_amount	percent	amount_contribution	attribution_effect
0	0	242	9.126804e+07	0.039	0.000603	0.015457
1	1	406	3.296860e+08	0.066	0.002178	0.032993
2	2	680	2.013465e+10	0.110	0.132987	1.208974
3	3	1041	2.292236e+10	0.168	0.151400	0.901188
4	4	1306	2.092082e+09	0.211	0.013818	0.065488
5	5	776	1.168669e+09	0.126	0.007719	0.061261
6	6	987	2.863698e+09	0.170	0.018914	0.111261
7	8	0	0.000000e+00	0.000	0.000000	0.000000
8	7	741	1.018006e+11	0.120	0.672382	5.603180

최빈성 Class

1. 각 등급별 매출액(total_buy_amount)을 전체 매출액으로 나누어 매출 기여도(amount_contribution) 계산
2. 매출 기여도를 각 등급 별 구성 비율(percent)로 나누어 기여 효과(attribution_effect) 계산
3. F, M Class에 대해서도 위와 동일하게 기여 효과 계산
4. 각 클래스 별 기여 효과의 합을 계산한 뒤, 클래스 별 기여 효과 비중을 가중치로 사용
5. R-가중치 * R + F-가중치 * F + M-가중치 * M 식을 사용하여 고객이 소비한 제품 별 RFM 지수를 계산

2 고객 및 상품군 분석

온라인 및 오프라인 행동 기반 고객 군집화



How

온라인 및 오프라인 기반 고객 세분화 방법

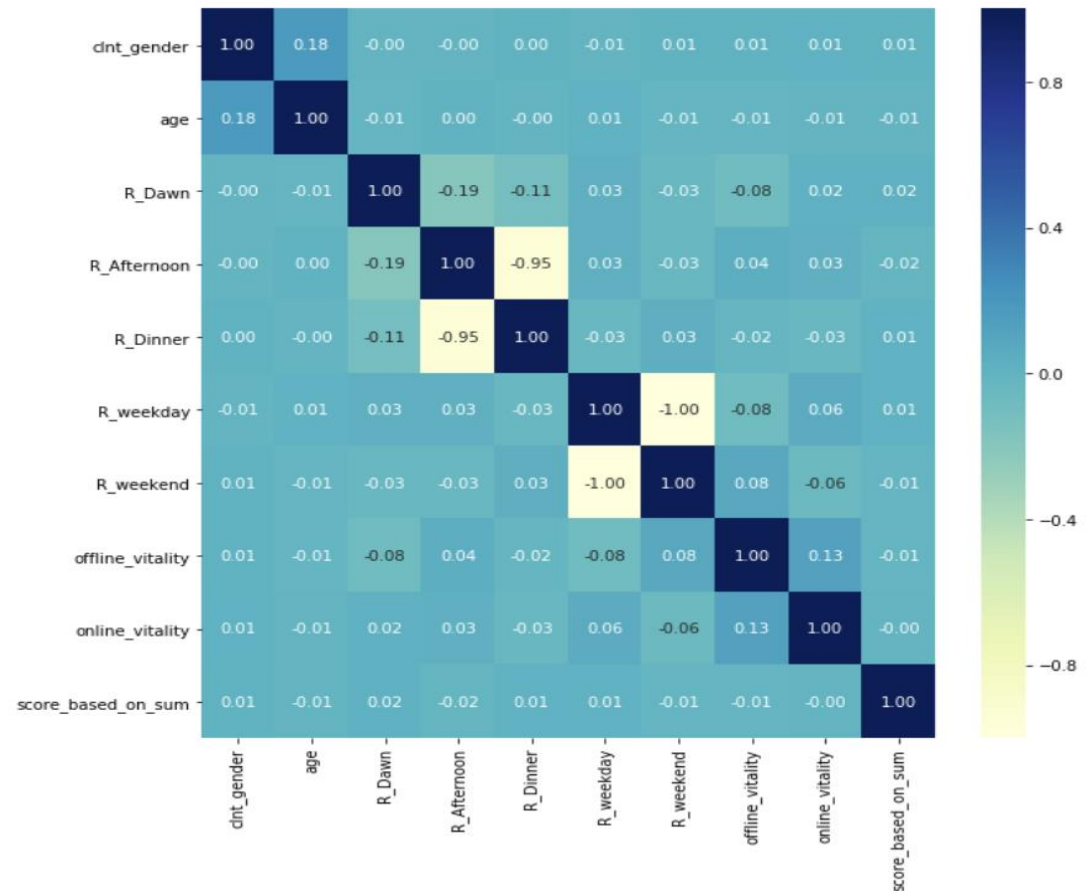
오프라인 활동성 변수

- Offline_Vitality
 - 구매 빈도 : 한 번 들어왔을때 얼마나 구매하는 가
 - 소비자 별 (구매 수량) \times (구매 빈도) $\times \log_{10}$ (총 구매 금액)
 - (구매 빈도) = (소비자 별 전체 거래 시도 횟수) \times (거래 시도 별 구매 횟수)
 - 소비자 별 전체 거래 시도 횟수 = 소비자 별 'trans_id' 개수
 - 거래 시도 별 구매 횟수 = 'trans_id' 별 'trans_seq'의 개수

온라인 활동성 변수

- Online_Vitality
 - 세션 별 (Max Hit) \times (Max Page View) $\times \log_{10}$ (Total Time)
 - Max Hit : 세션 별 'hit_seq' 값 중 최대값
 - Max Page View : 세션 별 'tot_pag_view_ct' 중 최대값
 - Total Time : 'tot_sess_ht_v'

전체 변수 Heatmap



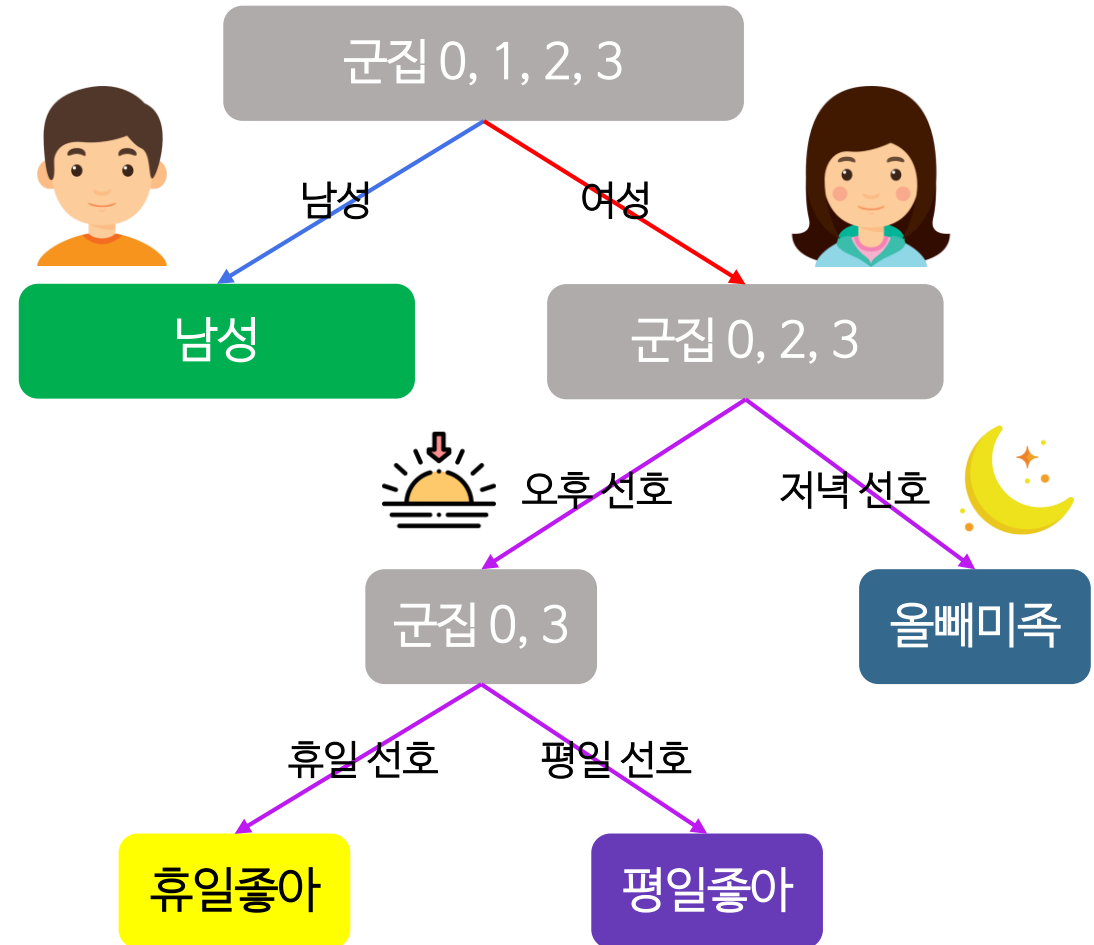
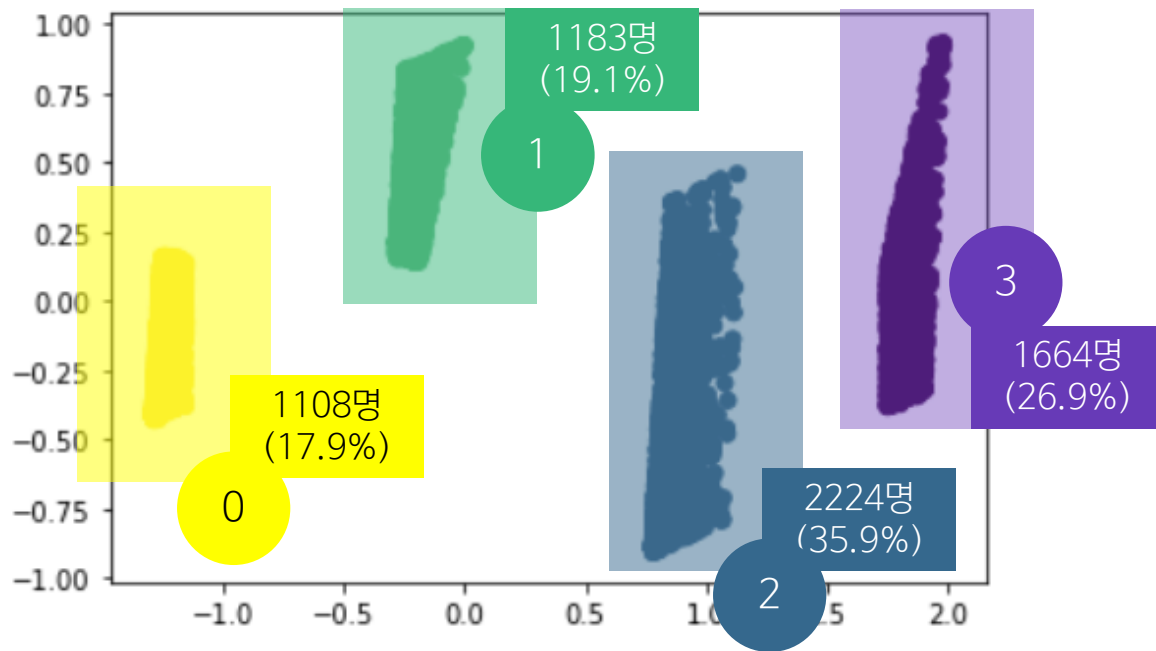
2 고객 및 상품군 분석

온라인 및 오프라인 행동 기반 고객 군집화



Result 군집화 결과

전체 : 6179(명)



2 고객 및 상품군 분석

상품 소비 기반 고객 군집화



How

상품 소비 기반 고객 군집화

RFM 변수

Issue

대대분류 별 가격 편차가 큰 문제 발생

- 가장 많은 고객이 구매한 대대분류 군인 Food가 Furniture, Beauty와 같은 대대분류와의 평균 가격 편차가 매우 커 RFM 지수만으로는 군집이 명확하게 나뉘지지 않음.

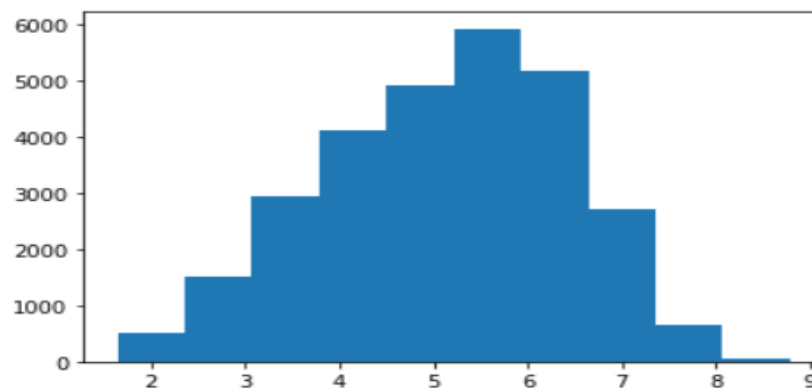
Solve

대분류의 상품 가격 차이를 고려한 RFM 점수 조정

- 상품 가격 차이를 고려한 새로운 구매 금액 지수($c_{buyamount}$) 생성
- $$c_{buyamount} = \frac{\text{구매자 별 해당 대대분류의 총 구매 가격 평균}}{\text{전체 구매자의 해당 대대분류의 총 구매 가격 평균}}$$
- $c_{buyamount} < 1$: 다른 구매자 대비 해당 상품군에 지출하는 금액 ↓
- $c_{buyamount} > 1$: 다른 구매자 대비 해당 상품군에 지출하는 금액 ↑

결과 및 해석

clnt_id	clac	total_buy_times	c_buy_amount	last_buy_date	R_Class	F_Class	M_Class	score_based_on_sum	
0	9	Beauty	1	0.334868	20190708	1	4	3	2.548356
1	9	Beverage	3	1.173843	20190727	3	6	6	5.001881
2	9	Food	7	0.967787	20190914	5	7	6	5.881063
3	20	Beverage	6	0.955858	20190930	9	7	6	7.211888
4	20	Food	6	1.105053	20190930	9	7	6	7.211888
...
28426	72424	Furniture	1	0.374800	20190718	2	4	3	2.881063
28427	72424	Service	1	0.417458	20190727	3	4	4	3.667294
28428	72424	Beverage	1	0.672426	20190727	3	4	5	4.120818
28429	72424	Food	3	1.118382	20190924	7	6	6	6.332706
28430	72424	Sports	2	2.562386	20190826	4	5	8	6.027868



- 상품군의 가격 차이에 의한 불균형 분포의 문제 해결

2 고객 및 상품군 분석

상품 소비 기반 고객 군집화



How

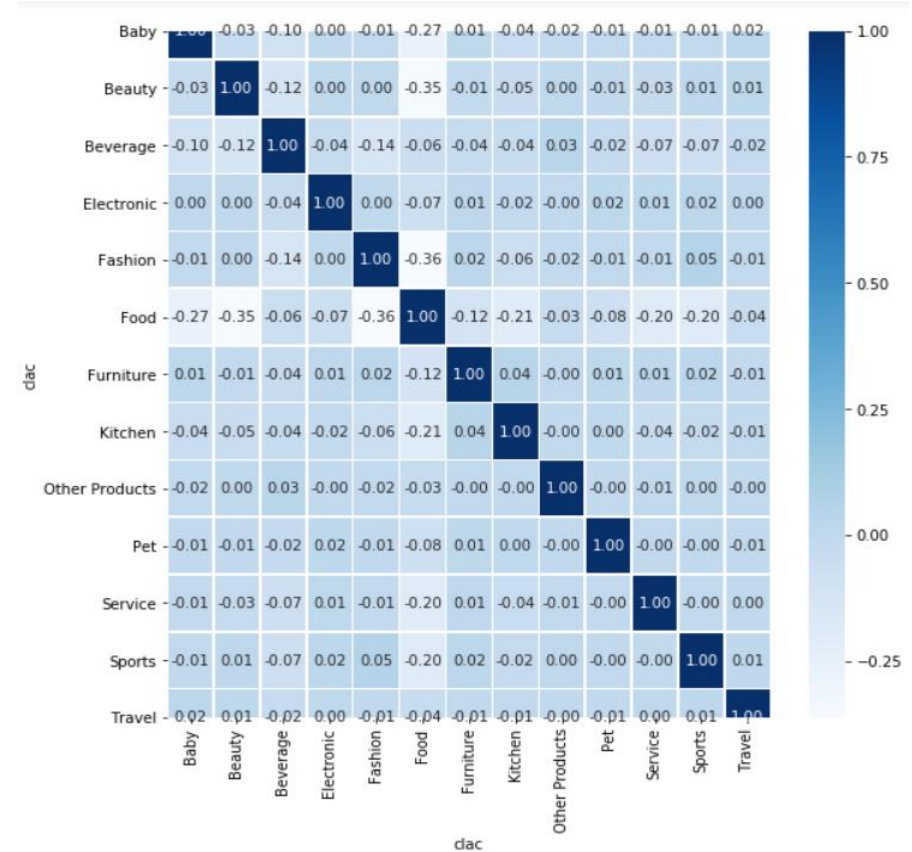
상품 소비 기반 고객 군집화

선택도 점수 (pf_{score})

- 구매 총액 ('amount') = 상품 구매 금액 × 상품 구매 수량 = 0 인 기록 삭제
- $pf_{score} = \frac{\text{구매자 별 해당 대대분류의 총 구매 수량}}{\text{구매자 별 전체 상품 구매 수량}} \times \log_{10}(\text{구매자별 해당 대대분류의 구매 총액}) \times RFM \text{ 지수}$

clac	clnt_id	Baby	Beauty	Beverage	Electronic	Fashion	Food	Furniture	Kitchen	Other Products	Pet	Service	Sports	Travel
0	9	0.000000	0.638228	2.741178	0.000000	0.000000	63.310913	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0
1	20	0.000000	0.000000	20.828435	0.000000	0.000000	63.759909	0.000000	0.976039	0.000000	0.0	0.000000	0.000000	0.0
2	23	0.887176	0.000000	2.330688	0.000000	0.000000	67.492302	0.000000	0.455432	0.000000	0.0	0.000000	0.000000	0.0
3	24	0.000000	0.000000	3.701753	0.000000	0.000000	58.412592	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0
4	29	0.000000	0.000000	16.494615	0.000000	0.000000	40.279895	0.000000	3.993226	0.000000	0.0	0.000000	0.000000	0.0
...
6174	72340	0.063231	4.593830	10.007796	0.000000	0.000000	43.945271	0.000000	9.531504	1.342073	0.0	0.125336	0.000000	0.0
6175	72356	0.000000	3.841081	15.670891	7.110346	1.807486	10.374199	0.000000	10.171152	0.000000	0.0	1.564488	15.973506	0.0
6176	72410	0.441161	0.267547	21.434391	0.000000	1.100046	47.730403	0.000000	1.724305	0.000000	0.0	0.000000	0.000000	0.0
6177	72423	0.000000	0.000000	0.000000	0.000000	0.000000	66.352723	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0
6178	72424	0.000000	0.000000	2.134358	0.000000	0.000000	50.917632	0.363501	9.713593	0.000000	0.0	1.433102	3.025423	0.0

전체 변수 Heatmap



- 대대분류들이 대체로 독립재 (커피와 소금과 같이 관련 없이 독자적으로 사용되는 재화)
- Food 와 Beauty, Fashion이 역의 관계를 보임

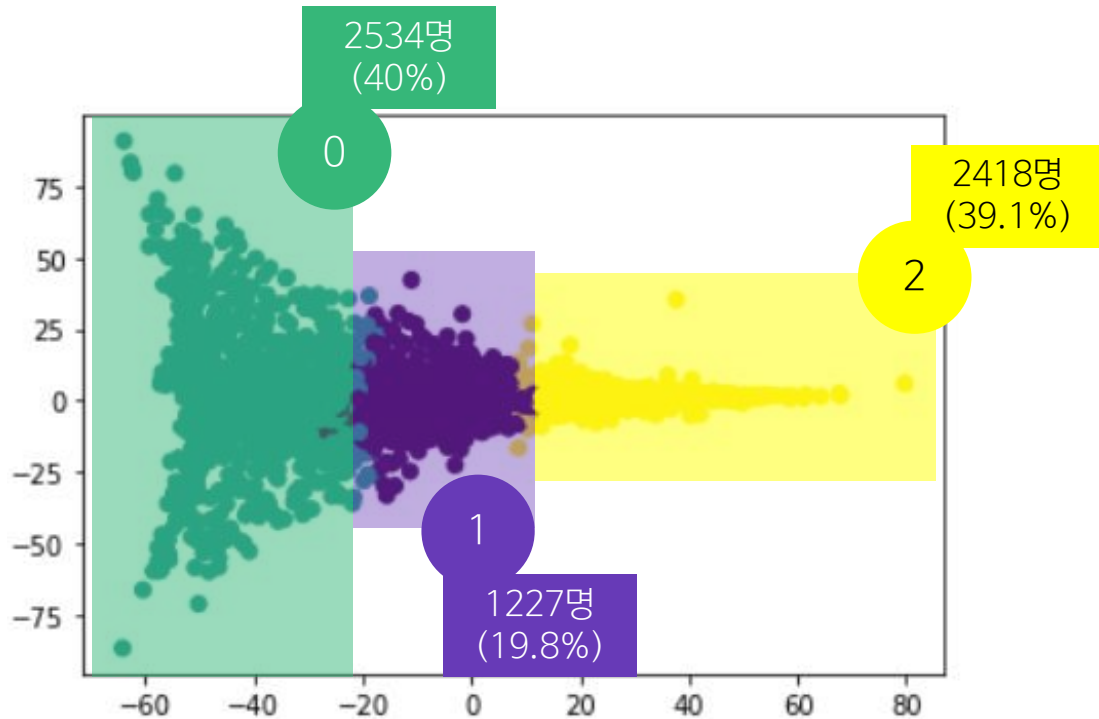
2 고객 및 상품군 분석

상품 소비 기반 고객 군집화

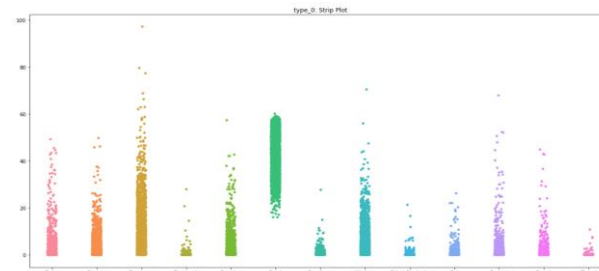


Result 군집화 결과

전체 : 6179(명)



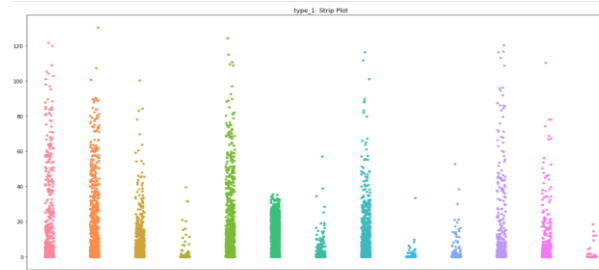
0번 군집 : 살림꾼



Beverage
Food
Furniture



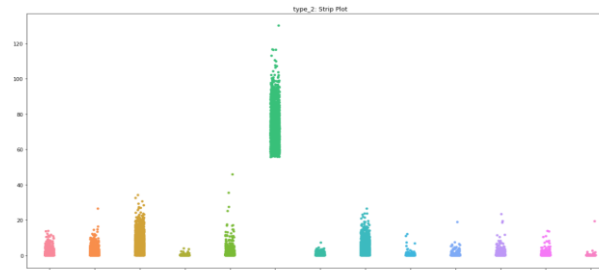
1번 군집 : YOLO족



Baby
Beauty
Fashion
Kitchen
Service
Sports



2번 군집 : 푸드러버



Food



2 고객 및 상품군 분석

L-Palette

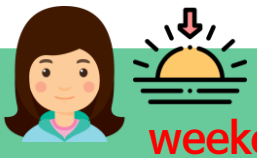





Result

최종 군집화 결과

온/오프라인 행동 기반

상품 기반

	weekend 			weekday 
살림꾼	502명	442명	671명	919명
YOLO족	226명	231명	326명	444명
푸드러버	455명	435명	667명	861명



온라인 및 오프라인 행동 기반, 상품 소비 기반의 고객 군집화의 조합을 통해
고객 성향 및 라이프 스타일에 대한 인사이트 도출

전체 : 6179(명)

3

추천 알고리즘 개발



- 고객 분류 별 맞춤 상품 추천 시스템
- 추천 시스템 시연



What

추천 시스템 알고리즘 (1)

▶ 아이템 기반 협업 필터링

- 고객이 좋아하는 아이템을 찾고, 좋아하는 아이템과 유사한 아이템을 추천하는 방식
- RFM 지수를 대대분류에서 소분류로 확장하여 각 소분류의 선호도 점수를 매김

3 추천 알고리즘 개발

고객 분류 별 맞춤 상품 추천 시스템

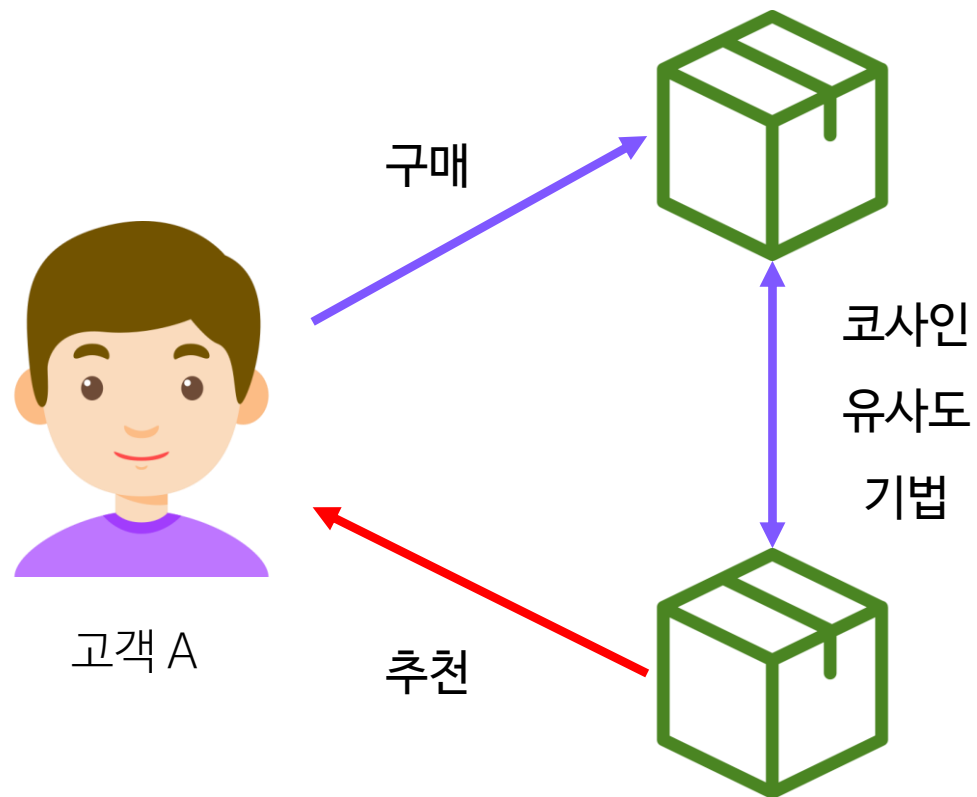


Method 아이템 기반 협업 필터링

	고객 A	고객 B	고객 C	고객 D	고객 E
사이다	5	4	5	3	5
커피	5	4	4		5
사과	4	2	1		2

➡ 각 제품에 대한 고객들의 **선호도**를 통해 아직 구매하지 않은 제품에 대해 추천이 가능

▶ 코사인 유사도 (Cosine Similarity)



➡ **코사인 유사도**를 이용해 고객이 구매한 제품과 유사도 값이 가장 비슷한 상품을 추천



What

추천 시스템 알고리즘 (2)

▶ 잠재 요인 기반 협업 필터링

- 데이터 속 숨은 요인 발견하기
- 사용자와 제품을 연결하는 잠재 요인 발굴

clac_nm3		Accessory Bags	Accident Prevention Equipment	Adhesive Tapes	Adhesives	Adult's Bed Covers and Skirts	Adult's Diapers	Adult's Disposable Briefes	Adults' Bed Fillings	Adults' Bedding Sets	...
clnt_id											
9	1.284126	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	1.217771	2.064066	...
20	0.000000	0.0	1.058024	0.000000	0.0	0.0	0.0	0.0	1.816715	1.279038	...
23	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.943401	4.381274	...
24	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	1.403654	...
29	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.783511	3.356294	...
41	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.821557	2.138476	...
43	1.645398	0.0	2.984640	0.000000	0.0	0.0	0.0	0.0	6.707434	6.889272	...
45	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	0.000000	...
49	0.000000	0.0	0.000000	2.186931	0.0	0.0	0.0	0.0	3.148192	8.176062	...
57	2.406380	0.0	0.000000	1.402851	0.0	0.0	0.0	0.0	0.000000	5.838208	...

사용자 - 구매 제품 데이터





Method

잠재 요인 기반 협업 필터링

- 고객이 아직 구매하지 않은 품목에 대해 **예측 평점** 산출

사용자 - 잠재 요인 행렬

	Item 1	Item 2	Item 3	Item 4	Item 5
유저 1	4			2	
유저 2		5		3	
유저 3			3	4	4
유저 4	5	2	1	2	

원본 행렬

사용자 - 아이템 평점 행렬

≈

	요인 1	요인 2
유저 1	0.94	0.96
유저 2	2.14	0.08
유저 3	1.93	1.79
유저 4	0.58	1.59

×

	Item 1	Item 2	Item 3	Item 4	Item 5
요인 1	1.7	2.3	1.41	1.36	0.41
요인 2	2.49	0.41	0.14	0.75	1.77

잠재 요인 - 아이템 행렬

매트릭스 분해

내적곱
결과

	Item 1	Item 2	Item 3	Item 4	Item 5
유저 1	4	2.56	1.46	2	2.08
유저 2	3.82	5	3.02	3	1.02
유저 3	5	5	3	4	4
유저 4	5	2	1	2	3.05

3 추천 알고리즘 개발

고객 분류 별 맞춤 상품 추천 시스템



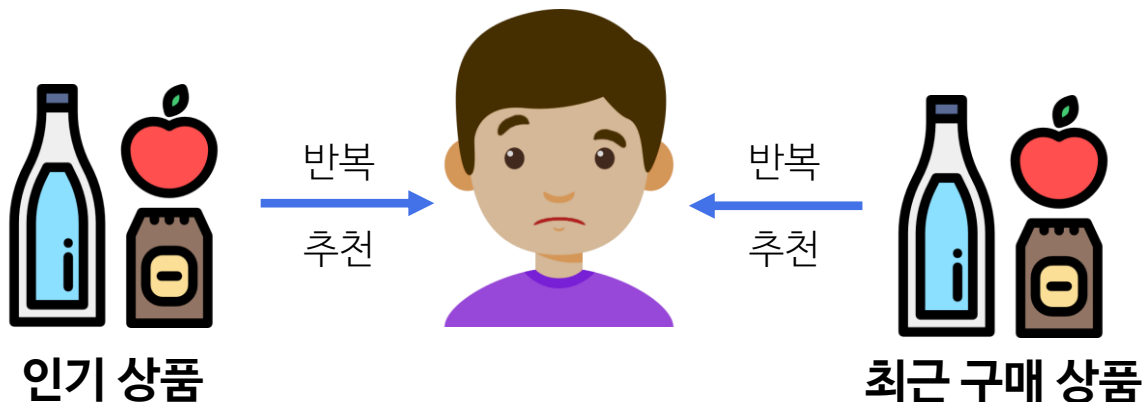
Issue

추천 시스템의 고질적 문제

인기 상품의 반복 추천
이미 구매한 상품의 반복 추천



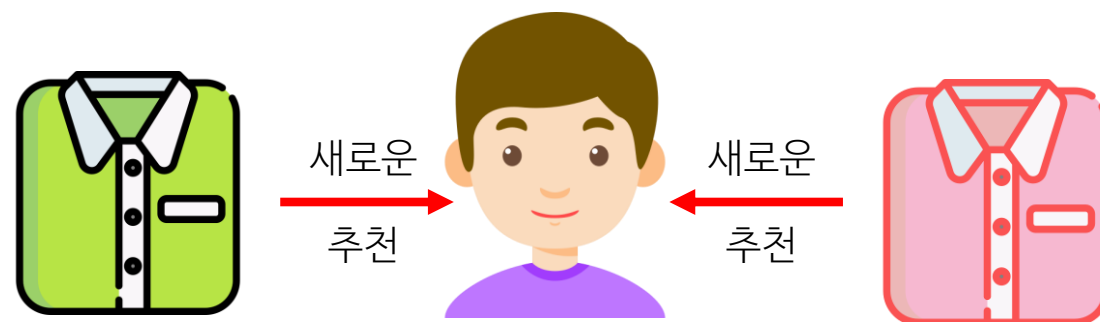
고객의 추천 상품 만족도 저하



인기 있는 상위 1%의 품목
최근 1주일 구매 품목



추천 대상 제품 군에서 제외



고객 분류 별 맞춤 상품 추천 시스템



Basic

아이템 기반 협업 필터링

- 구매를 하지 않아 선호도 점수가 없는 제품에 대해서도 선호도 점수를 도출하여 유사한 제품에 따라 추천
- 단순하며 설명이 가능한 직관적인 모델

▶ Memory base Algorithm

Plus

잠재 요인 기반 협업 필터링

- 특이값 분해(SVD)를 이용한 모델 기반 알고리즘
- 메모리 베이스 모델에 비하여 높은 정확도
- 잠재요소를 고려 가능 하지만 설명이 불가능한 블랙박스 모델

▶ Model base Algorithm

스테킹 기법으로 각 모델의 단점을 보완한

HYBRID 추천 시스템



Result

추천 시스템 결과

고객 9번	Gas Stations	Yogurt Drinks	Peppers	Grapes	Domestic Beer	Makgeolli	Persimmons	Jumbo Spring Onions	Kids' Casual Sport Socks and Hosiery	Flavored Milk
pred_score	52.0	50.0	49.0	49.0	49.0	48.0	48.0	48.0	48.0	48.0

Hybrid 추천 시스템

- 결과 코멘트 1 : 같은 대분류에 있는 제품만 추천 해주던
아이템 기반 추천 시스템의 문제점이 해결
- 결과 코멘트 2 : 단순

분석 결과 활용 방법

- RFM 기반 상품 선호도 점수를 통한 고객 추천 결과를 통해
'Offline to Online' 활성화 서비스에 추천 상품 목록 제공

분석 조건

- 데이터 셋의 개수 : 6179
- 함수 Feature : 1614 (전체 소분류)
- 모델 평가 방법 : 잠재요인 추천 - Hold Out (Test Size 0.2)

분석 수치

- RMSE : 아이템 기반 인접 TOP10 이웃 MSE = 22.01
잠재요인 추천 RMSE = 23.65
- 결과 코멘트 : 아이템 기반 추천 모델의 RMSE 의 경우 10개의 상위 유사도를 갖는 아이템에 대한 RMSE 인 것을 감안하면 잠재요인 추천모델의 RMSE 가 매우 낮으며 ,곧 더 정확한 추천 모델임을 알 수 있다



4

비즈니스 모델 제안



- 비즈니스 모델 1 : 오프라인 고객을 온라인으로 !
- 비즈니스 모델 2 : RFM 지수에 따른 고객 맞춤 마케팅 전략

4 신규 서비스 제안 오프라인 고객을 온라인으로 !



Why

오프라인 매장의 위기 (1)



20세기 유통 공룡 '시어스'
2018. 10 파산신청



20세기 장난감 천국 '토이저러스'
2017. 09 파산신청

두 회사의 “**공통점**”



오프라인 유통망에 안주, Only **오프라인 이용 소비자**만 타겟



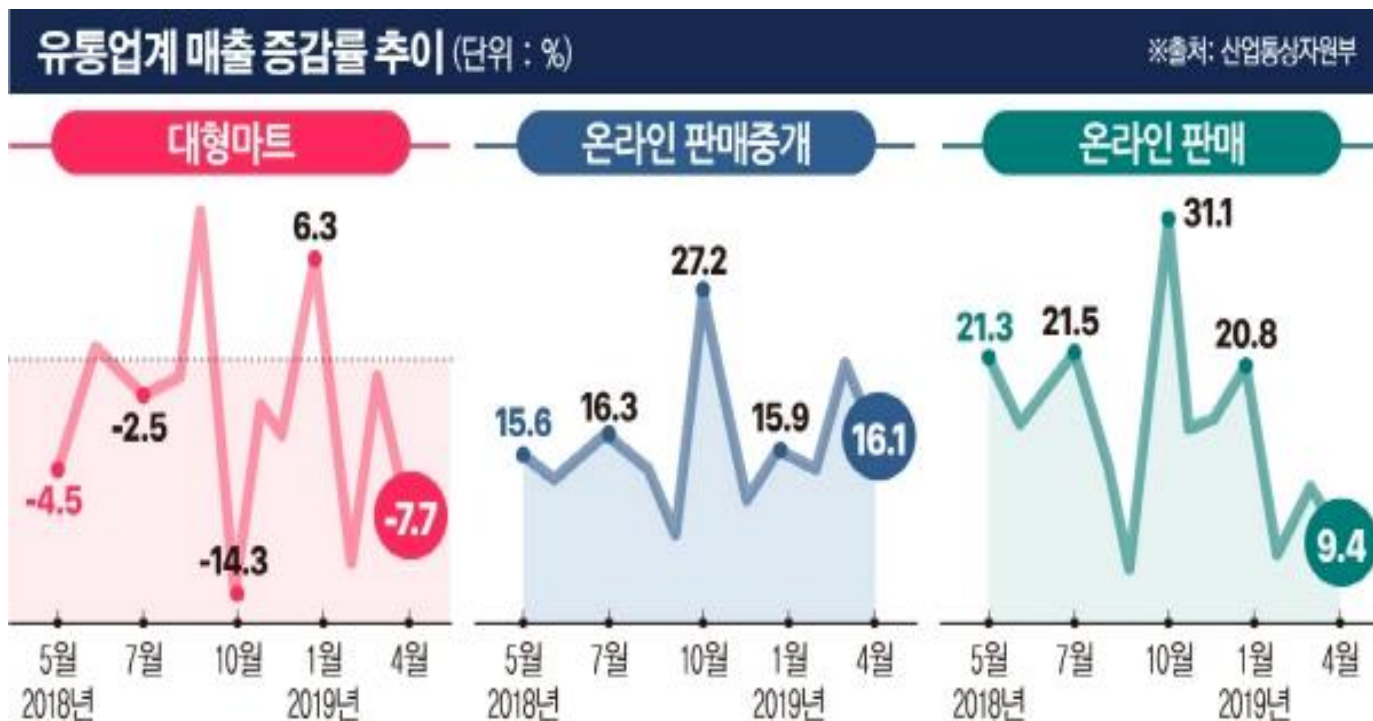
변화하는 시대의 **소비 트렌드**를 따라가지 못함

4 신규 서비스 제안 오프라인 고객을 온라인으로 !



Why

오프라인 매장의 위기 (2)



- 이마트, 롯데마트, 홈플러스 등 대형 유통업계의 **매출, 영업이익**이 꾸준히 **감소**
- 이로 인해 더 이상 오프라인 매장을 늘리지 않고, 온라인 판매 비중 확대

전자 상거래 (E-Commerce) 시장의 성장

‘모루밍족’, ‘쇼루밍족’의 증가

- ▶ 모루밍족 : 오프라인 매장에서 물건을 먼저 확인하고, 모바일 기기를 이용해 온라인에서 구입
- ▶ 쇼루밍족 : 오프라인 매장을 쇼룸처럼 이용하고 온라인에서 구입



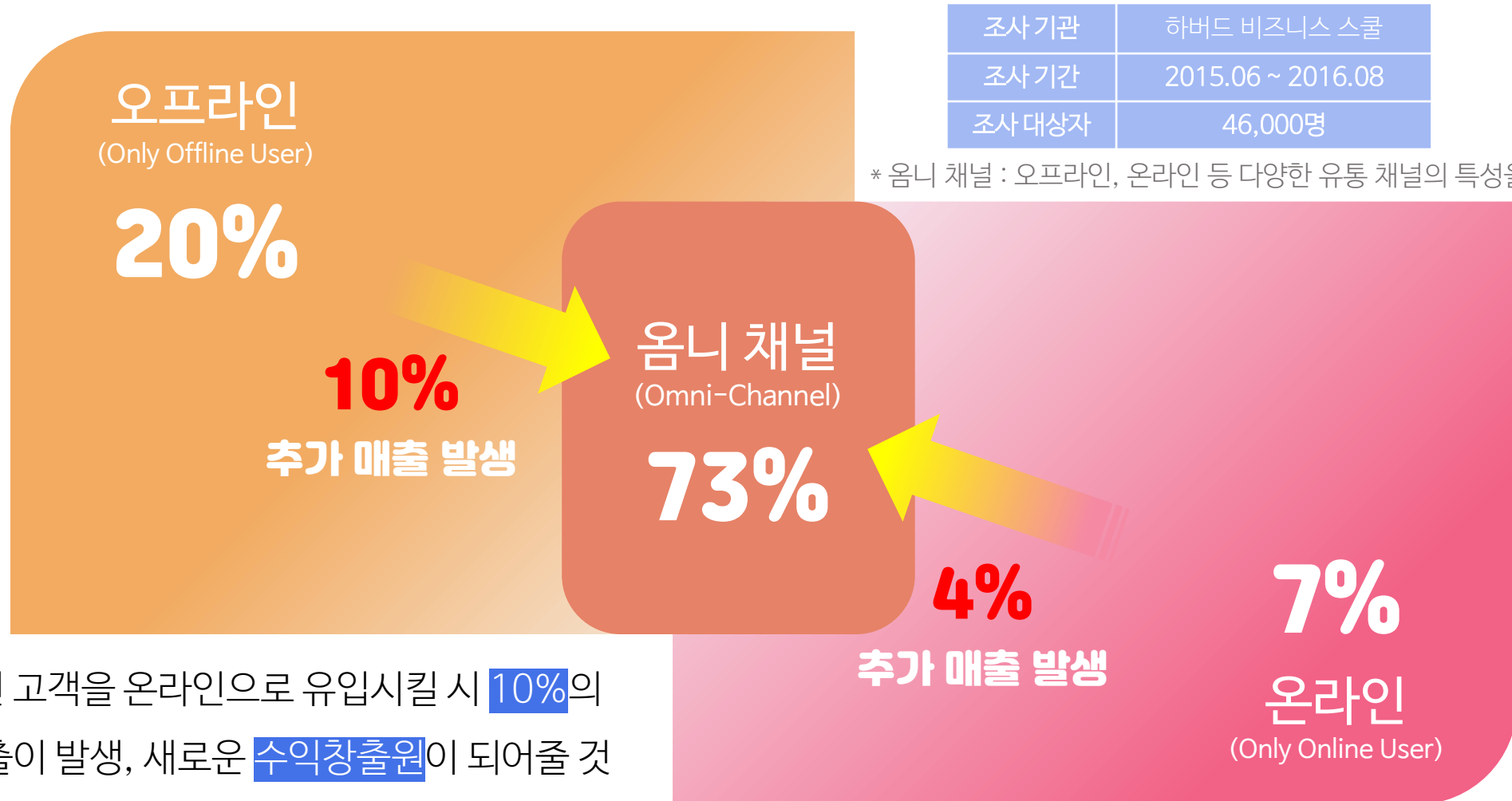
효과적인 추천 비즈니스 전략으로
오프라인 고객을 온라인으로 !

4 신규 서비스 제안 오프라인 고객을 온라인으로 !



Why

오프라인 고객을 온라인으로 유입했을 때의 이점



▶ 오프라인 고객을 온라인으로 유입시킬 시 10%의 추가 매출이 발생, 새로운 수익창출원이 되어줄 것
(온라인 고객을 오프라인으로 유입시키는 것에 비해 약 2.5배 효과적)

4 신규 서비스 제안 오프라인 고객을 온라인으로 !

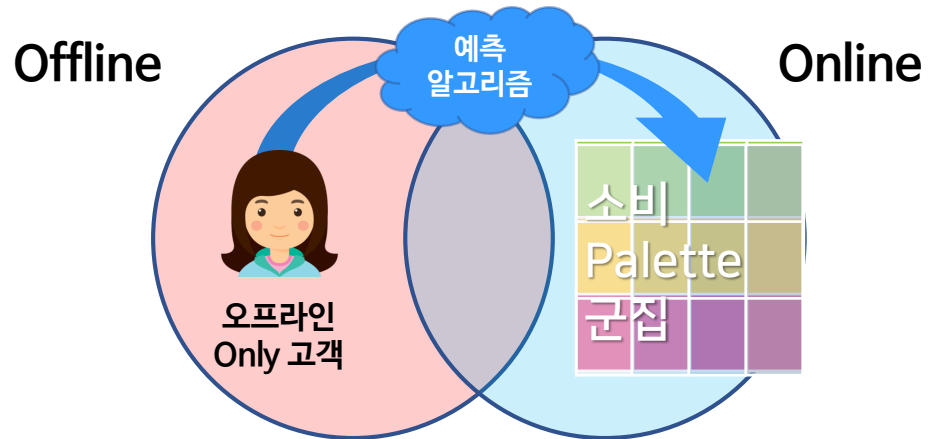


IDEA

Welcome to Offline!

▶ ‘오프온투’, ‘Offline to Online’ (가칭) 정의

오프라인만 이용하는 고객을 대상으로, ‘오프라인 구매 이력’ 과 ‘고객 행동 양식’을 바탕으로 추천 시스템을 구현하여 온라인 구매 유도를 위한 신규 서비스

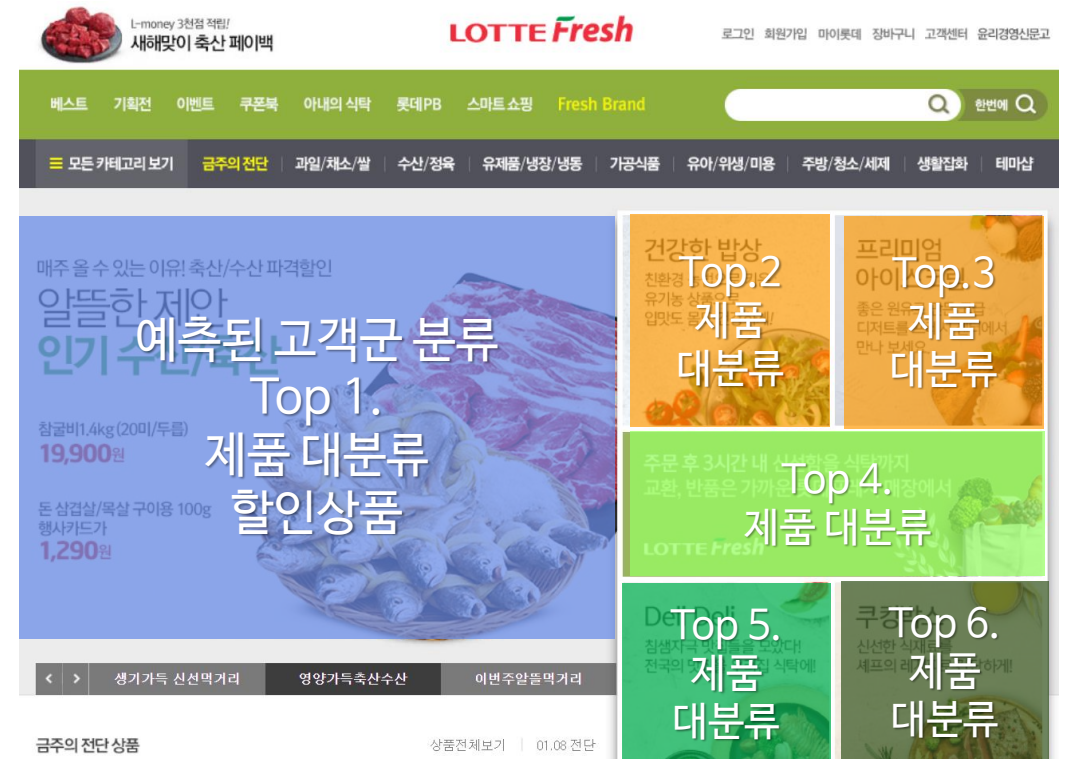


▶ 고객의 행동 정보를 랜덤 포레스트를 이용하여
신규 접속 고객 군집 예측

▶ 상세 예시 (1)



오프라인 고객층인 "엘보순" 고객의
롯데프레시 온라인 최초 접속 화면



▶ 예측된 군집군이 선호하는 제품군에 대한 추천

4 신규 서비스 제안 오프라인 고객을 온라인으로 !



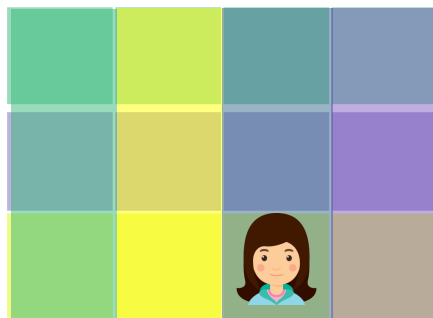
Scenario 차별화된 고객 맞춤 추천 시스템



오프라인
이용 고객



L.pay를
통한 구매



아 ! 엘포이님은

올빼미족 + 푸드러버

시군요 !

고객이 속할 군집 예측

- 행동 양식과 구매 상품 기록을 기반으로 12가지의 군집 중 가장 적합한 군집 예측



PUSH 알람 발생

- 행동 양식, 구매 상품 기반 추천 아이템을 구매 직후 푸시 알람 통해 추가 구매 유도



Top 10 아이템 추천

4 신규 서비스 제안 오프라인 고객을 온라인으로 !



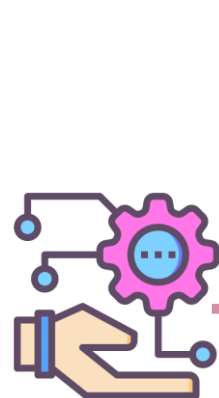
Scenario

차별화된 고객 맞춤 추천 시스템



모바일 홈 접속 화면

Top 10 아이템 추천



추천 알고리즘

Hybrid 추천 시스템

추천 제품
Top. 10



소비
Palette
군집

L.Palette
선호 제품
대분류 1
대분류 2
대분류 3

Hybrid 협업 필터링 추천 알고리즘을 통해
선별한 추천 제품 10가지에 대하여

고객의 L-Palette 군집의
제품 대분류 선호도에 따라 배열



특가상품 페이지

특가 상품 중
추천 알고리즘 상위점수
제품을 우선 노출



감사합니다