

# 구매 감소 고객 예측 및 추천 솔루션

TEAM UNIVERSE

최예림, 양주희, 이정현A, 진청아

# 목차

개요

1

2

탐색적 분석(EDA)

모델링

3

4

군집화

마케팅 제안\_ 추천시스템

5

UNIVERSE

개요

개발 환경

구축 환경



통합 개발 환경



Database



개발 언어



활용 기술

PYCARET

surprise





# 개요

## 팀원 소개

### 최애립

- 탐색적 분석
- 데이터 전처리
- 군집화



### 양주희

- 탐색적 분석
- 데이터 전처리
- 추천시스템



### 이정현

- 탐색적 분석
- 데이터 전처리
- 모델링



### 진청아

- 탐색적 분석
- 데이터 전처리
- 추천시스템



# 개요



## 분석 대상

**2014~2015년도에  
L사에 4개의 계열사에서  
구매한 고객**

## 분석 내용

**유지 고객 대상으로  
2년간의 구매 패턴을 파악하여  
구매 감소 고객 예측 모델 생성**

## 마케팅 제언

**감소 고객 대상  
추천 솔루션을 통한 마케팅  
제언**

UNIVERSE

# 탐색적 분석(EDA)

# 탐색적 분석 \_외부요인 분석

L사 마트 2015년 유통 키워드

UNIVERSE

**H**

(Hesitate to Buy)

**A**

(Anytime, Anywhere)

**R**

(Renew Everything)

**D**

(Desire to Safe)

# 탐색적 분석 \_외부요인 분석

## L사 마트 2015년 유통 키워드

UNIVERSE

L사 마트는 2015년 유통 키워드를 '어려웠다'란 의미의 'H.A.R.D'를 선정

### H

경기 침체가 장기화되며 소비 심리가 위축

유통업체들은 구매활성화를 위해 대형 할인 행사를 지속해서 선보였으며 정부 주도 행사에도 적극 동참하는 등 소비 심리 회복을 위해 노력

ex) 코리아 블랙프라이데이, K-세일데이 행사

### R

온라인 유통업체의 성장에 대응하기 위해  
오프라인 매장만의 강점을 어필

ex) 스타필드, 롯데타워, 더현대

### A

소셜커머스 회사에서 시작된 배송전쟁이 온-오프라인 유통업계 전반으로 확산  
드라이브&픽 서비스, 오토바이 퀵배송 서비스 등 마트&백화점에서 차별화된 배송 서비스 제공

이에 따라 대형마트,SSM 등 유통업체 온라인 전용 물류센터 구축

**\*윜니채널 시스템** 등 기존 온라인 배송과는 차별화되는 서비스 제공

### D

15년 6월 '메르스(MERS)', 10월 **육가공육 이슈** 등의 영향으로 대규모 다중시설인 백화점, 대형마트 등 유통업체를 방문하는 고객 수요가 급감

15년 6월 대형마트는 전년 동월대비 10.2%, 백화점은 11.9% 매출 감소

**\*윜니채널:**

소비자가 온라인, 오프라인, 모바일 등 다양한 경로를 넘나들며 상품을 검색하고 구매할 수 있도록 한 서비스 각 유통 채널의 특성을 결합해 어떤 채널에서든 같은 매장을 이용하는 것처럼 느낄 수 있도록 한 쇼핑 환경





## 고객 DEMO

- 고객번호
- 성별
- 연령대
- 거주지역



## 경쟁사이용

- 고객번호
- 제휴사
- 경쟁사
- 이용연월



## 멤버십 여부

- 고객번호
- 멤버십명
- 가입년월



## 상품분류

- 제휴사
- 대분류코드
- 중분류코드
- 소분류코드
- 중분류명
- 소분류명



## 채널 이용

- 고객번호
- 제휴사
- 이용횟수



## 구매데이터

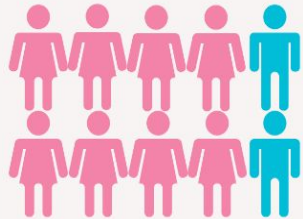
- 제휴사
- 영수증번호
- 대분류코드
- 중분류코드
- 소분류코드
- 고객번호
- 점포코드
- 구매날짜
- 구매시간
- 구매금액

# 탐색적 분석 \_데이터 탐색

고객 특징 및 데이터 분포 확인

UNIVERSE

## 고객 특징 데이터

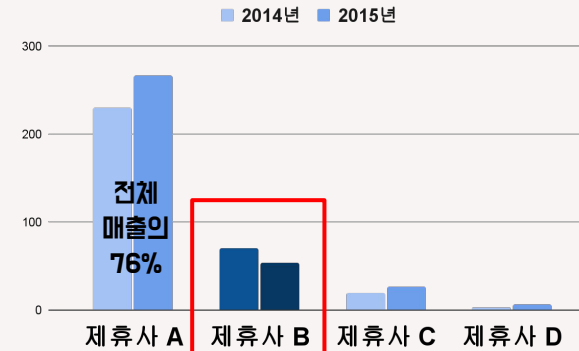


전체 고객  
80% 여성 20% 남성

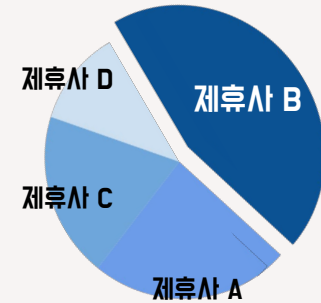


전체 고객  
70% 35세~54세 고객

## 제휴사 및 경쟁사 데이터



제휴사 별 매출 분석



경쟁사 이용 횟수

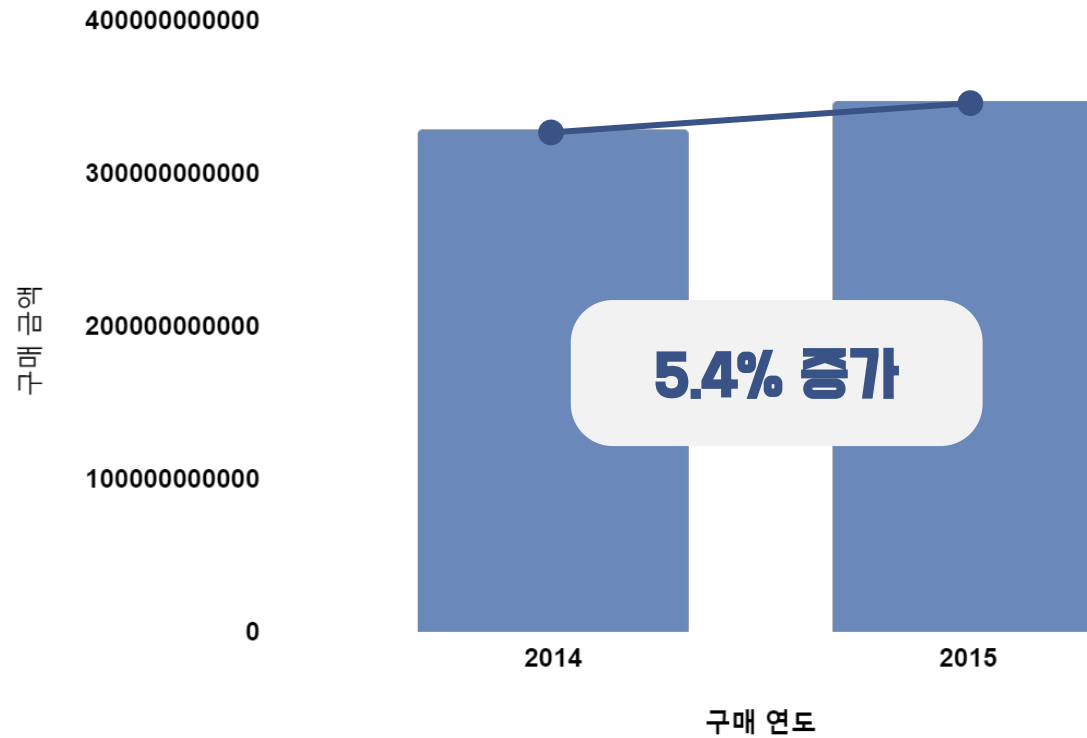
**“제휴사 B 2014년 대비 2015년 매출 감소,  
제휴사 중 2015년 경쟁사 이용 횟수 가장 높음”**

# 탐색적 분석 \_테이터 탐색

매출 확인\_연도별 매출

UNIVERSE

## 연도별 구매 금액 변화

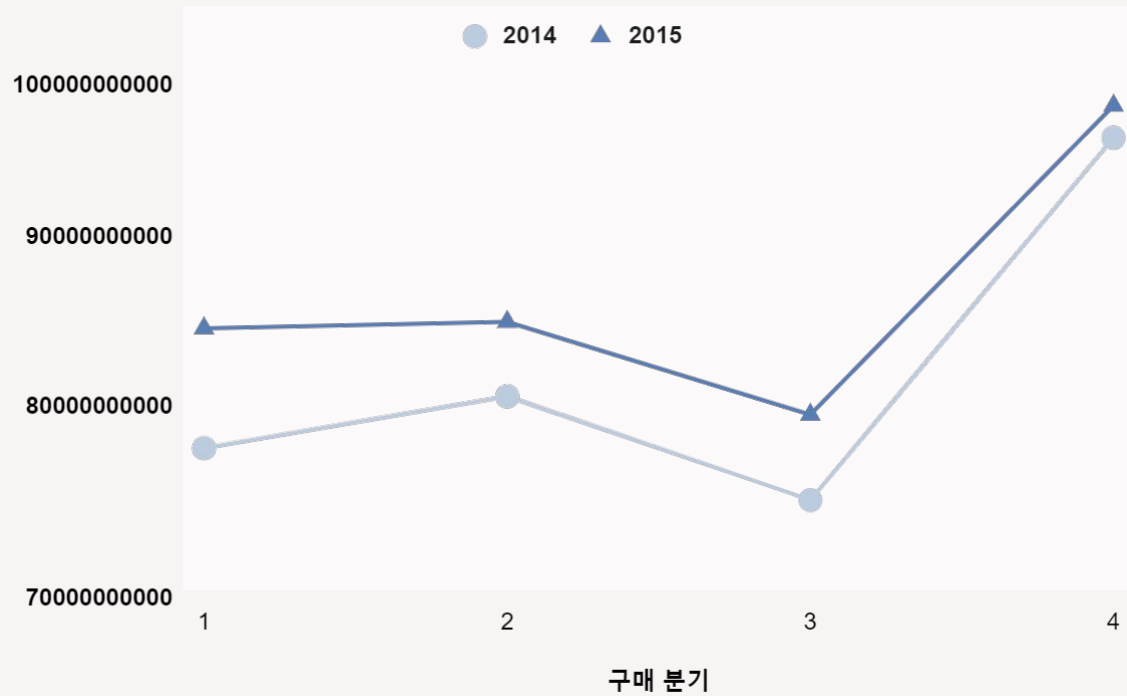


# 탐색적 분석 \_테이터 탐색

매출 확인\_연도별 매출

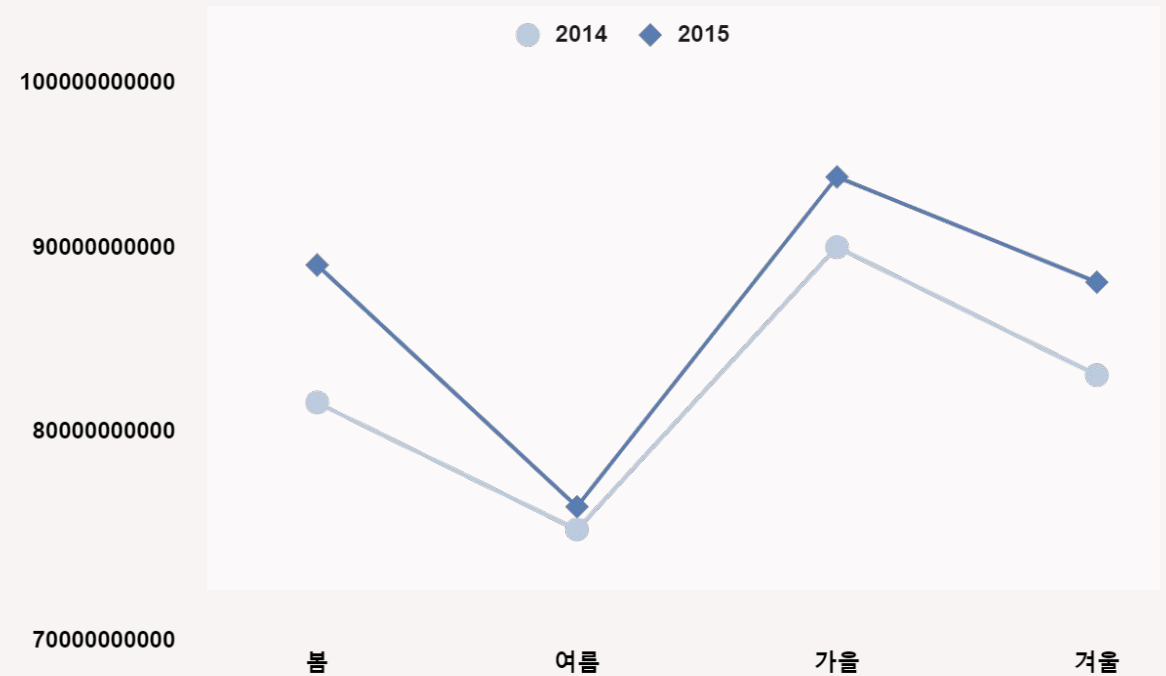
UNIVERSE

## 연도/분기별 구매 금액 변화



전년 동기 대비 증가폭은 1분기가 가장 크고, 4분기가 가장 작음

## 연도/계절별 구매 금액 변화



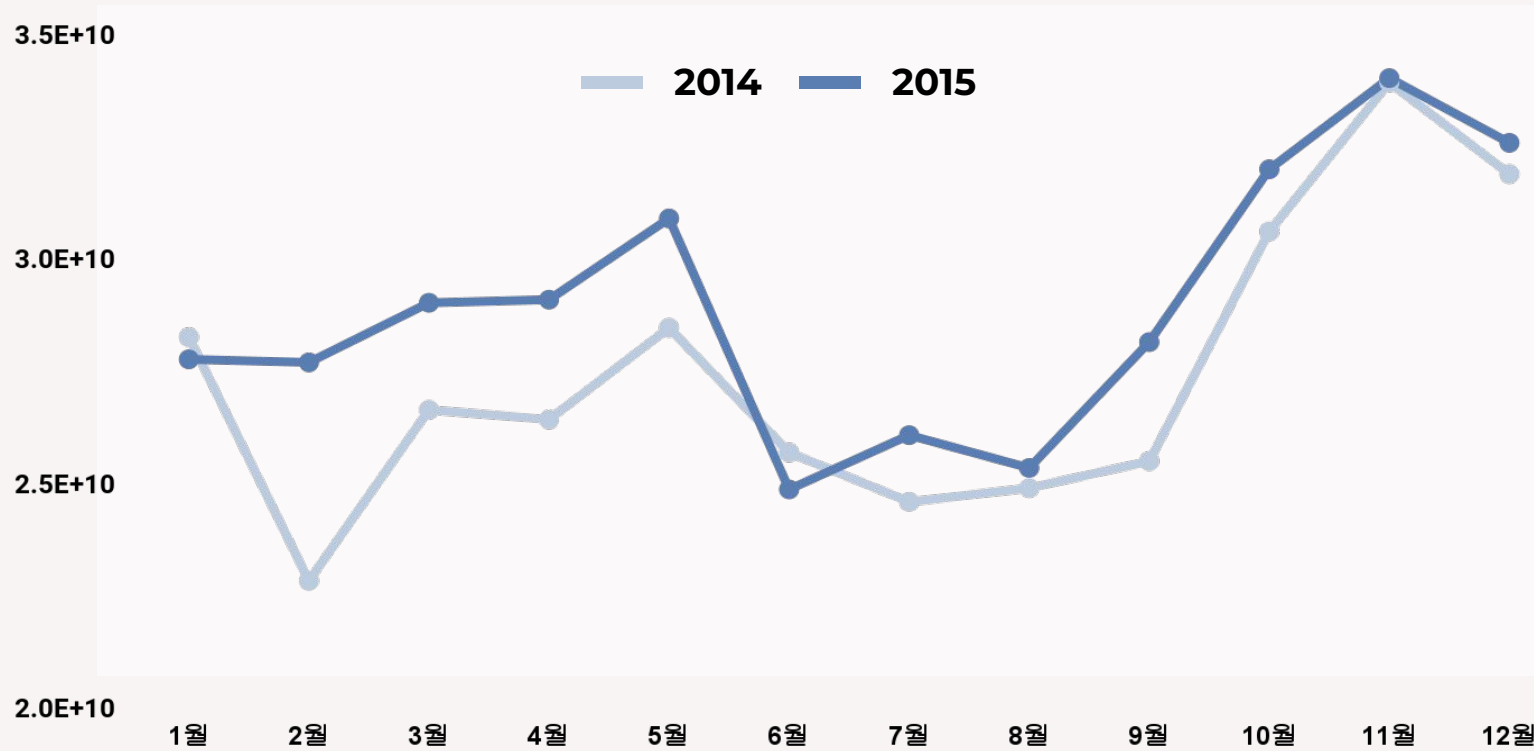
전년 동기 대비 증가폭은 봄이 가장 크고, 여름이 가장 작음

# 탐색적 분석 -테이터 탐색

매출 확인\_연도별/월별 매출

UNIVERSE

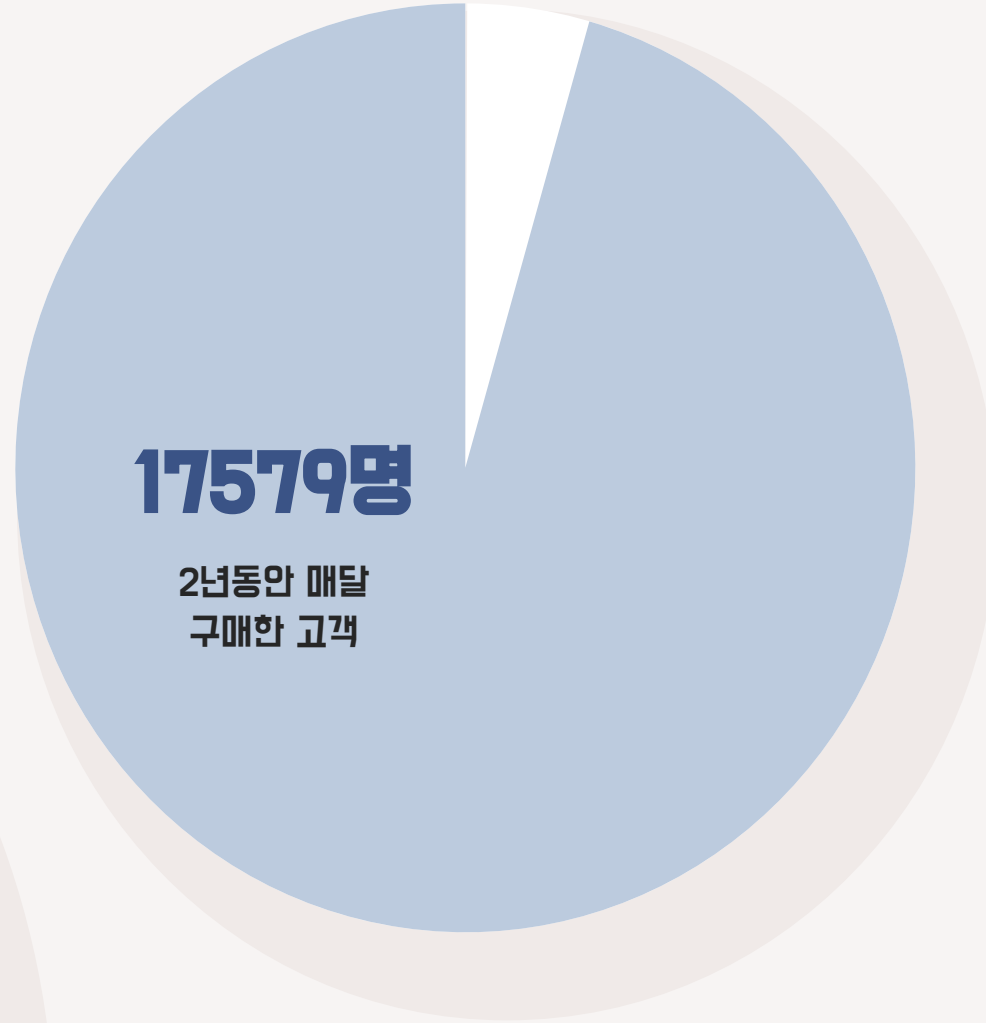
## 연도/월별 구매액 변화



계절적 특징을 보임

# 탐색적 분석 \_데이터 탐색

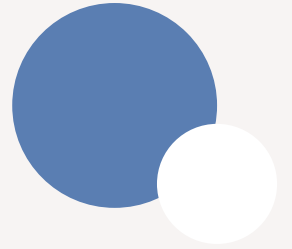
월 단위 유지 고객 확인



전체 고객(19383)의

91%

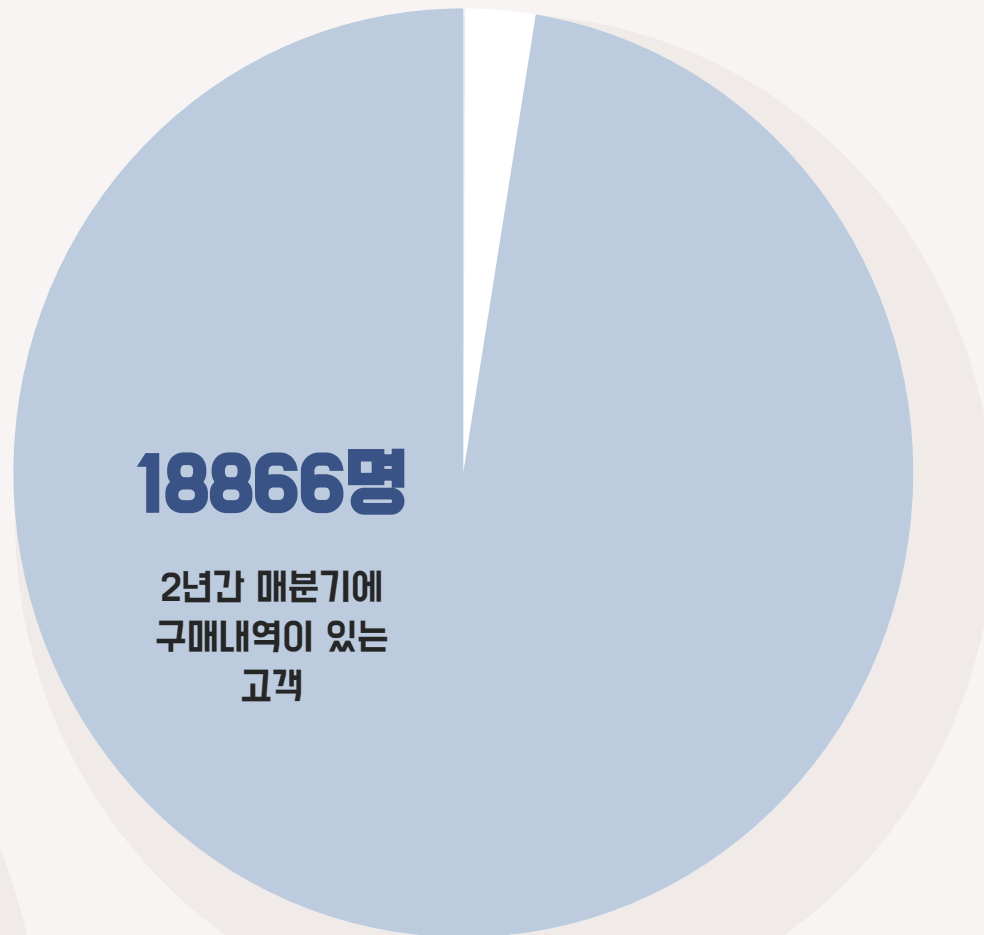
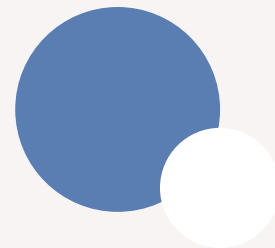
UNIVERSE



# 탐색적 분석 - 데이터 탐색

분기 단위 유지 고객 확인

UNIVERSE



전체 고객(193,833)의

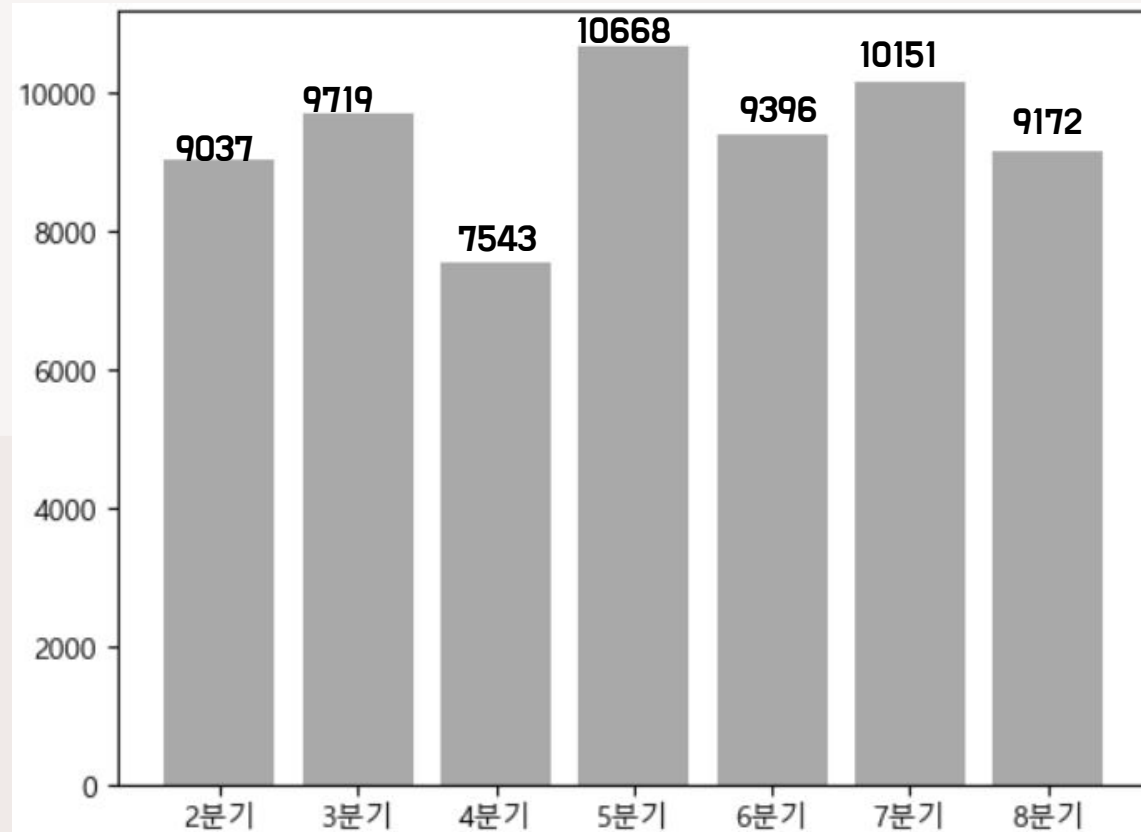
97%

분석 모델 단위 "유지 고객"으로 정의

# 탐색적 분석 \_테이터 탐색

분기별 이전분기 대비 감소고객 수 확인

UNIVERSE

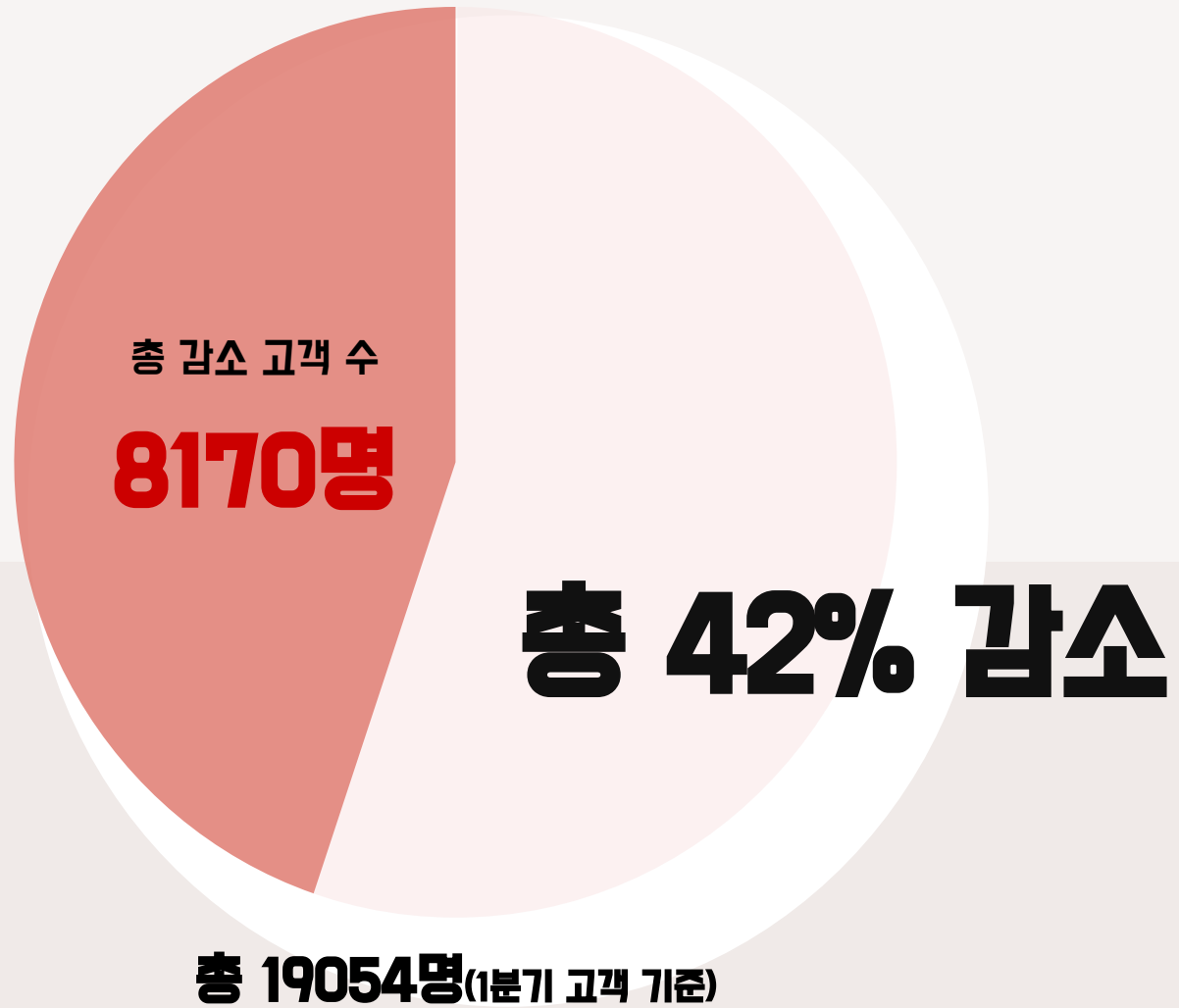




# 탐색적 분석 \_테이터 탐색

1분기 대비 8분기 감소고객 수 확인

UNIVERSE



# 탐색적 분석 - 데이터 탐색

감소고객 정의

UNIVERSE

## 감소고객이란?

$$\text{1분기 구매 금액} \times \text{*평균 증감률} = \text{기준금액}$$

7분기 기준금액

실제 7분기 구매금액

### 기준금액 ▼

평균 증감율을 고려한 기준금액보다  
적은 구매를 한 고객

$$\text{*평균 증감률} = \frac{\text{7분기 구매 금액}}{(\text{1분기 구매 금액} + \text{7분기 구매 금액})}$$

## 주제 선정

6분기 테이터로 다음 분기의 감소고객을 예측하여  
감소고객의 구매패턴을 확인 후  
**매출 증가**를 위한 추천 시스템 활용

# 탐색적 분석 \_ 변수 설정

## 대대분류 카테고리 생성

UNIVERSE

1

소분류별로 각 min, max, mean 매출액 표시 (구매 고객 데이터)

고객번호	제휴사	대분류명	중분류명	소분류명	구매금액_ MIN	구매금액_ MAX	구매금액_ MEAN
4294	D	잡화	양말류	여성 타이즈	2950	7450	5338.9
4295	D	잡화	여행용품	여행용 소품	3430	230000	44773.9
4296	D	잡화	우산/양산 류	우비	17400	58000	39371.4
4297	D	잡화	우산/양산 류	장우산	4000	29000	10140.6

# 탐색적 분석 - 변수 설정

## 대대분류 카테고리 생성

UNIVERSE

2

소분류별로 금액타입 표시 (저가, 중가, 고가\_분류로직 데이터)

	제휴사	대분류명	중분류명	소분류명	금액타입	단가	평균 금액
0	D	잡화	양말류	여성 타이즈	저가	5539	138394
50	D	잡화	여행용품	여행용 소품	저가	44774	138394
100	D	잡화	우산/양산 류	우비	중가	11330	138394
200	D	잡화	우산/양산 류	장우산	저가	10742	138394

# 탐색적 분석 \_ 변수 설정

## 대대분류 카테고리 생성

UNIVERSE

3

대분류, 금액 타입 기준으로 분류  
→ 총 13개의 대분류, 3개의 금액타입  
(총 39개의 종류로 나뉨)

### INPUT

```
category_sort_need['대분류명_x'].unique() #롯데에서 나눈 금액 타입
```

### OUTPUT

```
array(['식품', '멀티샵', '잡화', '이미용', '아동', '의류', '숙웃', '레저취미', '가전',  
      '주방', '가구인테리어', '교육문화', '생활', '웨딩', '침구'], dtype=object)
```

### 3

대분류, 금액 타입 기준으로 분류  
→ 총 13개의 대분류, 3개의 금액타입  
(총 39개의 종류로 나뉘짐)

#### INPUT

```
category_sort_need['대분류명_y'].unique() #웨딩, 속옷이 빠짐 (구매 고객 데이터)
```

#### OUTPUT

```
array(['식품', '멀티샵', '잡화', '이미용', '아동', '의류', '레저취미', '가전',  
      '주방', '가구인테리어', '교육문화', '생활', '침구'], dtype=object)
```

# 탐색적 분석 \_ 변수 설정

## 대대분류 카테고리 생성

UNIVERSE

### 1. 소분류 개수

제휴사별

- 대분류 카테고리
- 금액타입별 소분류 개수

### 2. 제휴사별 매출

제휴사별

- 대분류 카테고리
- 금액타입별 매출 평균값

### 3. 대분류별 매출

대분류 카테고리  
금액타입별

매출액 mean,  
max, min값

예시)

금액 타입	고가	저가	중가	고가	저가	중가	고가	저가	중가	저가	중가
대분류명_y											
가구인테리어	10.0	NaN	NaN	NaN	42.0	8.0	NaN	NaN	NaN	NaN	NaN
가전	6.0	NaN	7.0	8.0	93.0	13.0	NaN	57.0	NaN	2.0	NaN
교육문화	NaN	NaN	2.0	21.0	51.0	87.0	NaN	NaN	NaN	NaN	NaN
레저취미	55.0	NaN	79.0	NaN	128.0	35.0	NaN	100.0	NaN	3.0	NaN
멀티샵	NaN	NaN	9.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
생활	12.0	NaN	NaN	24.0	18.0	183.0	9.0	23.0	72.0	3.0	24.0
식품	35.0	NaN	87.0	214.0	333.0	1000.0	70.0	200.0	340.0	35.0	7.0
아동	NaN	NaN	14.0	NaN	48.0	8.0	NaN	NaN	NaN	NaN	NaN
의류	27.0	3.0	165.0	NaN	115.0	6.0	NaN	38.0	NaN	NaN	NaN
이미용	12.0	NaN	NaN	4.0	97.0	9.0	NaN	15.0	NaN	55.0	9.0
잡화	26.0	22.0	33.0	NaN	183.0	19.0	NaN	1.0	NaN	7.0	NaN
주방	15.0	NaN	10.0	NaN	43.0	10.0	NaN	62.0	NaN	NaN	NaN
침구	10.0	NaN	NaN	NaN	25.0	20.0	NaN	NaN	NaN	NaN	NaN

소분류 개수, 매출 mean, max, min 값에서 공통적인 특징을 가진 대분류끼리 join



# 탐색적 분석 \_ 변수 설정

대대분류 카테고리 생성

UNIVERSE

## 6개의 대대분류 카테고리 생성



# UNIVERSE

- 분석 기간 단위: 분기
- 발생 문제점: 계절별 구매금액 차이 패턴이 존재함

계절별 차이 제거를 위한 연도/분기/카테고리/금액타입별 구매 금액 평균을 구해서 가중치를 계산하여 원본 데이터에 적용하는 과정 진행

## 1. 분기/카테고리 및 금액타입별 계산된 가중치

BBCL_PRICE_TYPE	가구인	가구인	가구인	가전/	가전/	가전/	교육문	교육문	교육문	식품/	식품/	식품/	의류/	의류/	의류/	침구/	침구/	침구/
	테리	테리	테리	레저취	레저취	레저취	화/아	화/아	화/아	식품/	식품/	식품/	의류/	의류/	의류/	침구/	침구/	침구/
	어/잡	어/잡	어/잡	미/멀	미/멀	미/멀	동_고	동_저	동_중	생활_	생활_	생활_	이미용	이미용	이미용	주방_	주방_	주방_
	화_고	화_저	화_중	티삽_	티삽_	티삽_	가	가	가	고가	저가	중가	_고가	_저가	_중가	고가	저가	중가
	가	가	가	고가	저가	중가												
1	1.1332	1.0382	1.1037	1.1826	1.0626	1.0564	1.1646	0.9368	1.0426	0.9261	1.0345	0.9971	1.0050	1.0633	1.0280	1.1741	0.9681	1.0510
2	1.0104	1.0996	0.9170	0.9806	0.9719	0.9665	1.1419	1.0480	0.9882	1.1664	0.9965	1.0032	1.0615	1.0036	1.0459	0.9192	0.9973	0.9984
3	1.1289	1.1769	1.0824	1.1330	1.0144	1.2749	1.1817	1.0512	1.1682	0.8745	0.9745	0.9473	1.1819	1.0515	1.2950	1.0934	0.9084	1.0767
4	0.8051	0.7827	0.9264	0.7987	0.9577	0.8102	0.7046	0.9736	0.8526	1.0877	0.9963	1.0587	0.8218	0.8992	0.7699	0.8727	1.1581	0.8943

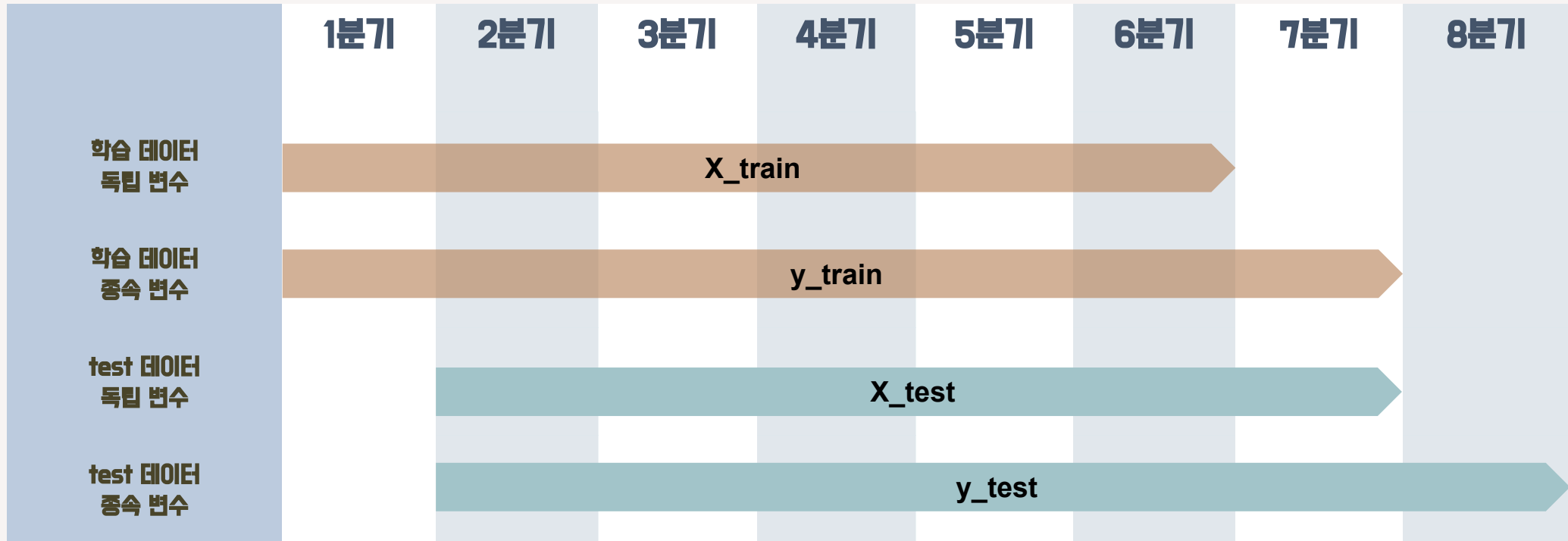
**예시) 2014년도 1분기**

## 2. 가중치가 곱해진 '분기별 구매데이터 생성'

[illegible]

# 모델 학습 예측 & 평가

## TARGET 선정



# 모델링 \_감소고객 예측 모델 생성

Feature 선정

UNIVERSE

## 고객 정보

1. 연령 5세 단위
  2. ENGEL 지수
    - 소득을 추정하기 위한 지수.
- 식품 카테고리 구매액/총 구매액

## 주말 비율

train, test data의 분기 기간을  
기준으로 주말 비율을 산정

## Monetary\_추세선 기울기

train, test data의 각각의  
분기 기간 동안의 카테고리별  
구매 금액 추세선의 기울기 산정

## Frequency\_추세선 기울기

train, test data의 각각의  
분기 기간 동안의 카테고리별  
구매 횟수 추세선의 기울기 산정

# 모델링 \_감소고객 예측 모델 생성

Feature 선정

UNIVERSE

## Monetary\_변동폭, 절대 변동폭

### Monetary 등급 기준

상위 1%	1등급
상위 5%	2등급
상위 10%	3등급
상위 20%	4등급
상위 40%	5등급
상위 70%	6등급
상위 100%	7등급
구매 없음	8등급

1. 고객별로 분기·카테고리별 구매 금액에 대한 등급을 정의

2. 2분기 간의 등급의 차이를 계산

1) 일반적인 차이의 계산 방법

2) 차이에 대해 절댓값을 씌운 방법

예시)

계산 방법	Monetary_가전/ 레저취미/멀티샵_ 고가_등급	일반 차이 계산	절댓값 씌운 방법
1분기	6	-1	1
2분기	5		

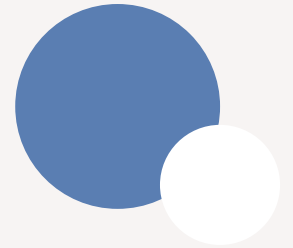
3. 6분기 간의 변동의 합 구하기

- 최근 분기일수록 가중치 높게 부여

# 모델링 \_ 감소고객 예측 모델 생성

최적의 모델 및 파라미터 선정

UNIVERSE



독립변수 : 1분기 ~ 6분기 데이터

종속변수 : 1분기 대비 7분기 증감여부 \_ 8 : 2 비율로 train, test 분리

모델	정확도	AUC	F1
<b>Light Gradient Boosting Machine</b> <small>best_params = {n_estimators : 270, learning_rate : 0.05, max_depth : -1, min_child_samples : 66, num_leaves : 70, reg_lambda : 0.05}</small>	<b>0.74</b>	<b>0.81</b>	<b>0.73</b>
<b>Gradient Boosting Classifier</b>	<b>0.73</b>	<b>0.80</b>	<b>0.71</b>
<b>Random Forest Classifier</b>	<b>0.72</b>	<b>0.79</b>	<b>0.71</b>
<b>Extreme Gradient Boosting</b>	<b>0.72</b>	<b>0.79</b>	<b>0.70</b>

# 모델링 \_ 감소고객 예측 모델 생성

모델 생성 후 예측 평가

UNIVERSE

**train 데이터**

**독립 변수: 1분기 ~ 6분기 데이터**  
**종속 변수 : 1분기 대비 7분기 증감여부**

**test 데이터**

**독립 변수: 2분기 ~ 7분기 데이터**  
**종속 변수 : 2분기 대비 8분기 증감여부**

“ **Light Gradient Boosting Machine** 모델 학습

예측 정확도: 약 73%, AUC : 약 79%, F1 : 약 73%

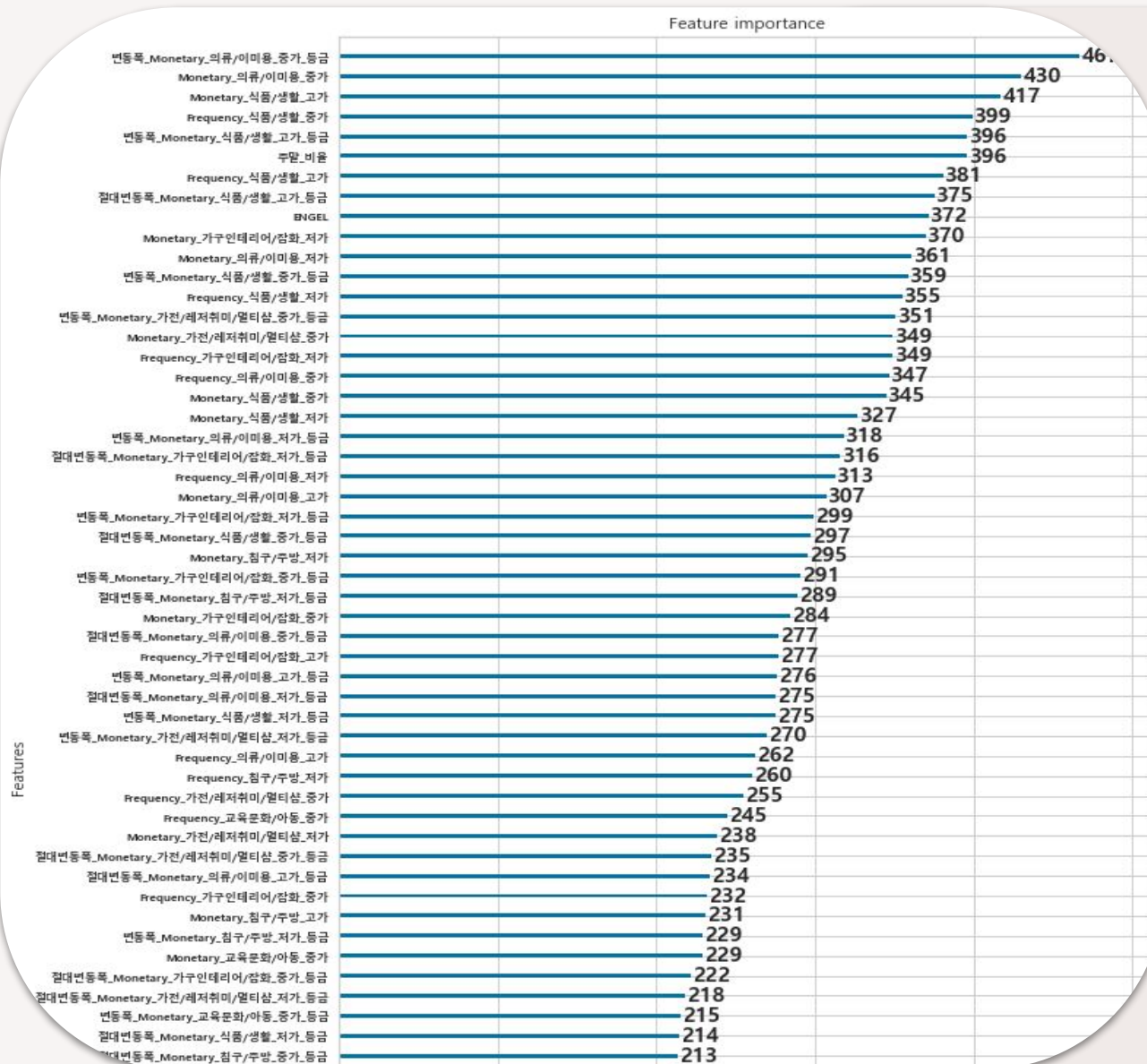
”



# 모델링 \_ 감소고객 예측 모델 생성

## LGBM 변수 중요도

UNIVERSE



# FEATURE IMPORTANCE

모델링 \_감소고객 예측 모델 생성

LGBM 변수 중요도

고객 속성	주말 비율	Frequency _추세선	Monetary_추세선	변동폭_등급	절대변동폭_등급
연령 5세 단위		식품/생활_저가	인류/이미용_저가	식품/생활_중가	식품/생활_고가
성별		식품/생활_중가	인류/이미용_중가	식품/생활_고가	
ENGEL		식품/생활_고가	식품/생활_저가	가전/레저취미/멀티샵 _중가	
		가구인테리어/잡화 _저가	식품/생활_중가	인류/이미용_중가	
		인류/이미용_중가	가구인테리어/잡화 _저가		
			가전/레저취미/멀티샵 _중가		

# 모델링 \_최종 감소고객 예측

훈련 모델 적용

UNIVERSE

```
df['predict_label'] = lgbm.predict("3분기 ~ 8분기 독립변수 데이터")
```

3~8분기 독립변수

학습한 모델로 예측값 출력

3분기 대비 9분기  
증감여부 예측



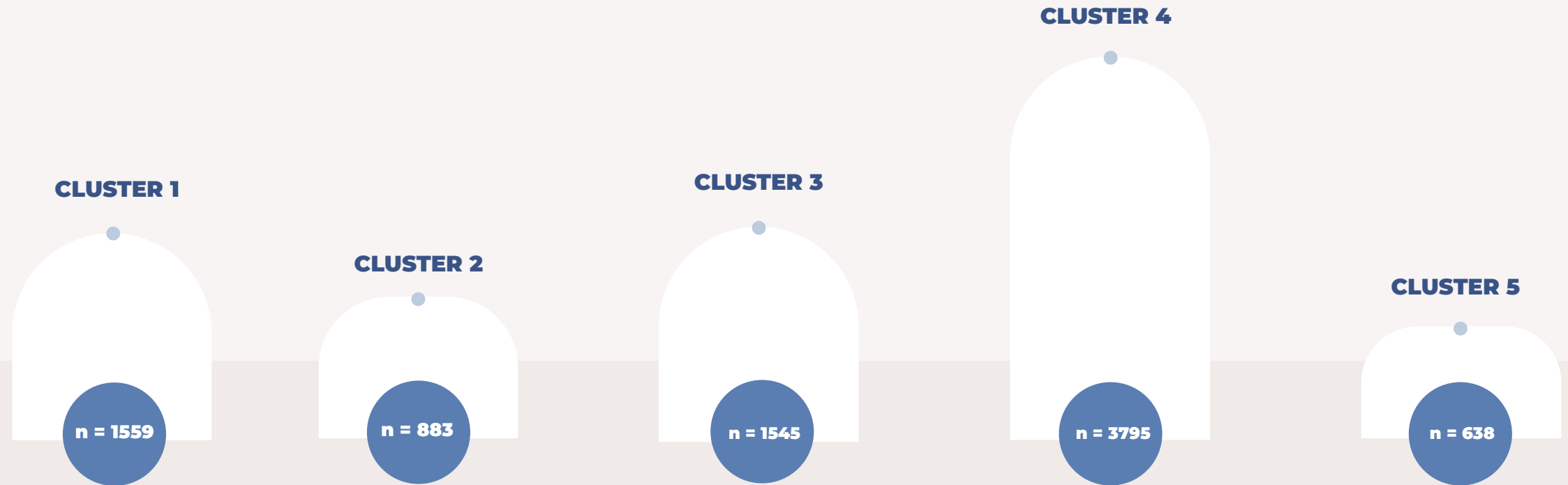
# 군집화

## \_마케팅 제언

# 구매감소고객 군집화

UNIVERSE

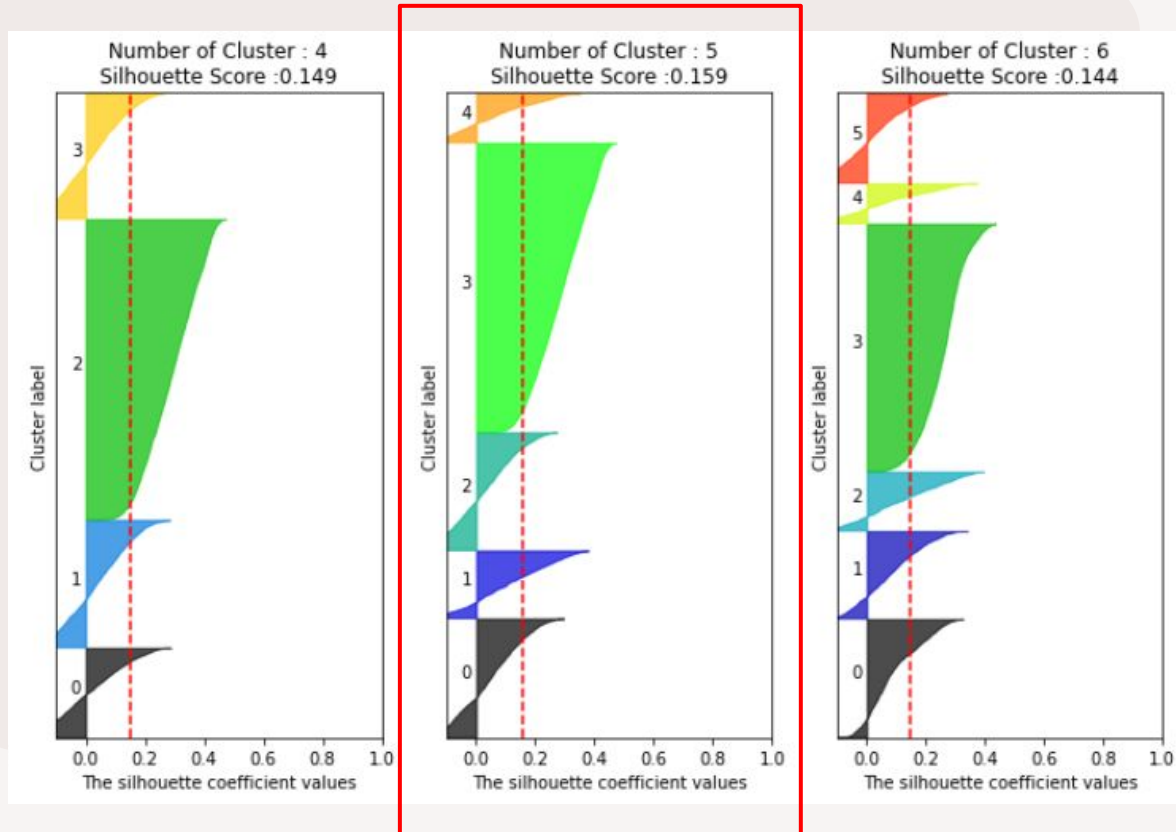
유지고객 18,866명 중 3-8분기 데이터 기반 감소 예측 모델 적용 시  
감소 예측으로 분류된 고객 8,150명 대상으로 군집화



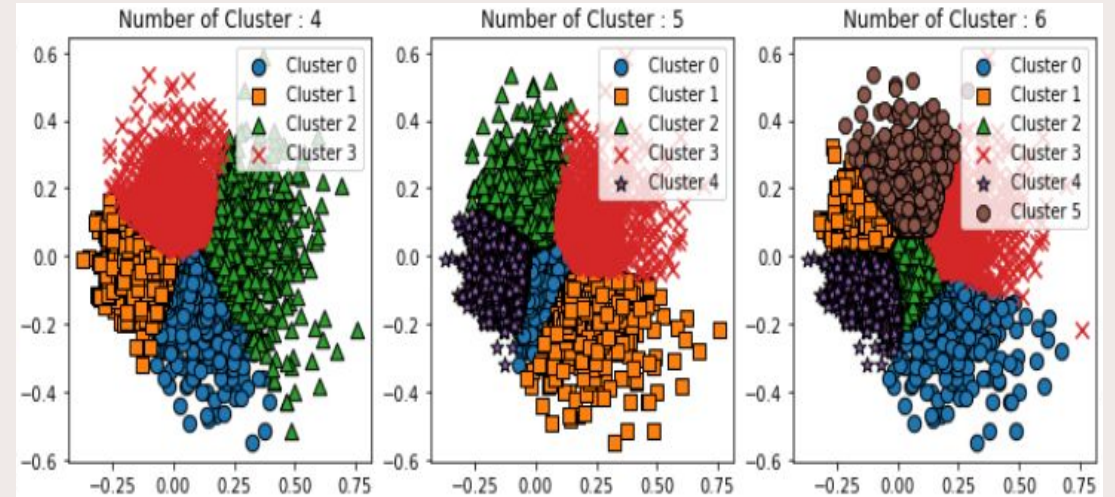
# 구매감소고객 군집화

## 군집 평가 - 실루엣계수

UNIVERSE



## PCA를 활용하여 2차원으로 축소후 군집 시각화 비교



## Cluster 1 (n = 1559)

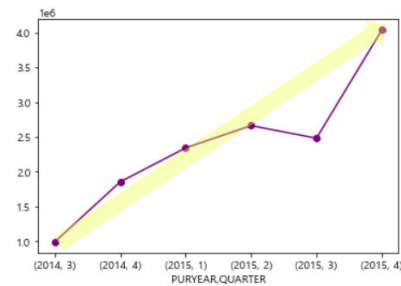
28%

평균 주말 구매 비율

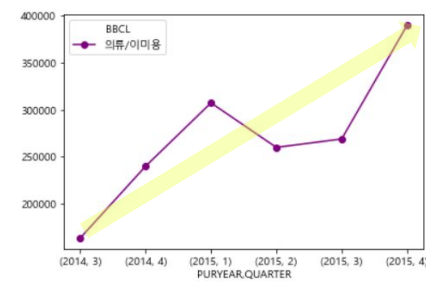
38.20

엔겔 지수

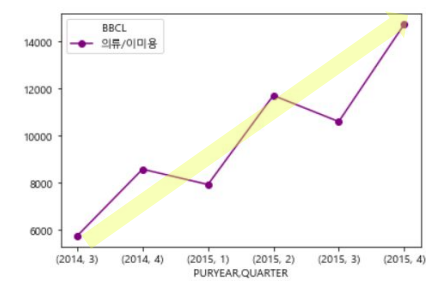
고객별 평균 구매금액



해당 분기 평균 구매금액



전체 고객 총 구매 횟수



의류/이미용\_증가

# 구매감소고객 군집화\_군집별 추천 마케팅 제언

UNIVERSE

## Cluster 1 (n = 1559)

28%

평균 주말 구매 비율

38.20

엔겔 지수

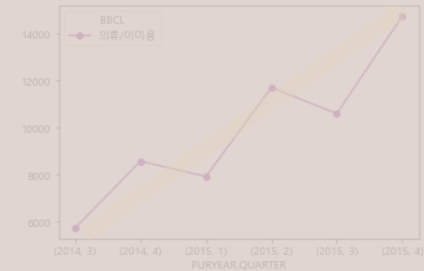
인류/이미용\_증가  
상품에 대한 관여도 높음

개인의 과거 구매 패턴을 고려하여  
인류/이미용\_증가 상품 중



새로운 것을 주기적으로 추천

전체 고객 총 구매 횟수



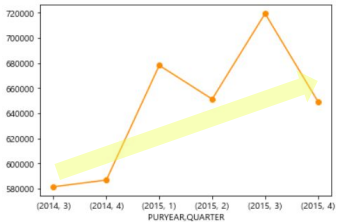


Cluster 2 (n = 883)

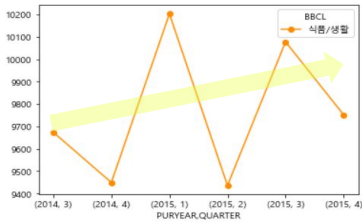
53%

평균 주말 구매 비율

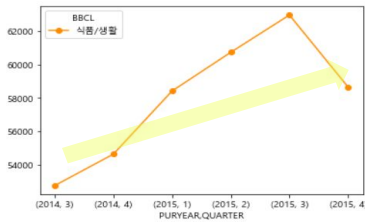
고객별 평균 구매금액



해당 분기 평균 구매금액

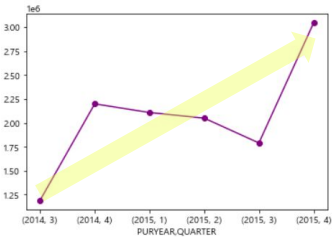


전체 고객 총 구매 횟수

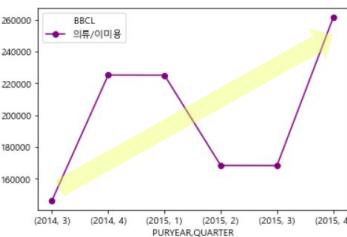


식품/생활\_증가

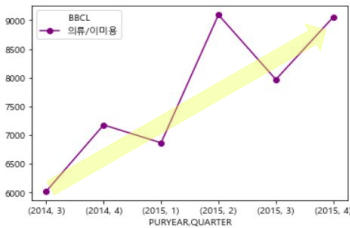
고객별 평균 구매금액



해당 분기 평균 구매금액



전체 고객 총 구매 횟수



의류/이미용\_증가

27.05

영끌 지수

## Cluster 2 (n = 883)

### 식품/생활\_증가

구매액, 구매횟수 모두 감소하므로  
증가를 위한 구매로 이어질 수 있도록  
기준에 구매한 이력이 있는 상품 중  
구매횟수가 잦았던 상품을 추천 적용

### 인류/이미용\_증가

8분기에 구매액, 구매횟수  
모두 증가 추세이므로  
매출 증대로 이어지기 위해  
구매 이력이 없는 상품을 추천 적용

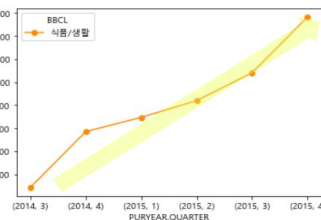
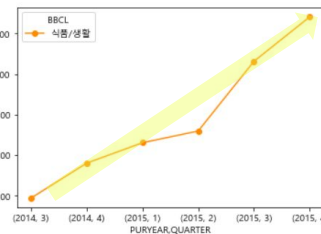
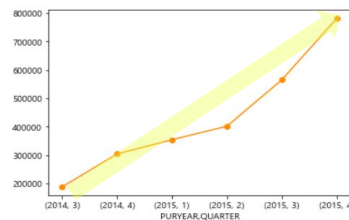
## Cluster 3 (n = 1545)

26%

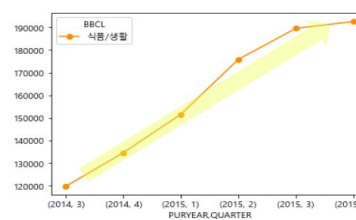
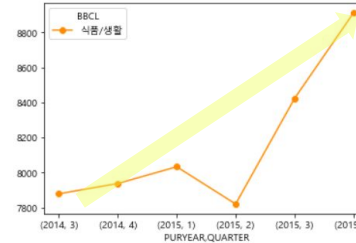
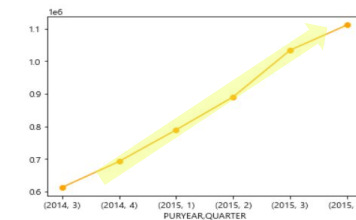
평균 주말 구매 비율

51.26

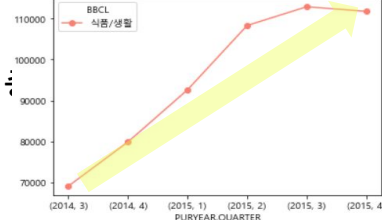
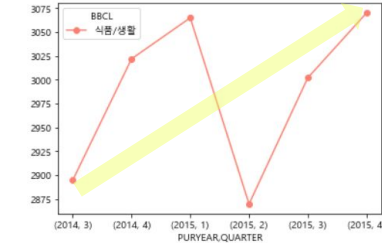
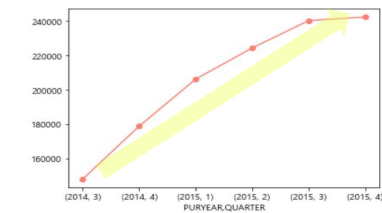
영끌 지수



식품/생활\_고가



식품/생활\_증가



식품/생활\_저가

# 구매 감소고객 군집화\_군집별 추천 마케팅 제언

UNIVERSE

## Cluster 3 (n = 1545)



26%

평균 주말 구매 비율

51.26

식품 고가, 중가, 저가에 대해 미구매/구매 상품들을 다양하게 추천하는 것으로 함

### '식품 레저화'

고객들이 다양한 식품을 소비하고 있음을 보여줌



식품/생활\_고가

식품/생활\_중가

식품/생활\_저가

평균 구매금액  
평균 구매금액  
평균 구매 횟수



## Cluster 4 (n = 3795)

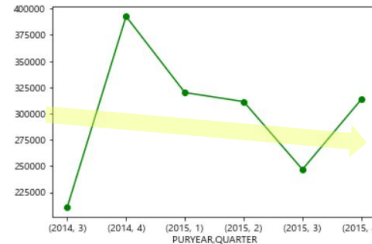
25%

평균 주말 구매 비율

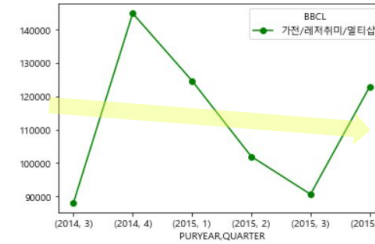
58.83

영끌 지수

고객별 평균 구매금액

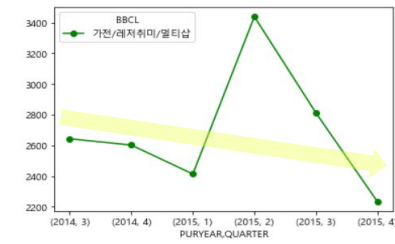


해당 분기 평균 구매금액

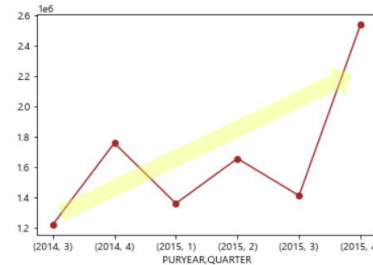


가전/레저취미/멀티샵\_증가

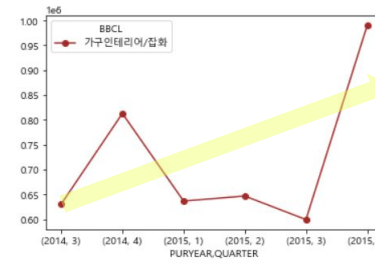
전체 고객 총 구매 횟수



고객별 평균 구매금액

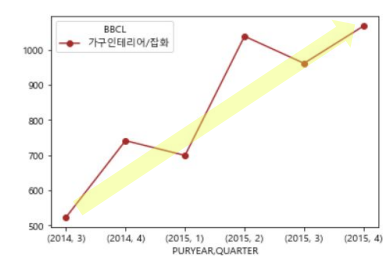


해당 분기 평균 구매금액



가구인테리어/잡화\_고가

전체 고객 총 구매 횟수

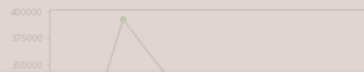


# 구매 감소고객 군집화\_군집별 추천 마케팅 제언

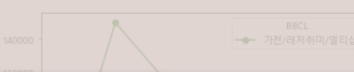
UNIVERSE

## Cluster 4 (n = 3795)

고객별 평균 구매금액



전체 고객 평균 구매금액



전체 고객 총 구매 횟수

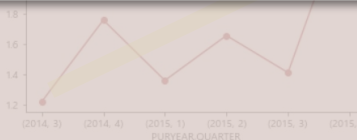


최근 가구인테리어에  
관심도가 높아진 것으로 보임



소비이력이 없는  
새로운 가구인테리어/잡화 제품을 추천

영끌 지수



가구인테리어/잡화\_고가



- 가구인테리어/잡화\_고가  
절대변동폭도 0이 75프로 이상

## Cluster 5 (n = 638)

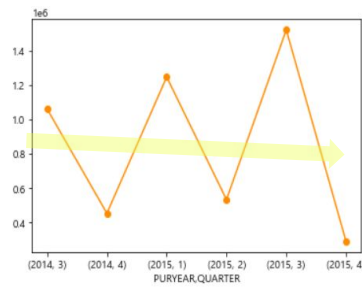
32%

평균 주말 구매 비율

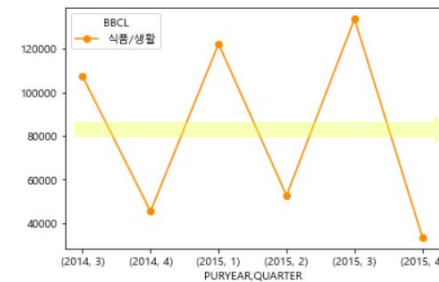
32.37

영끌 지수

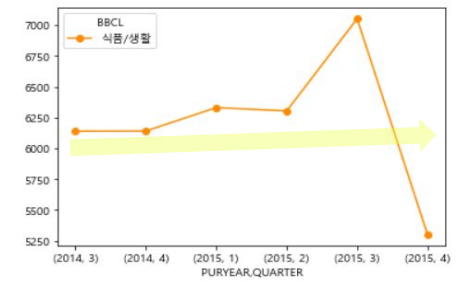
고객별 평균 구매금액



해당 분기 평균 구매금액

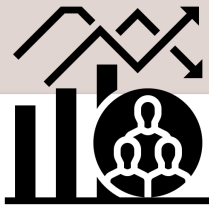


전체 고객 총 구매 횟수



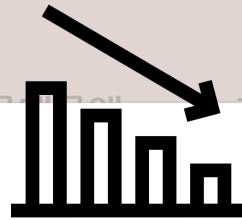
식품/생활\_고가

## Cluster 5 (n = 638)



식품/생활\_고가의 절대 변동폭이  
가장 크게 나타난 군집이며,  
평균 구매 금액의 증감폭이  
가장 크게 나타남

영끌 지수



8분기의 현재 매우 감소세

식품/생활\_고가



증가를 위한 구매로  
이어질 수 있도록  
기존에 소비했던 상품 중  
식품/생활\_고가 상품을  
추천 적용





**surpr!se**

**마케팅 제안  
\_추천 시스템**

# 구매감소를 위한 추천 시스템

SURPRISE API 추천 모델

UNIVERSE

## 고객별 상품별 구매 지수

$$((\text{상품별 구매 건수}) \times \text{상품별 단가 가중치}) / \text{고객별 구매 건수 합}$$

\*상품별 단가 가중치 = 단가/평균단가

CUSTNO	PRD	SCORE
1	A_4대 B/D	0.0
1	A_5 ON THE GO	0.0
1	A_ACC Bloom (1F)	0.0
1	A_ACC Bloom (3F)	0.0
1	A_AK골프	0.0

## 상품 추천 모델

SURPRISE API를 활용한  
잠재요인 협업 필터링 기반의  
개인화된 추천 시스템 모델 학습

## 구매 지수 예측

해당 고객의 미구매/구매 상품의  
구매 지수를 예측

# 구매감소를 위한 추천 시스템

## SURPRISE API 추천 모델

### 고객별 TOP 100개 추천 상품

미구매/구매 상품의 구매 지수를  
내림차순으로 정렬 후 TOP 100개 추천 상품 추출



### 고객을 위한 카테고리별 상품 추천

가장 관심이 높을 것으로 예측한 TOP 100개 추천 상품을  
카테고리별로 분류한 후, 해당 고객 속성 정보 및  
상품별 상세 정보와 함께 추천 결과 출력

### 추천 시스템 결과 예시) 군집 1번의 고객 6번

UNIVERSE

고객 ID	성별	연령대	앵겔지수	총 구매금액	6-8분기 총 구매횟수	
0	6	여자	60세이상	33.5	9,491,804원	1,593회

#### 이미용 추천 제품 ▼

	제휴사	대분류명	중분류명	추천 상품	금액 타입	예측 구매 지수
1	A	이미용	화장품	기초 화장품	고가	0.027408
23	A	이미용	화장품	색조 화장품	고가	0.015143
46	A	이미용	화장품	기초A	고가	0.005459
53	B	이미용	기초화장품	일반화장품	저가	0.004600

#### 생활 추천 제품 ▼

	제휴사	대분류명	중분류명	추천 상품	금액 타입	예측 구매 지수
14	A	생활	생활잡화	단기행사	고가	0.016865
27	A	생활	생활잡화	욕실용품	고가	0.011674
38	A	생활	생활잡화	타월	고가	0.006664

#### 의류 추천 제품 ▼

	제휴사	대분류명	중분류명	추천 상품	금액 타입	예측 구매 지수
15	A	의류	남성 트렌디	단기행사	중가	0.016865
32	A	의류	캐주얼	global SPA	중가	0.010227
51	A	의류	란제리/내의	패션내의	고가	0.005148

# 구매감소를 위한 추천 시스템

## SVD 모델 학습 평가 결과

Cluster1 평가지표

모델 평가	Fold1	Fold2	Fold3	Fold4	Fold5	평균
RMSE(testset)	0.0050	0.0054	0.0052	0.0050	0.0049	0.0051
MAE(testset)	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007
Best_params	n_epochs: 40, n_factors: 200					

Cluster2 평가지표

모델 평가	Fold1	Fold2	Fold3	Fold4	Fold5	평균
RMSE(testset)	0.0067	0.0067	0.0075	0.0062	0.0062	0.0067
MAE(testset)	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
Best_params	n_epochs: 40 , n_factors: 100					

Cluster3 평가지표

모델 평가	Fold1	Fold2	Fold3	Fold4	Fold5	평균
RMSE(testset)	0.0054	0.0054	0.0086	0.0078	0.0050	0.0064
MAE(testset)	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006
Best_params	n_epochs: 40 , n_factors:200					

# 구매감소를 위한 추천 시스템

## SVD 모델 학습 평가 결과

Cluster4 평가지표

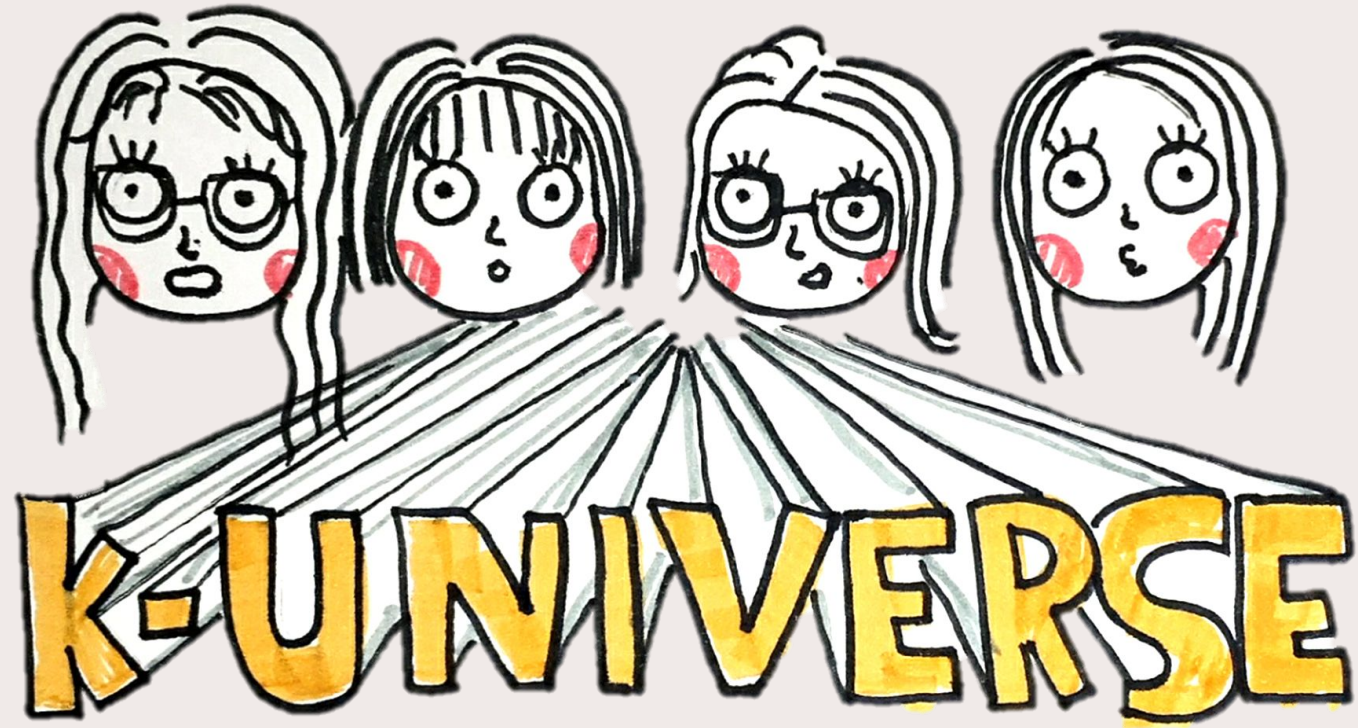
모델 평가	Fold1	Fold2	Fold3	Fold4	Fold5	평균
RMSE(testset)	0.0038	0.0048	0.0050	0.0039	0.0046	0.0044
MAE(testset)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
Best_params	n_epochs: 40, n_factors: 200					

Cluster5 평가지표

모델 평가	Fold1	Fold2	Fold3	Fold4	Fold5	평균
RMSE(testset)	0.0066	0.0065	0.0147	0.0064	0.0070	0.0082
MAE(testset)	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
Best_params	n_epochs: 60, n_factors: 100					



Q&A



감사합니다