

# 구매감소고객 예측 및 맞춤 솔루션

6조 : 최한결 조현정 이종현 배송이

# CONTENTS



## 분석 과제 정의

L사 데이터 탐색  
문제 정의  
과제 도출



## 분석 과제화

데이터 전처리  
변수 정의



## 모델링 및 평가

모델링  
검증 및 평가

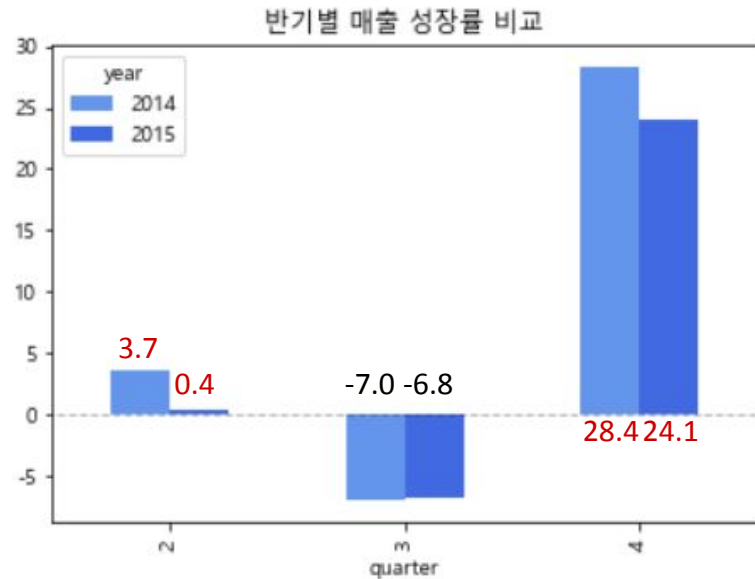
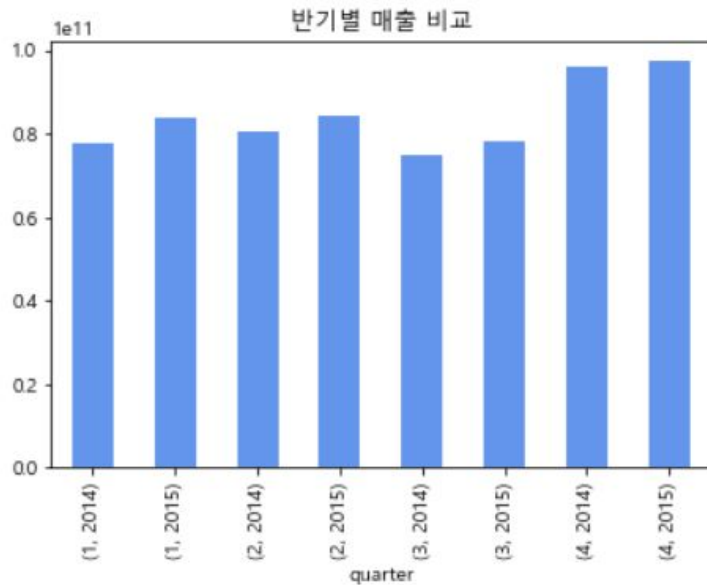


## 솔루션 제시

고객 분류 및 대응방안



# L사 매출 현황 및 문제점 파악



반기별 매출 증가폭이 감소하는 추세

## L사 매출 현황 및 문제점 파악

### 기존고객

2014년, 2015년  
매 반기 구매를 한 고객  
(제휴사 A, B, C 대상)

### 구매감소고객

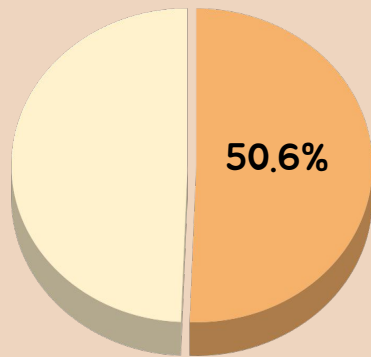
기존고객 기준  
비교기간 매출 평균 성장율 보다  
구매실적 성장율이 하락한 고객을  
구매 감소 고객으로 정의

※ 구매증감 : 개인별증감율/전체증감율

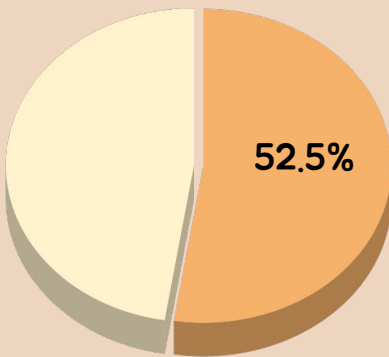
※ 증감율 : (비교대상-기준)/기준

신규/이탈 고객 수 276명 → 전체 고객의 1.42%, 전체 구매 건수의 0.74%

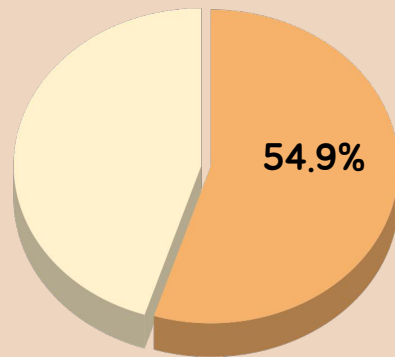
## L사 매출 현황 및 문제점 파악



14년 상반기 -> 14년 하반기



14년 하반기 -> 15년 상반기

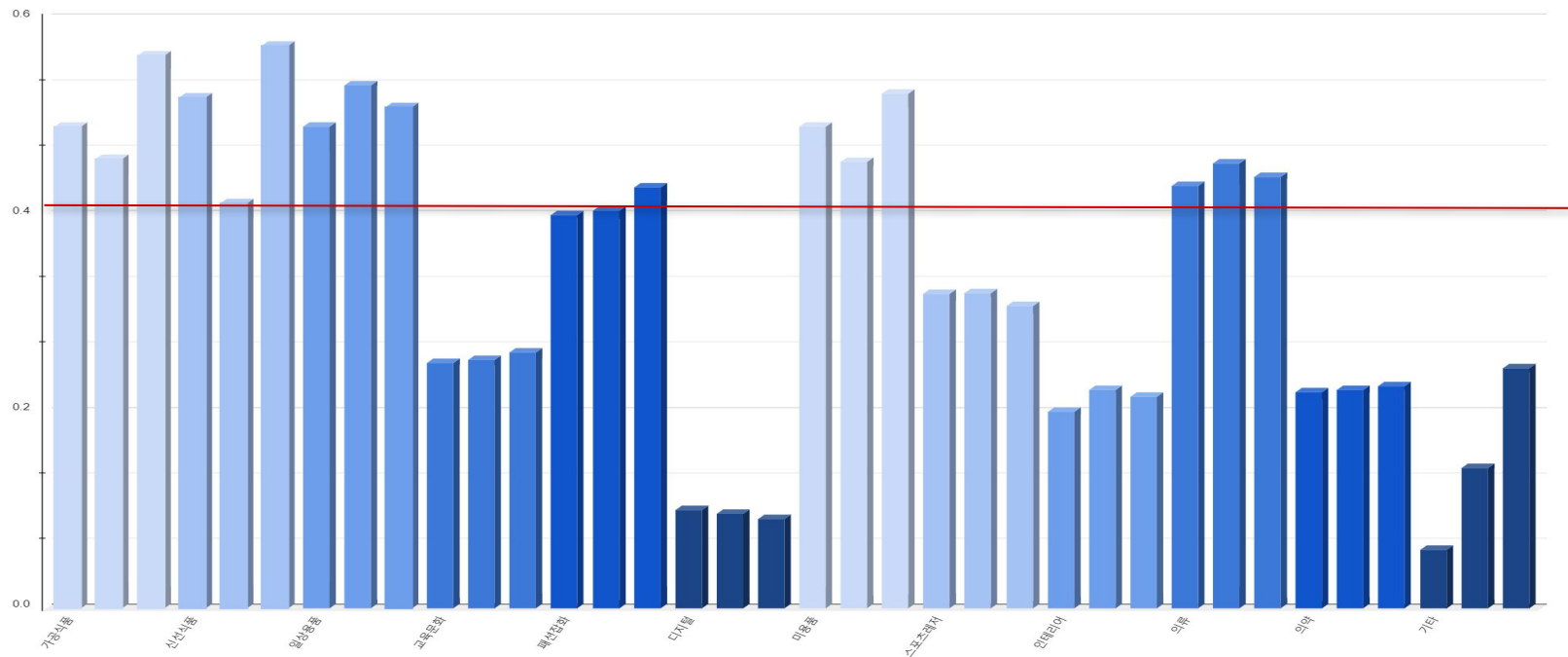


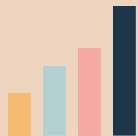
15년 상반기 -> 15년 하반기

**반기별 구매감소고객의 비율이 증가**

# L사 매출 현황 및 문제점 파악

## 품목별 구매감소고객 비율





## 과제 정의

**고객의 구매패턴을 나타내는 요인을 독립변수로 설정하여  
구매감소고객을 예측 및 대응하는 것을 과제로 정의함**

# 데이터 전처리

소분류명	중분류명	대분류명	분류1	상품분류
사이다	탄산음료	가공식품	편의품	가공식품저가
사이다	탄산음료	가공식품	편의품	가공식품저가
사이다	탄산음료	가공식품	편의품	가공식품저가
사이다	탄산음료	가공식품	편의품	가공식품저가
사이다	탄산음료	가공식품	편의품	가공식품저가
...	...	...	...	...
온라인열매채소	온라인버섯/열매채소	신선식품	편의품	신선식품중가
온라인해물연체류	온라인해물연체	신선식품	편의품	신선식품고가
포트형정수기	정수기	디지털	선매품	디지털저가
온라인구색과일	온라인저장과일	신선식품	편의품	신선식품중가
온라인돼지고기	온라인돼지고기	신선식품	편의품	신선식품중가

## 대분류명

- 상품DB의 상품분류기준을 참고하여 상품의 대분류를 재정의

## 분류 1

- 일상품 : 가격평균 10만원 이하
- 선매품 : 가격평균 10만원 이상
- 전문품 : 중분류명이 귀금속, 명품 혹은 수입물품

## 상품분류

- 고가 :  $2 \times (\text{대분류 평균금액}) \leq \text{중분류평균금액}$
- 중가 :  $0.5 \times (\text{대분류 평균금액}) < \text{중분류평균금액} < 2 \times (\text{대분류 평균금액})$
- 저가 :  $\text{중분류평균금액} \leq 0.5 \times (\text{대분류 평균금액})$

※ 상품을 세가지 기준으로 분류 진행



## 변수 정의

### 고객 속성 데이터

성별  
연령대  
거주지역  
channel 이용유무

### 구매데이터 - 증감율

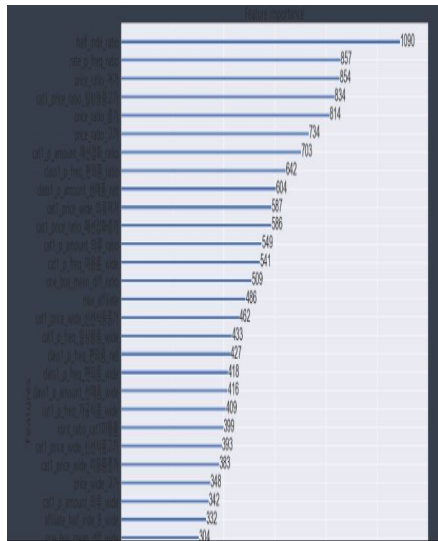
구매횟수증감  
대분류별구매금액증감  
대분류별구매물품수증감  
최대구매제휴사변동  
구매금액증감  
제휴사별 구매금액 증감  
제휴사별 구매횟수 증감  
평균구매금액 증감  
월별구매횟수증감

### 구매데이터 - 변동율

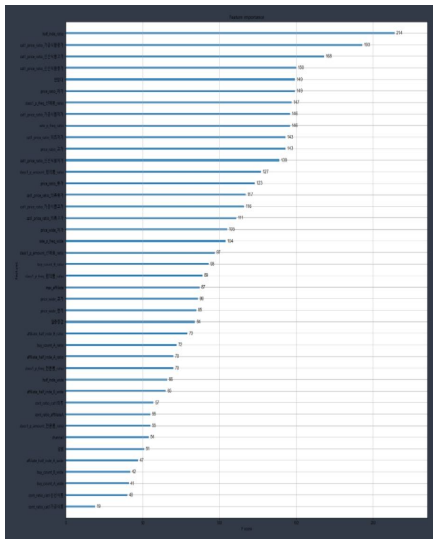
구매횟수변동  
대분류별구매금액변동  
대분류별구매물품수변동  
최대구매제휴사변동  
구매금액변동  
제휴사별 구매금액 변동  
제휴사별 구매횟수 변동  
평균구매금액 변동  
월별구매횟수변동

# 변수 정의

Feature Importance



XGB



DecisionTree

## 최종 독립변수

연령대

대분류별구매금액증감(NET)

대분류별구매금액변동(ABS)

최대구매제휴사변동

제휴사별 구매금액 증감율(A,B)

제휴사별 구매횟수 증감율(B)

월별구매횟수증감율

독립변수 52개 → 42개

※ P\_values < 0.05 조건에 충족하는 변수(LR) 및 Feature Importance가 높은 변수(XGB/DT) 선정

## 변수 정의

### 증감율 범주화

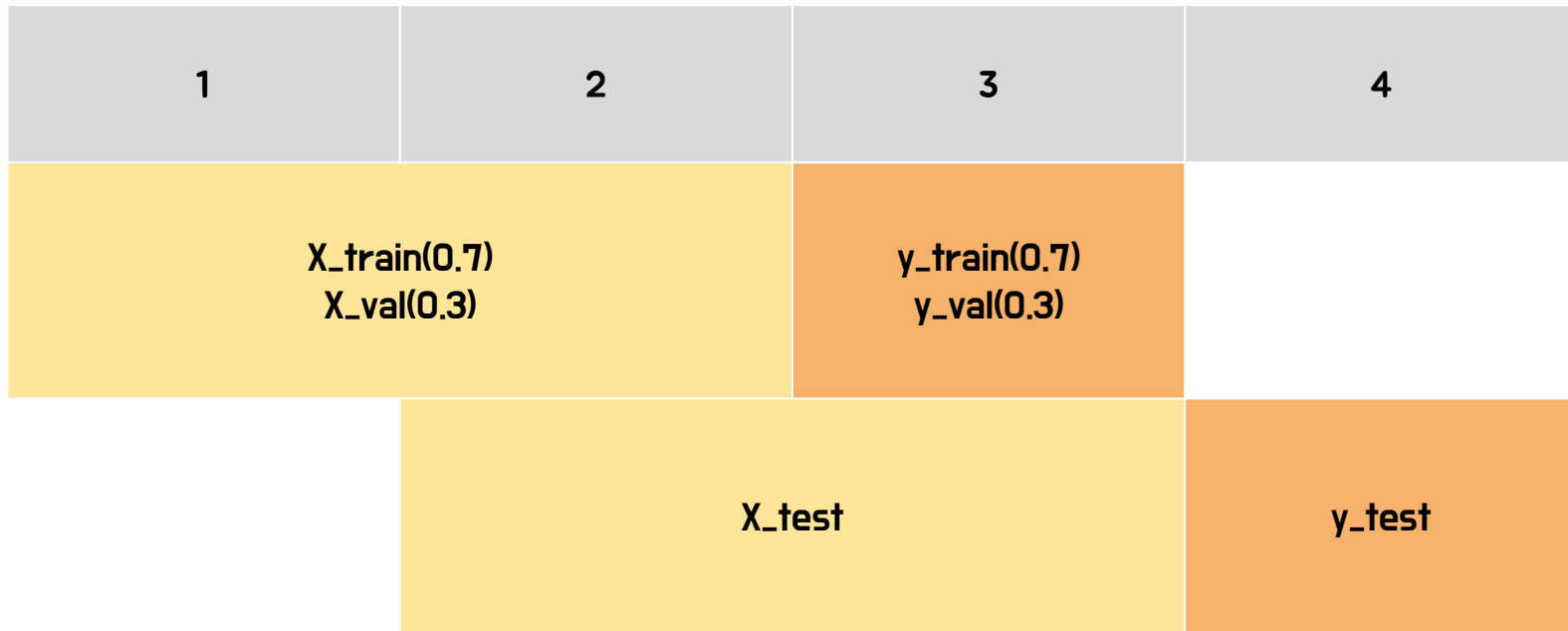
- 감소 : 1 ~ 4
- 증가 : 5 ~ 8, 9(100% 이상 증가)
- 두 반기 모두 0일 경우 : 부정 → 10
- 전반기가 0일 경우 : 불능 → 11

### 변동폭 범주화

- 사분위를 기준 : 1 ~ 4
- 낮은 이상치 = 0
- 높은 이상치 = 5

※ '값'을 기준으로 균일하게 나눠 카테고리화를 진행

# 모델링 - 데이터 분리



※ 1 : 2014년 1~2분기    2 : 2014년 3~4분기    3 : 2015년 1~2분기    4 : 2015년 3~4분기

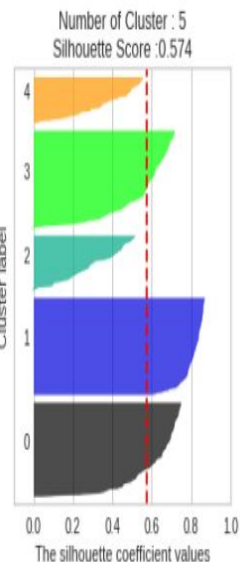
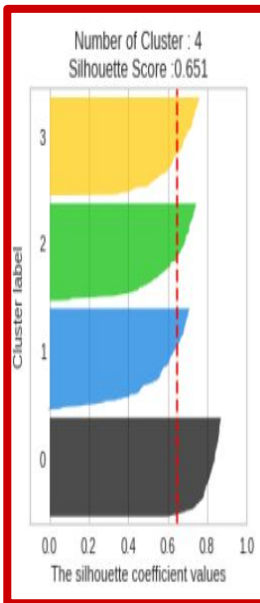
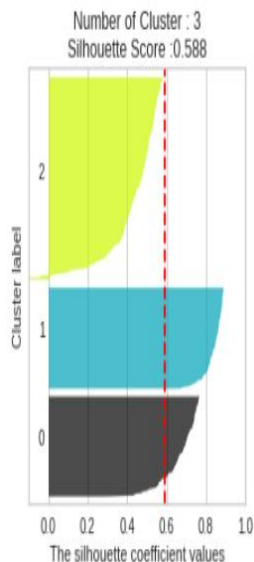
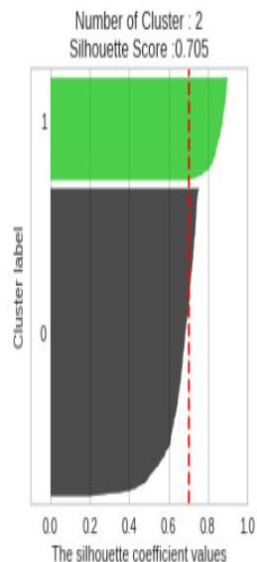
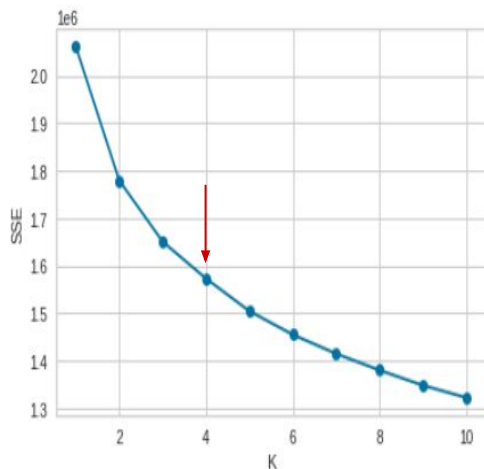
## 모델링 - 검증 및 평가

	LR	RF	DT	XGB
정확도	0.7	0.7	0.62	0.7
정밀도	0.69	0.69	0.61	0.69
재현율	0.7	0.7	0.6	0.7

학습 성능이 가장 좋은 XGB 모델로 선정

# 구매감소고객 분류 및 대응방안

## 최적 군집수 설정



Elbow method, 실루엣계수 → 군집 수 4개가 적절

# 구매감소고객 분류 및 대응방안

A제휴사 비중 62%

A제휴사 비중이 높으나  
평균구매금액이 높지 않은 편

주 구매품목 식품/인류

스포츠레저 감소/인테리어 증가



평균구매금액을 끌어올리기 위해  
식품과 인류에서 묶음상품 행사의  
활용이 용이할 것으로 보임

Cluster0

C제휴사 비중 40%

방문 빈도가 높음

주 구매품목 식품(83%)



방문 빈도가 높으므로,  
일회평균구매금액을 늘릴 수  
있도록 구매상품 연관성 분석후  
상품 추천

Cluster1

A제휴사 비중 64%

멤버십 가입률 46%로 가장 높음

주 구매품목 식품/인류



모바일 마케팅  
고객 모바일 장바구니 분석후  
개인에 맞는 상품 추천

Cluster2

A제휴사 비중 96%

평균구매금액이 높음

주 구매품목 인류(41%)

경쟁사이용 고객비율 40%로 높음



구매금액이 가장 높은 고객인 동시에  
경쟁사 이용률이 높게 보이므로  
고객의 주 구매 품목에 대해  
할인 행사등을 통해 충성률을  
올리는 방안을 추천

Cluster3

# 감사합니다

6조 : 최한결 조현정 이종현 배송이