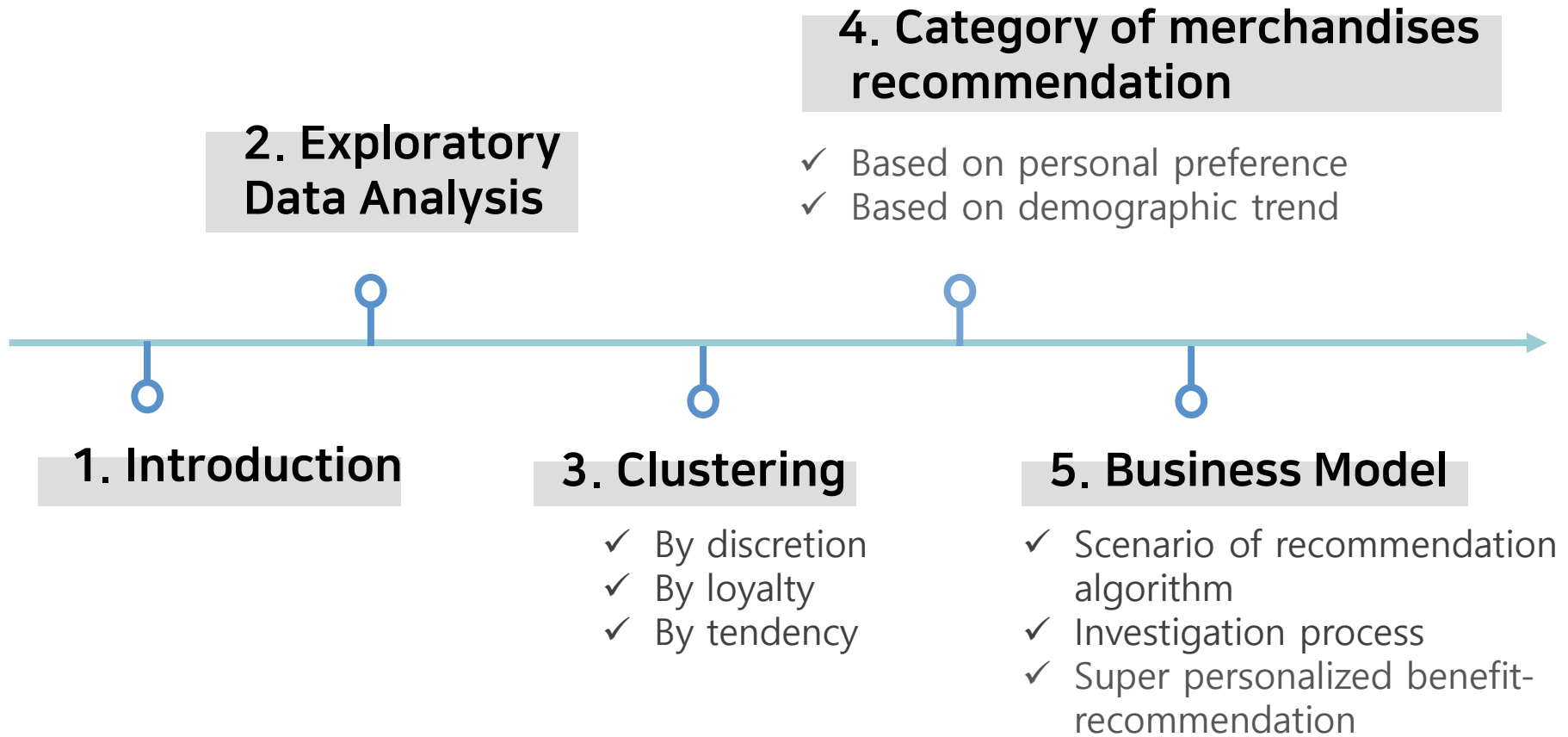


# Super personalized benefit-recommendation based on analysis of customer

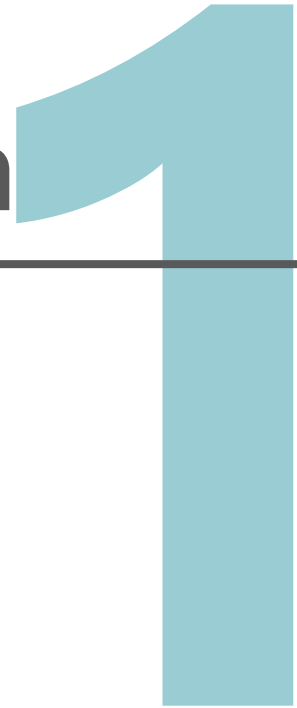


# CONTENTS



# Introduction

---



# Derection of analysis

## Super personalized benefit-recommendation


### Problem

“최근 온라인상의 고객 행동이 복잡해져 감에 따라 시시각각 변하는 고객의 행동정보를 활용하지 못하고 있으며, 선호에 영향을 미칠 수 있는 고객의 성향 정보가 적절하게 반영되지 못하는 문제점이 있습니다.”

### Solution

“이런 문제점들을 해결하고자 기존의 전략에서 더 나아가 저희가 만든 SSTP 전략을 이용하여 세분화된 그룹 속에서 또 한번의 초 개인화 작업을 통하여 최적화된 상품을 추천하고 그 결과를 이용한 개인별 혜택 제공을 해줄 수 있는 알고리즘 을 만들었습니다.”

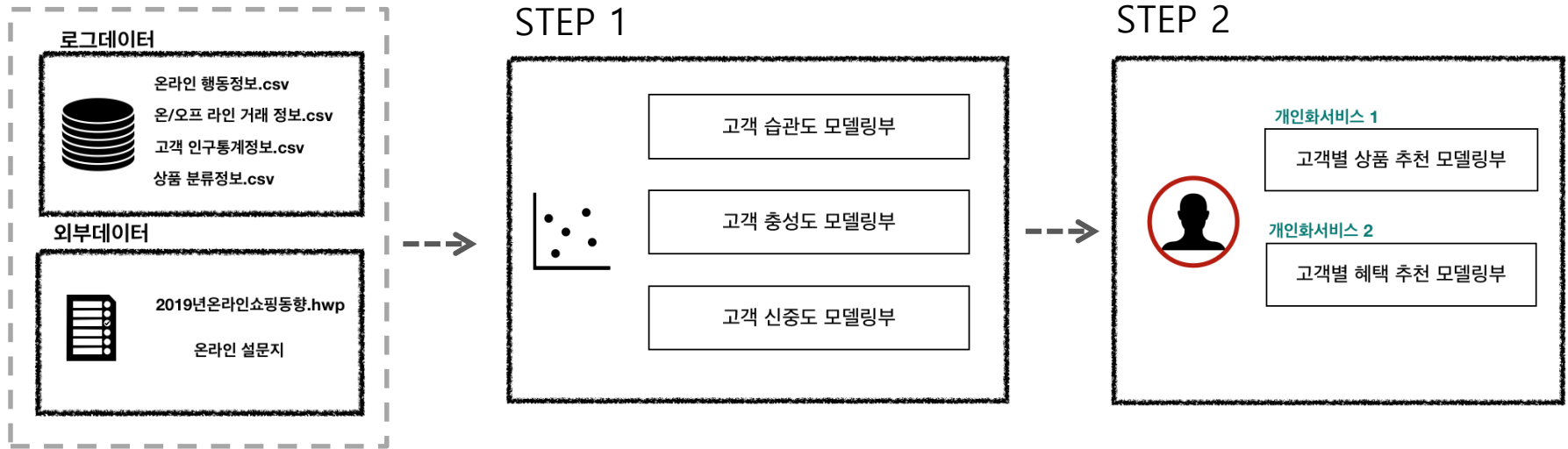
### SSTP (Segmentation Super Targeting Positioning) Strategy

- 
- ✓ Service 1  
Recommend optimized products with personalization algorithms
  - ✓ Service 2  
Provide super-personalized benefits based on one's shopping preferences



# Derection of analysis

## Super personalized benefit-recommendation



### STEP 1: Clustering & Recommendation

제공된 분석용 베포데이터 이외에 통계청에서 제공된 2019년 온라인쇼핑동향과 온라인 설문지를 사용하여 고객을 분류 하였습니다.

다양한 사용자 경험 기반 데이터와 반응 정보를 결합을 함으로써 효과적으로 고객들의 정보를 변수화 하여 기준을 세우고 유저를 그룹 별로 분리하는 단계입니다.

### STEP 2: Service based on personalization

제공된 분석용 베포데이터 이외에 통계청에서 제공된 2019년 온라인쇼핑동향과 온라인 설문지를 사용하여 고객을 분류 하였습니다.

다양한 사용자 경험 기반 데이터와 반응 정보를 결합을 함으로써 효과적으로 고객들의 정보를 변수화 하여 기준을 세우고 유저를 그룹 별로 분리하는 단계입니다.

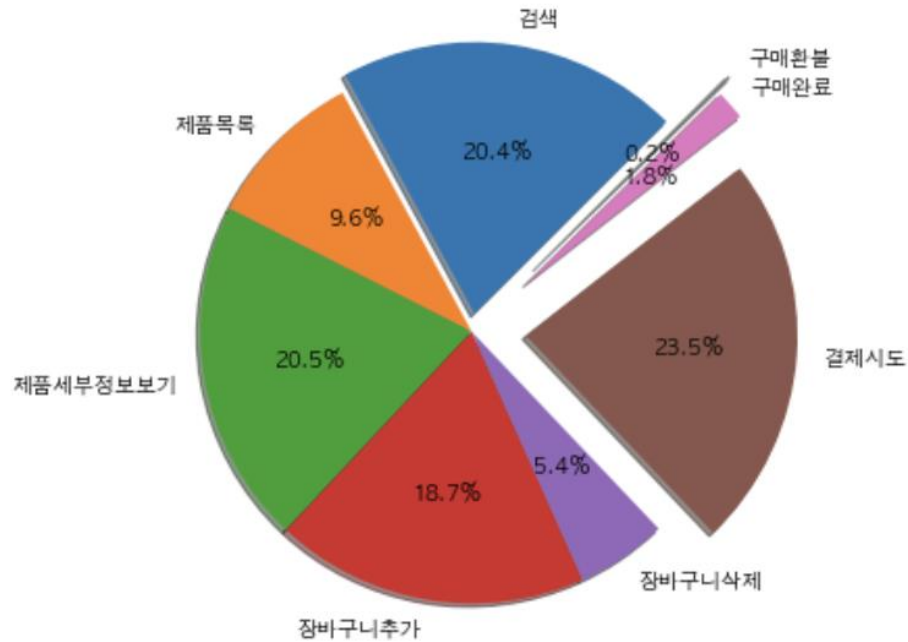
# Exploration data analysis

---



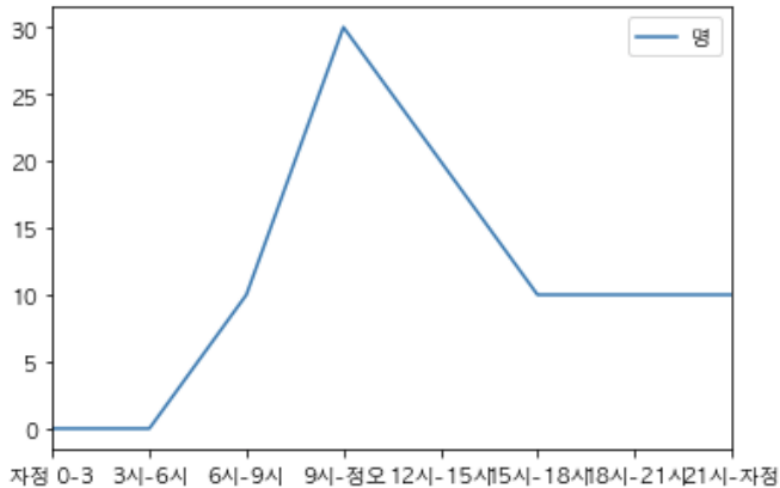
# Exploration data analysis

## ✓ 행동유형 (action\_type)

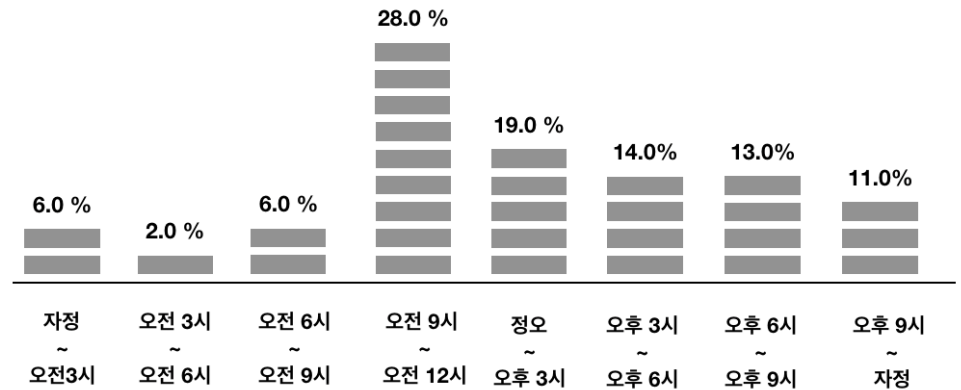


- 구매환불 비율은 너무 적어서 추후 사용하지 않았다.
- 구매완료와 결제시도가 매치 되지 않아서 결제시도를 구매완료와 합쳐서 처리하였다.

# Exploration data analysis



<시간대별 온라인 구매비중>



<시간대별 온라인쇼핑 이용률>

시간대에 따른 온라인 접속과 온라인 구매는 비슷한 양상을 보이고 있음을 알 수 있다  
또한, 오전 9시~오전 12시에 고객들이 몰리고 있어 시간대가 중요한 요소임을 알 수 있다



# Exploration data analysis

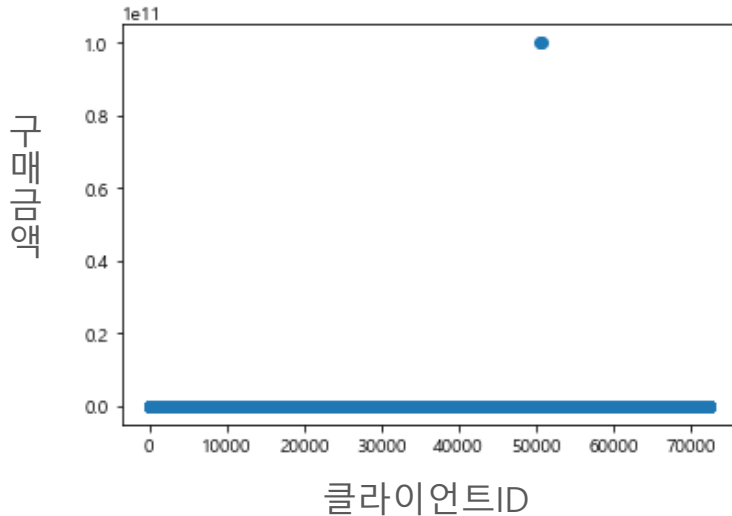
## ✓ 거래정보 (action\_type)

행동유형(action\_type)은 '온라인 고객'의 행동을 8개의 유형을 구분한 것입니다.

클라이언트ID	날짜	상품대분류	수량	가격
53031	2019-07-10 16:39:00	Digiter, Appliances	1	2960000
	2019-07-10 16:55:00	Digiter, Appliances	1	2960000
	2019-07-10 17:20:00	Digiter, Appliances	1	2960000

100만원이 넘는 가격 중 [ 같은 날짜, 상품대분류 수량] 이 똑같다면, 중복값으로 생각하여 삭제하였습니다.

# Exploration data analysis



구매금액이 3달동안 1억원을 넘는  
것은  
이상치라 생각하여 제거

date	clnt_id	biz_unit	clac_nm0	buy_ct	buy_am
2019-09-14 12:07:00	47093	A03	Foods	0	0
2019-08-06 22:22:00	46906	A03	Foods	0	0
2019-07-14 19:01:00	4267	A03	Foods	0	0
2019-07-03 22:42:00	37463	A03	Foods	0	0
2019-09-15 19:48:00	5879	A03	Foods	0	0
...	...	...	...	...	...
2019-08-18 14:18:00	39915	A02	Living, Health	1	0
2019-08-10 21:01:00	54813	A02	Living, Health	1	0
2019-07-09 22:00:00	53569	A02	Living, Health	1	0
2019-08-12 18:37:00	55686	A02	Living, Health	1	0
2019-09-16 20:32:00	2886	A02	Living, Health	1	0



거래 ID는 존재하지만  
구매금액과 구매수량이 0개인 것은  
이상치라 생각하여 제거

# Clustering

---

1. By discretion
2. By loyalty
3. By tendency

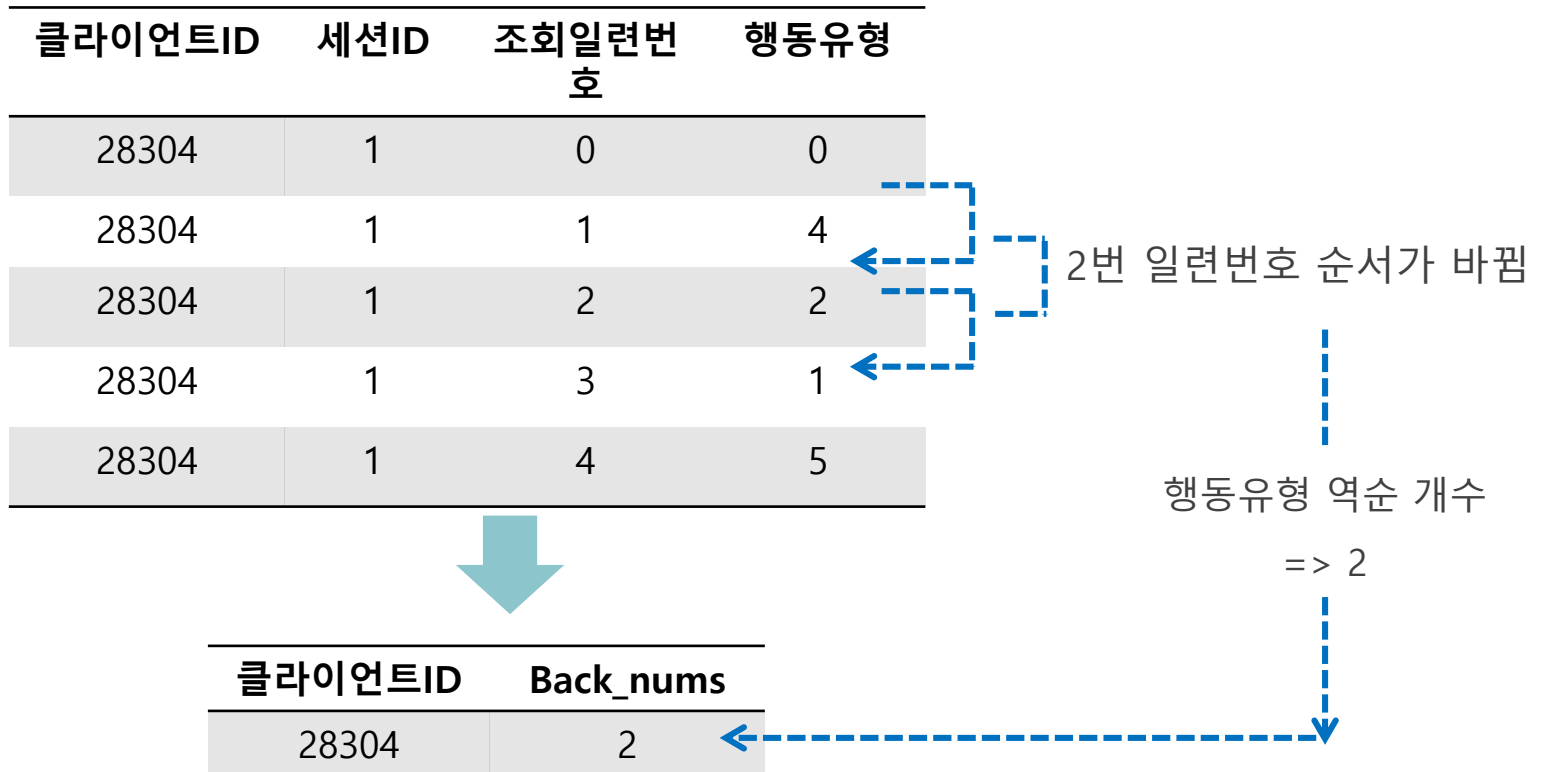


# Clustering

By discretion

## ✓ Number of press back\_button

신중한 고객이라면 행동 유형 0번 부터 6까지 번호가 순차적으로 이어져 한 번에 구매로 이어지기 보단 구매까지 가기 전 여러 가지 행동을 보여 행동유형 번호가 계속해서 바뀔 것이라 판단! 따라서, 6까지는 순차적으로 증가 해야 하는 행동유형 순서가 바뀔 때마다 개수를 count!



# Clustering

By discretion

## ✓ Period that stay in sessions

신중한 고객이라면 행동 유형 6번(구매)가기 전까지 행동유형 0~5번에서 머무르는 시간이 길 것이라 판단!

클라이언트ID	세션 ID	조회일련번호	행동유형	조회경과시간
2	8	1	0	624578
2	8	2	2	791905
2	8	3	6	306944
2	8	4	5	3251794
2	8	5	6	306944
2	10	1	6	3251794

세션ID에서 처음 행동유형 6이 나온  
조회경과시간에서 **최댓값**을  
구매까지\_조회시간으로 만들어준다

클라이언트 ID	구매까지_조회경과시간
28304	3251794

# Clustering

By discretion

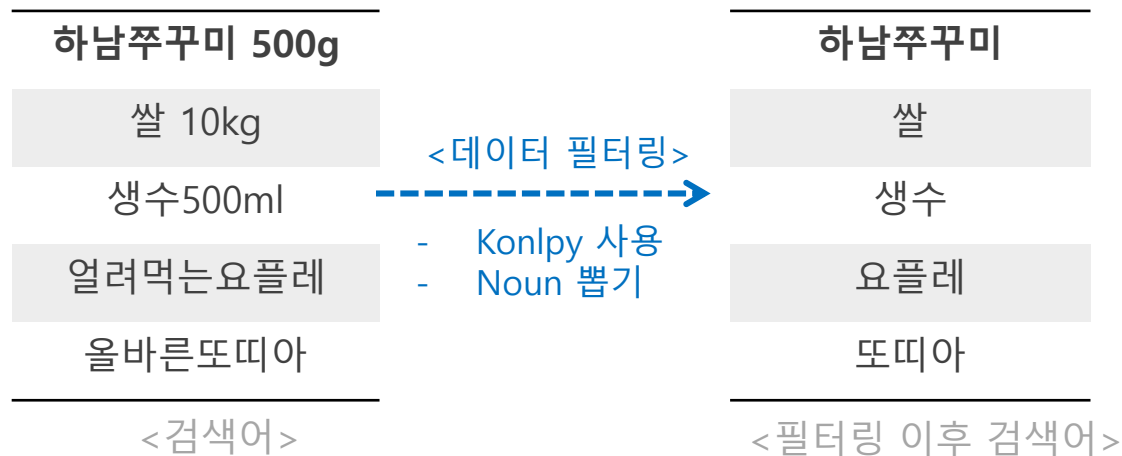
## ✓ Degree of similarity of keywords

신중한 고객이라면 유사한 검색어를 반복적으로 입력하였을 것이라 판단!

따라서, 각 주, 각 월 마다 평균 얼마만큼에 유사한 검색어를 입력하였는지 측정하는 지표를 생성

ex) (패딩, 노스페이스패딩, 남성패딩) => 3번 유사한 검색어를 검색하였다!

<검색어 유사도 생성 과정 - process1>



# Clustering

By discretion

## ✓ Degree of similarity of keywords

<검색어 유사도 생성 과정 – process2>

1. "levenshtein metric distance" 사용하여  
단어들 사이 Lenv\_distance를 구함

클라이언트ID	sech_kwd_1	sech_kwd_2	Lenv_distance	유사 Yes/No	Caculate_sim
5535	두루마리휴지	구강청결제	0	No	
5535	식빵	식빵	1	Yes	29
5535	두부	순두부	0.8	Yes	
58655	돼지고기앞다리	잡채용돼지고기	0.533333	Yes	
58655	비비고만두	만두	0.57142857	Yes	64

2. 단어들 길이마다 다른 threshold를 주어  
유사한지 안한지 결정하도록 함

‘돼지고기앞다리 / 잡채용돼지고기’ vs ‘두부/순두부’  
‘돼지고기앞다리 / 잡채용돼지고기’가 음절 하나 하나 비교하면서  
삭제하고 첨가해야 할 개수가 더 많기 때문에 threshold를 더 작게 잡아준다

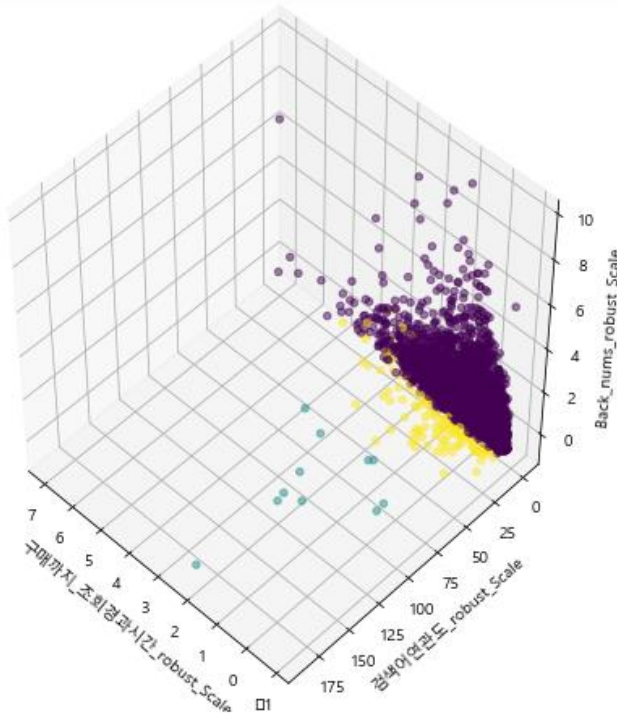
하나라도 단어길이 1~2이면 threshold = 0.6  
하나라도 단어길이 3이상이면 threshold = 0.4

3. 클라이언트ID마다 YES개수를  
count하여 caculate\_sim 완성!

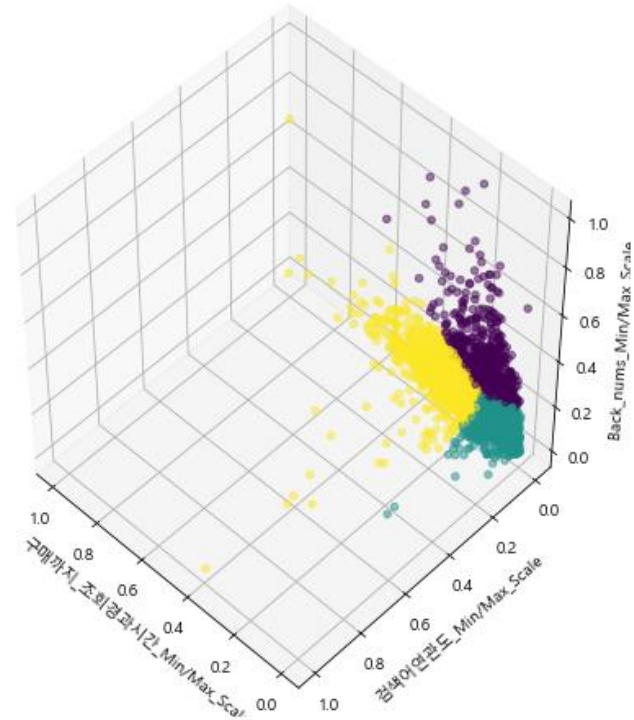
# Clustering

By discretion

## Robust\_Scaling



## Min/Max\_Scaling



한눈에 보아도 Min/Max\_Scaling이 더 고르게 군집화된 것을 볼 수 있다.

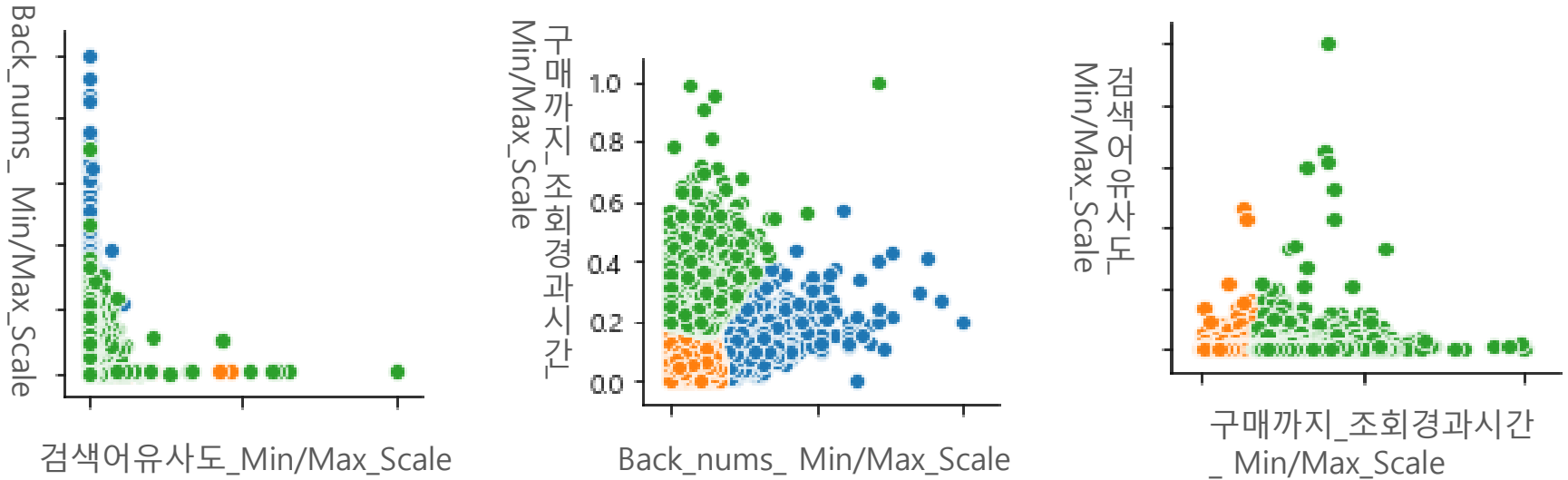
Robust\_Scaling은 아웃라이너의 영향을 최소화한 scaling이기 때문에  
군집을 균일하게 잡지 않는 것을 볼 수 있다.



# Clustering

By discretion

n = 3 Clustering



	Back_nums	검색어유사도	구매까지_조회경과시간	재정의
A	↓	↑	↑	구매에 신중하면서 구매물품이 뚜렷한 소비자
B	↓	—	↓	구매에 신중하지 않은 소비자
C	↑	↓	—	구매물품은 뚜렷하지 않지만 구매까지 신중한 소비자

# Clustering

By loyalty

## ✓ Degree of interest

엘포인트에 충성도가 높은 소비자들은 엘포인트 에 주기적으로 접속할 것이라 판단!

20190701부터 20190930까지 일주일 단위로 나눠 접속빈도를 세어 관심도 변수를 만들어준다

클라이언트 ID	일자1	일자2	...	일자12	일자13
7	1.0	0.0	...	0.0	0.0
9	6.0	0.0	...	1.0	1.0
49	4.0	4.0	...	7.0	4.0
41	6.0	3.0	...	12.0	5.0

클라이언트ID	관심도
7	0.384615
9	2.538462
49	5.384615
41	4.615385

일자1 => 20190701~ 20190708에 접속한 횟수

일자2 => 20190708~ 20190715에 접속한 횟수

⋮

일자13 => 20190923~ 20190930에 접속한 횟수

$$\sum_{i=1}^{13} \text{일자}(i) / 13 = \text{관심도}$$

ex) '관심도 < 1' 이면,  
'평균 일주일 동안 한번도 접속 하지 않았다'라는 의미

# Clustering

By loyalty

## ✓ Rate of using merchandise code to search

검색어 중 상품코드를 검색한것으로 추정되는 검색어가 존재함을 발견하고 상품코드를 검색하고 들어온 사람들은 구매의사가 명확한 소비자라 판단되어 **검색어 중 상품코드의 기여도를 판단**하고자 한다

ex) 8801007202709, wa20r7870gv

상품코드를 검색하고 구매까지 한 사람 수

상품코드를 검색한 사람 수

0.6

VS

상품명을 검색하고 구매까지 한 사람 수

상품명을 검색한 사람 수

0.1

상품코드를 검색한 소비자 / 상품명을 검색한 소비자들의

**구매전환율**에 차이가 있음을 알 수 있다.

따라서, 상품코드가 구매에 중요한 영향을 미쳐 상품코드가 구매에 기여하는 비율을 구한다

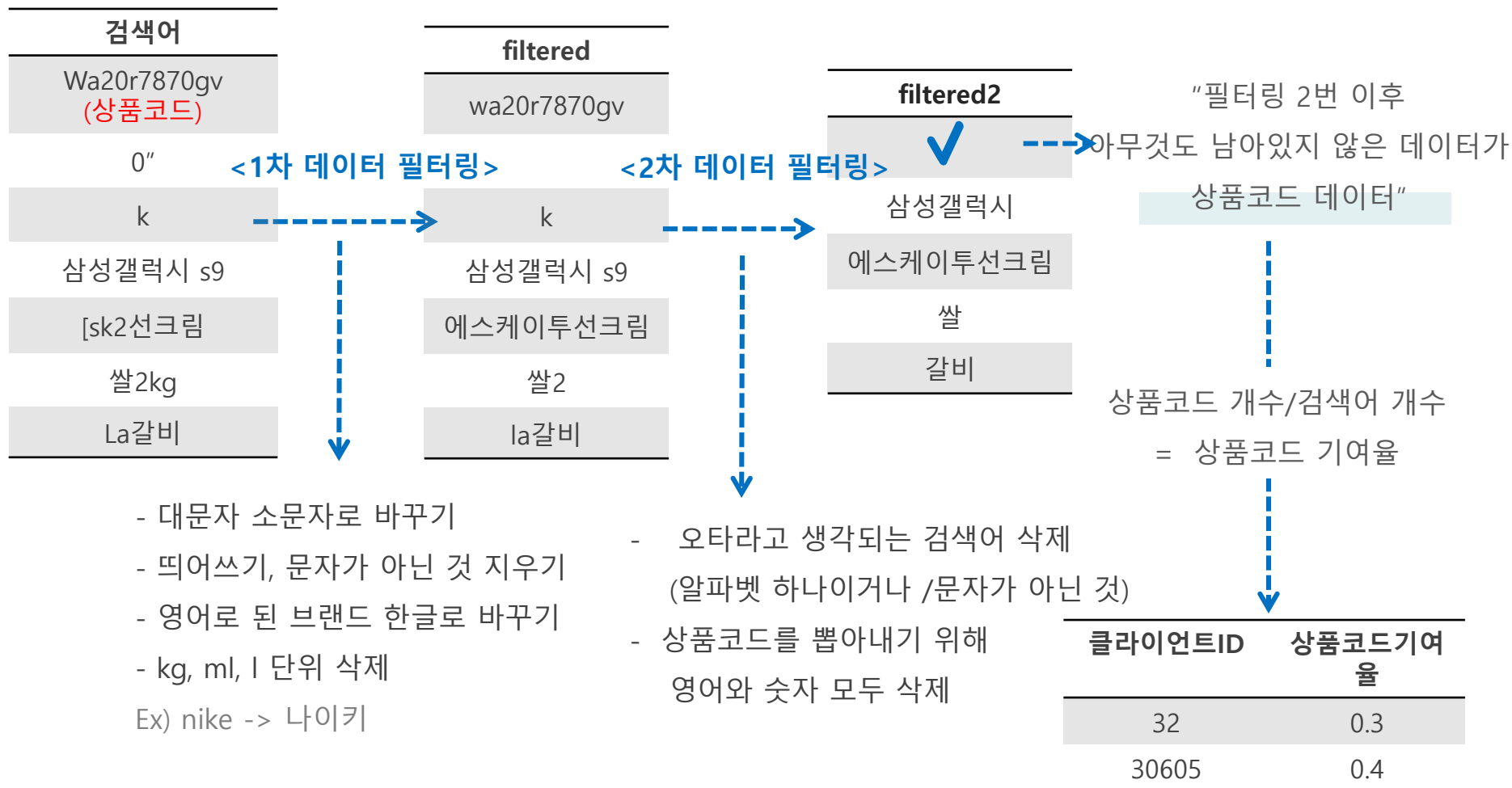
# Clustering

By loyalty

## ✓ Rate of using merchandise code to search

<상품코드 생성 과정>

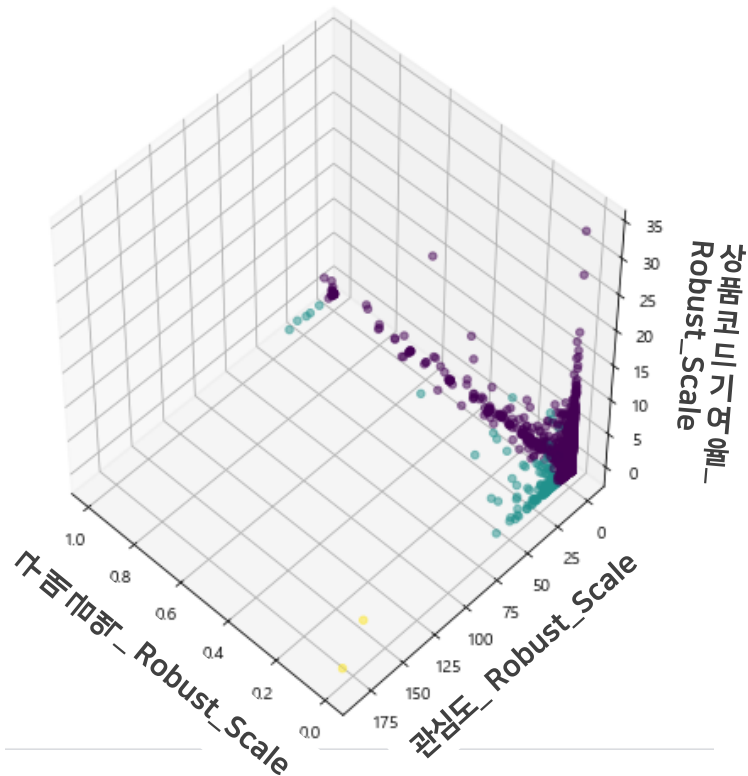
상품코드는 보통 영어+숫자 또는 숫자로 구성. 따라서, 데이터 필터링 후 영어와 숫자를 다 지워 아무것도 남지 않았을 때가 상품 코드 데이터



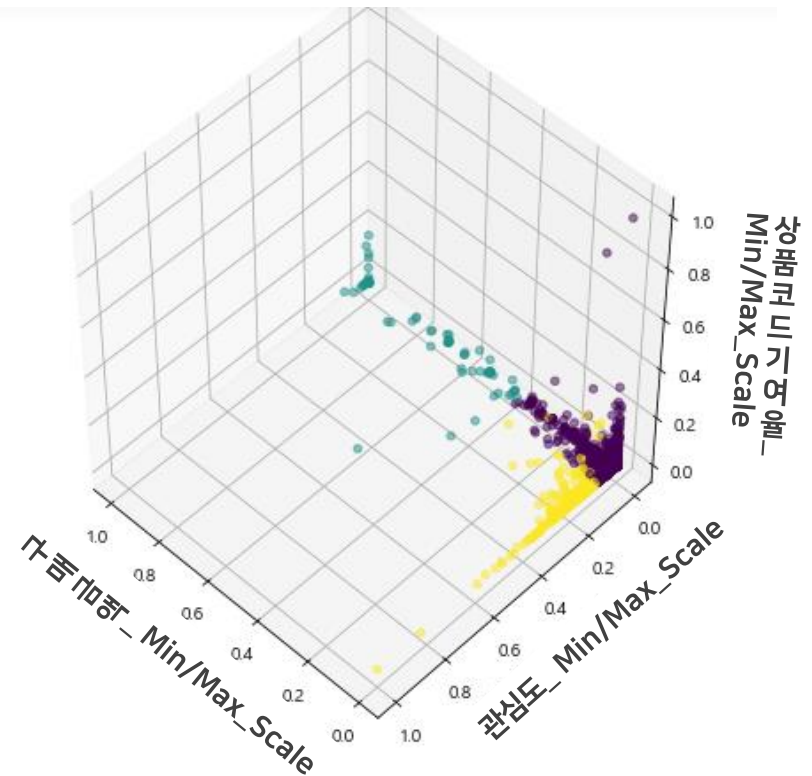
# Clustering

By loyalty

## Robust\_Scaling



## Min/Max\_Scaling



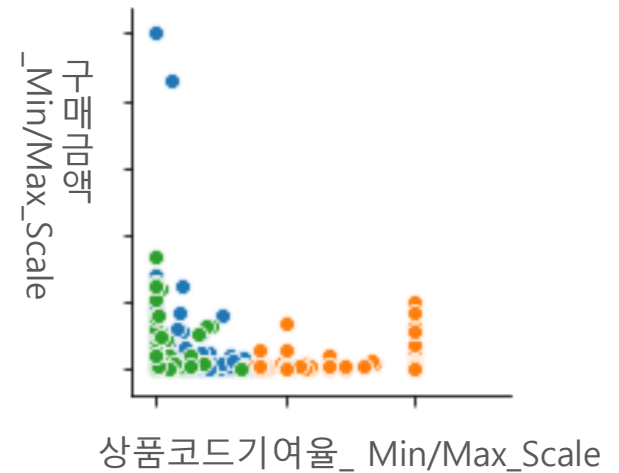
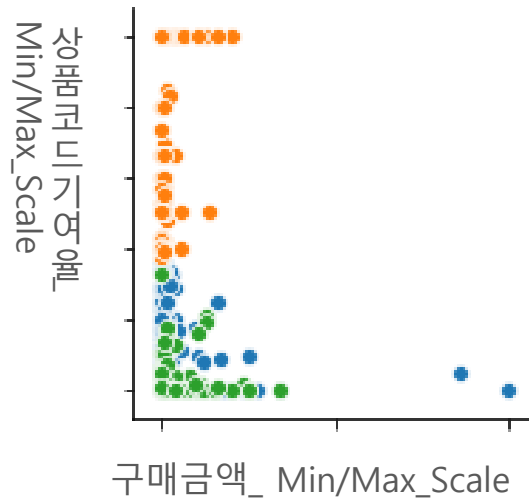
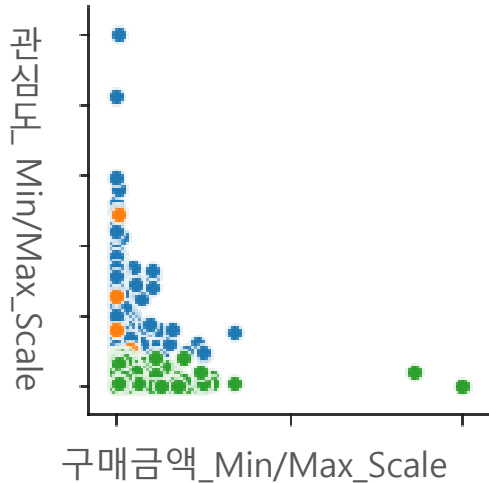
Robust\_Scaling에선 구매금액 기준으로 나뉘지지 않아

한눈에 보아도 Min/Max\_Scaling이 더 고르게 군집화된 것을 볼 수 있다.

# Clustering

By loyalty

n = 3 Clustering



	관심도	상품코드기여율	구매금액
A	↑	↓	↑
B	↓	↑	↓
C	↓	—	↑↑



재정의

특정물건을 사고자 엘포인트에 들어오는 소비자는 아니지만 엘포인트관심과 구매가 많은 소비자

상품코드를 검색할때만(특정 물건을 목표) 엘포인트에 접속하는 소비자

접속 빈도가 낮고  
한번에 비싼 물건을 사는 고객

# Clustering

By tendency

## ✓ Degree of purchase repeatably

습관형 소비를 하는 사람들은 주기적으로 오프포인트에 들어가 **특정 상품을 반복적으로 구매할 것이라 판단!**

클라이언트ID	거래ID	상품소분류명
28304	104999	Instant Rice
28304	104999	Water
28304	104907	Instant Rice
28304	105124	Instant Rice
28304	105124	Car Paint
28304	105124	Water

Instant Rice=> 3번

Water=> 2번

클라이언트ID	반복구매
28304	3

Instant Rice를 3번,  
Water을 2번

구매했으므로 '최댓값'을  
반복구매로 넣어준다

# Clustering

By tendency

## ✓ Degree of access time concentration

EDA를 바탕으로 소비자들마다 들어오는 시간대에 차이가 있음을 발견하고 습관형 소비를 하는 사람들은 엘포인트에 접속하는 **시간에 주기성을 띄고 몰릴 것으로 판단!**

클라이언트ID	6-12	12-18	18-24	24-6
21289	0	238	144	179
48990	4	49	0	0
30547	0	4	20	1

--> '6-12' : 3달동안 6AM-12PM에 접속했던 횟수를 다 더한 값  
'12-18' : 3달동안 12PM-18PM에 접속했던 횟수를 다 더한 값  
'18-24' : 3달동안 18PM-24PM에 접속했던 횟수를 다 더한 값  
'24-6' : 3달동안 24PM-6AM에 접속했던 횟수를 다 더한 값

"시간대에 몰리는 크기와 정도를 수치화하기 위해  
6-12/12-18/18-24/24-6값들의  
표준편차 를 시간집중도로 만들어준다"

클라이언트ID	시간집중도
21289	87.665
48990	420.705
30547	8.074

-->

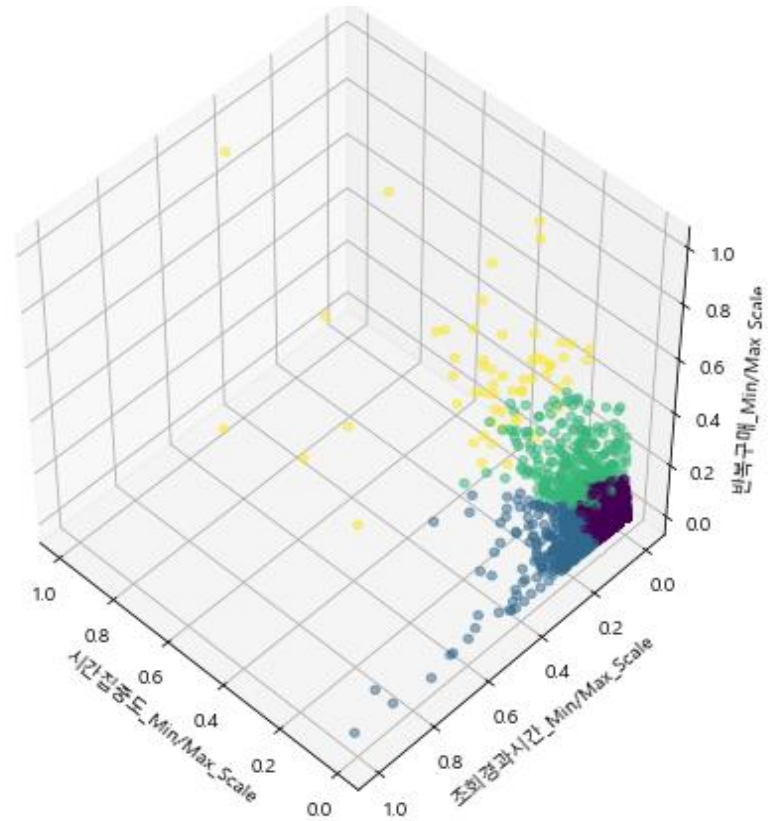
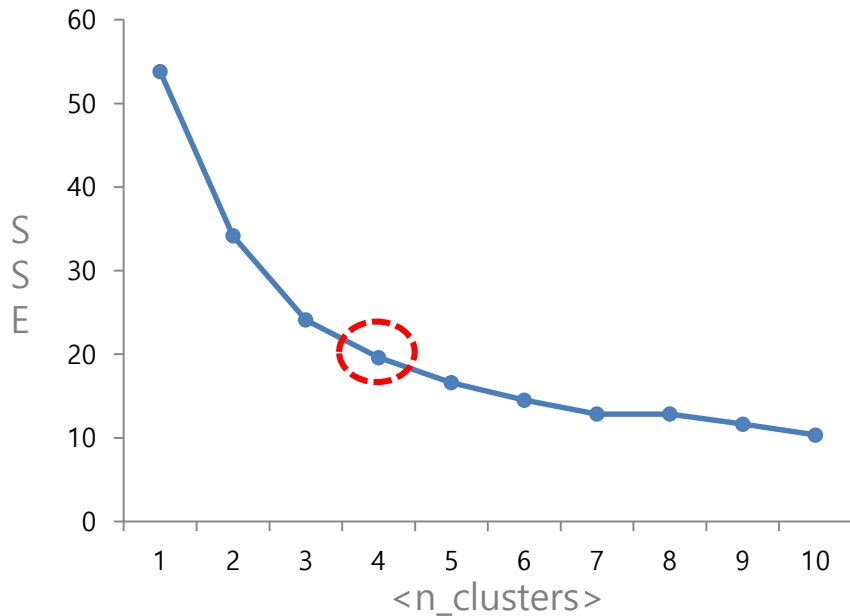
Ex)  
 $x = [100, 60, 80, 50]$   
 $\text{np.std}(x) = 19.2$   
 $x = [0, 0, 20, 0]$   
 $\text{np.std}(x) = 8.66$   
 $x = [100, 0, 20, 0]$   
 $\text{np.std}(x) = 41.23$   
 $x = [0, 0, 200, 0]$   
 $\text{np.std}(x) = 86.6$

표준편차들을 보면 "숫자들이 몰린 정도"와  
"숫자 자체의 크기"도 잘 나타냄을 보여준다



# Clustering

By tendency

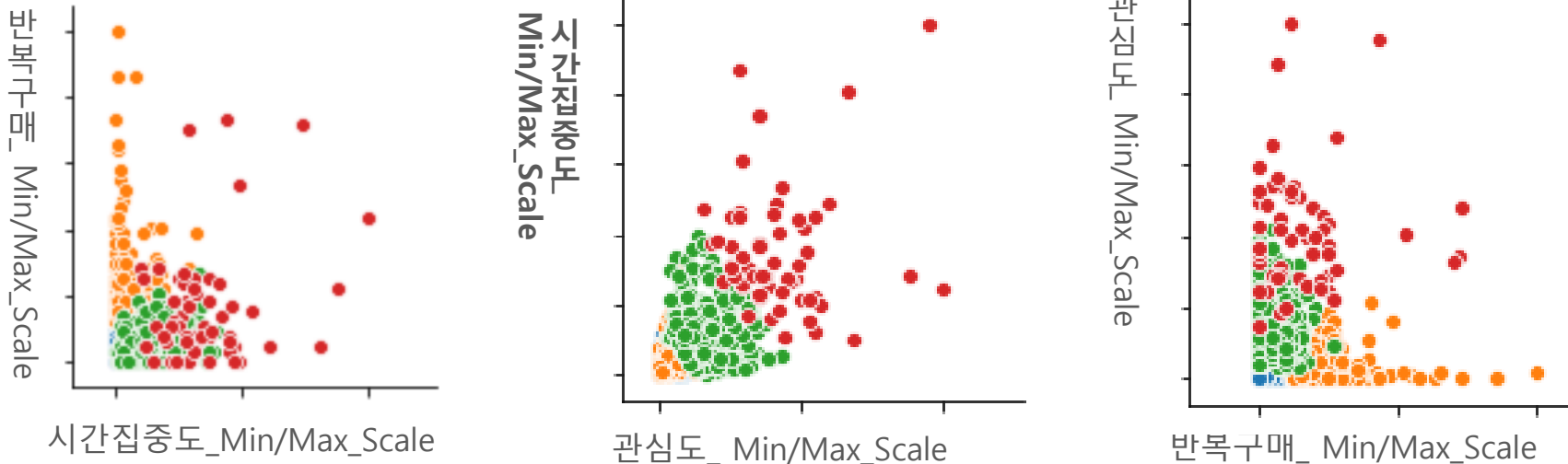


왼쪽 그래프는 클러스터의 개수에 따른 SSE값(응집도)을 나타내고 있다. 클러스터가 3개 또는 4개일 때 SSE값이 확 꺾이는 모습을 보이고 있기 때문에 적합한 클러스터의 개수는 3개 또는 4개라 할 수 있다. 오른쪽 그래프는 클러스터를 4개로 하고 3D로 그린 그래프이다. 그래프에서 볼 수 있듯 데이터들이 균일하게 잘 나뉜 모습을 볼 수 있다.

# Clustering

By tendency

n = 4 Clustering



	반복구매	시간집중도	관심도	재정의
A	↓↓	↓↓	↓↓	엘포인트 관심과 구매가 부족한 소비자
B	↑↑	↓	↓	엘포인트에 자주 들어오지는 않지만 사던 상품만 사는 반복구매자
C	↓	↑	↑	엘포인트에서 아이쇼핑만 하는 구매자
D	↑	↑↑	↑↑	습관적 소비를 하는 소비자

# Category of merchandises recommendation

---

- 01. Based on personal preferences
- 02. Based on demographic trend



# Category recommendation

Based on personal preferences

## "Purpose"

구매이력을 바탕으로 함께 구입한 상품들간의 관계들의 규칙을 찾아(연관 규칙) 상품을 추천해주기 위해서!

ex) 기저기를 사면 맥주를 함께 구매하는 연관규칙을 찾으면 어떤 소비자가 기저기를 구입할 시 맥주를 추천해준다

<연관규칙을 통한 선호 상품군 도출 과정 >

클라이 언트ID	Baby	Beauty	Living, Health	...	Food	Interior
7	1.0	0.0	0	...	0.0	0.0
9	6.0	0.0	0	...	1.0	1.0
49	4.0	4.0	2	...	7.0	4.0
41	6.0	3.0	9	...	12.0	5.0

< 클라이언트ID별로 구입했던 상품군의 개수 >

"Apriori 기법을 사용한 연관규칙 탐사"

base	=>	add	support	confidence	lift
Baby	=>	Foods	0.015608	0.143260	1.133969
Foods	=>	Baby	0.015608	0.123543	1.133969
Baby	=>	Interior	0.033038	0.303248	1.437726
Interir	=>	Baby	0.033038	0.156636	1.437726

Min\_support = 0.01    Min\_confidence = 0.1

# Category recommendation

Based on personal preferences

## "Interpretation of results"

base	=>	add	support	confidence	lift
Baby	==>	Foods	0.015608	0.143260	1.133969
Foods	==>	Baby	0.015608	0.123543	1.133969
Baby	==>	Interior	0.033038	0.303248	1.437726
Interior	==>	Baby	0.033038	0.156636	1.437726

지지도와 신뢰도는 내가 정한 최소 임계치(min\_support = 0.01, min\_confidence = 0.1)을 넘었기 때문에 **"향상도"** 위주로 살펴본다  
1보다 크므로 Baby만 샀을 때 보다, Interior과 같이 살 확률이 크다 => 추천 상품

만약 X가 발생하였을 때, Y도 같이 발생한다는 연관성을  $X \rightarrow Y$ 로 표기하고 이 때, X를 선행절 / Y를 후행절이라 한다.

**지지도(support) => 전체 거래에서 항목이 포함된 확률**

ex) 전체 거래에서 Foods는 100개의 상품 중 1개의 빈도를 보인다

**신뢰도(confidence) => 선행절이 포함된 거래 중 후행절의 항목들이 포함될 확률**

ex) Foods를 산 거래 중 Baby와 같이 살 확률이 0.12이다

**향상도 (lift) => 선행절과 후행절이 독립적인 경우 대비, 선행절과 후행절이 함께 발생할 확률**

향상도가 1이면 독립적인 관계이며, 1보다 크면 (+)양의 상관관계 1보다 작으면 (-)음의 상관관계이다.

ex) 1보다 크다는 것은 Foods만 단독으로 샀을 때 보다 Foods와 Baby를 같이 살 확률이 크다는 것이다

# Category recommendation

Based on demographic trend

## " Purpose "

연령별 쇼핑 트렌드에 따라 최근 구매가 증가하는 상품군을 추천하고자 한다!



Problem

L.Point 데이터에서 연령/성별에 따라 나누고자 하였으나

성별까지 나눌 경우 10대, 20대와 60대에서 구매인원이 너무 작아지기 때문에

Trend를 구하기 적절하지 않음

<10대 상품군별 구매인원>

대분류명	월	클라이언트 명수
Beauty	8	6
	9	5
Digiter, Appliances	7	2

<60대 상품군별 구매인원>

대분류명	월	클라이언트 명수
Baby	7	10
	8	2
	9	9

다른연령대는 평균 400

Solution



외부 데이터 사용  통계청

통계청 - 온라인 쇼핑 동향 자료를 활용하여  
구매인원이 부족한 부분을 대체하여 활용

2) 온·오프라인병행몰 거래액

구 분	합계	가 전			도 서		
		컴퓨터 및 주변기기	가전·전자 ·통신기기		서적	사무· 문구	
7	3,616,940	380,139	90,406	289,733	121,512	105,612	15,900
8	3,595,575	341,060	98,681	242,379	117,645	101,617	16,028
9	3,659,008	300,333	90,232	210,101	113,525	99,233	14,292
10p	3,899,302	364,693	97,443	267,250	106,211	89,781	16,430
11p	4,011,875	369,045	96,525	272,520	112,544	92,043	20,501

# Category recommendation

Based on demographic trend

<연령대별 trend 계산 과정-process1>

엘포인트에서 제공받은 데이터(고객 Demographic 정보, 상품 분류 정보)들로 구한 연령별 / 상품군별 / 월별 구매금액 합으로 지난달 대비 구매금액 증감율을 구해 엘포인트 안에서 나타나는 trend를 수치화한 **"trend lpoint"**를 만든다. 또한, 통계청의 온라인 동향 쇼핑 자료로 상품군별 지날달 대비 구매금액 증감율을 구해 사회 전반의 trend를 수치화한 **"trend Korea"**를 만든다.

단위 : 원			
나이	상품군	월	구매금액
10	Beauty	8	71,680
		9	63,690
	Digital, Appliances	7	68,530
		8	391,410
	Fashion Accessories	8	304,710
		9	226,830

<L.Point 연령별/상품군별/월별 구매금액 합>

지난달 대비 구매금액 증감율

**"trend Lpoint"**

단위 : 백만원		
상품군	월	구매금액
Beauty	7	696,542
	8	<b>740,689</b>
	9	<b>803,690</b>
	7	<b>380,139</b>
Digital, Appliances	8	<b>341,060</b>
	9	<b>300,333</b>

<통계청 상품군별/월별 구매금액 합>

지난 달 대비 구매금액 증감율

**"trend Korea"**

# Category recommendation

Based on demographic trend

<연령대별 trend 계산 과정-process2>

엘포인트 안의 트렌드를 나타내는 trend Lpoint와 사회 전반의 트렌드를 나타내는 trend Korea의 평균을 낸 "trend mean" 을 구해 월별 트렌드 상품군을 추천할 수 있도록 한다

6월 구매자료가 없기 때문에 trend\_korea로 대체

ex) 20대 트렌드

상품군	월	구매금액 _Lpoint	Trend_ Lpoint	구매금액 _Korea	Trend_ Korea
Beauty	7	4889170	0.050069	696,542	0.050069
	8	4757470	-0.020000	740,689	0.063380
	8	4117675	-0.130000	803,690	0.085057
Digital, Appliances	7	9044890	0.057181	380,139	0.057181
	8	4119730	-0.540000	341,060	-0.102802
	9	2404690	-0.410000	300,333	-0.119413

월	Trend_ mean	추천 상품 군
7	0.067508	Living, Health
8	0.285858	Others
9	-0.085348	Fashion Clothes

Trend\_mean 을 기준으로 순위를 매겨  
가장 높은 점수를 기록한 상품군을  
추천 상품군으로 선정

Lpoint와 Korea의 증감을 평균 내어 최종 상품군별 증감도를 구한다



# Business model

---

01. Scenario of recommendation algorithm
02. Investigation process
03. Super personalized benefit-recommendation



# Business model

## Scenario of recommendation algorithm

### SSTP 를 활용한 구체적 추천 적용 시나리오

#### Super Targeting



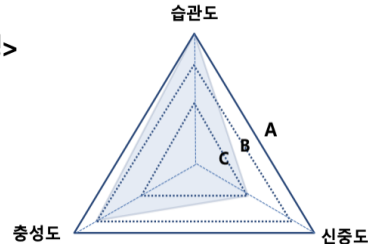
##### < 로그데이터를 이용한 A 의 특성 >

- 20대 남자
- 온라인 업종을 선호
- 주로 Direct 방식으로 들어오며 mobile\_web 이용

##### < SSTP 을 사용한 A 의 특성>



- 습관도 높음
- 충성도 보통
- 습관도 낮다



##### < SSTP 을 사용한 A 의 상품추천 >



- 선호 상품군: Food
- 선호 기준 연관 추천 상품군: Baby
- Trend 기반 추천 상품군: Fashion\_clothes

혜택 추천  
알고리즘

#### A 맞춤 최종 혜택

마일리지  
추가적립

맞춤 혜택1  
해당상품군 재 결제시

50% 더 적립!

중복가능  
할인쿠폰

맞춤 혜택2  
최저가에서 한번 더 할인

중복할인 가능!

10 %  
할인쿠폰

맞춤 혜택3  
수량 100 개

마음껏 할인!

1. Food 상품군 구매시 마일리지 2배 적립
2. Baby 상품군 구매시 5% 할인 쿠폰 5장
3. Fashion\_Clothes 상품군 구매시 사은품 제공

# Business model

## Scenario of recommendation algorithm

“Personalized recommendation service based on one’s demographic and shopping information”



### Problem

고객의 구매특성별로 클러스터링은 가능하지만 혜택에 대한 데이터가 없기 때문에 고객의 구매특성이 혜택 선호도를 묶어줄 수가 없음

### Solution



고객의 구매 특성과 혜택 선호도에 대한 설문조사를 시행하여 구매특성이 혜택 선호에 얼마나 영향을 미치는지를 구하여 구매특성과 혜택 선호도를 연결한다

# Business model

## Scenario of recommendation algorithm

### Ex) 고객 정보

: 20대 여성, 온라인 쇼핑을 선호하고 주로 Direct 방식으로 들어오며 주로 mobile\_web을 이용한다. 소비특성은 신중도는 A집단, 습관도는 B집단, 충성도는 C집단이다.

“고객 개별  
추천 상품군”



“구매특성별  
소속 집단 기반  
혜택 순위”

선호 상품군		연관 상품군		Trend 상품군				
Baby		Interial		Fashion_Clothes				
클라이언트 ID	신중도	습관도	충성도	혜택별 점수 합				
				혜택1	혜택2	혜택3	혜택4	혜택5
24	신중도_A	습관도_B	충성도_C	0.38	0.75	0.63	0.43	0.79



“혜택 추천”

1. Baby 상품군 구매 시 n% 할인 받을 수 있는 쿠폰 제공
2. Interial 상품군 구매 시 중복 할인이 가능한 쿠폰 제공
3. Fashion\_Clothes 상품군 구매 시 사용할 수 있는 배송비 쿠폰 제공

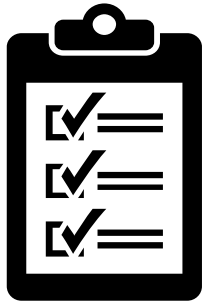
# Business model

## Questionnaire

설문 조사

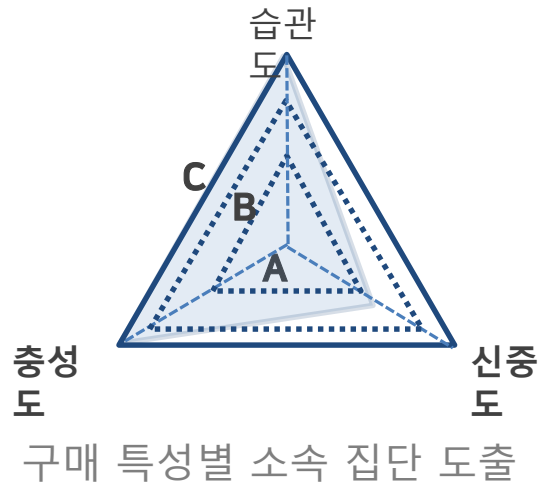
소속 집단 도출

혜택 선호 조사



- ✓ 신중도
- ✓ 습관도
- ✓ 충성도

클러스터링 기준들을  
활용하여 설문 실시



ex ) 클라이언트 ID =1  
신중도\_A, 습관도\_B, 충성도\_C 소속

혜택 1 혜택 2 혜택 3 혜택 4 혜택 5  
순위    **2**    **4**    **3**    **1**    **5**

소속 집단 별로 혜택 선호 조사

# Business model

## Kinds of benefits

N %  
할인쿠폰

맞춤 혜택 1

수량 100 개

마음껏 할인!

중복가능  
할인쿠폰

맞춤 혜택 4

최저가에서 한번 더 할인

중복할인 가능!

마일리지  
추가적립

맞춤 혜택 2

해당상품군 재 결제시

50% 더 적립!

재구매  
할인쿠폰

맞춤 혜택5

1주일 이내 재 결제시

10%더 할인!

배송비  
할인쿠폰

맞춤 혜택 3

식품 배송 시

빠른 총알배송!

### 혜택 설명

- 혜택 1: % 할인 쿠폰
- 혜택 2: 마일리지 추가 적립
- 혜택 3: 배송비 할인 쿠폰
- 혜택 4: 타 쿠폰과 중복해서 사용할 수 있는 할인 쿠폰
- 혜택 5: 구매 이후 해당 상품 재 구매 시 사용할 수 있는 쿠폰을 추가로 제공

# Business model

## Investigation process

	신중도 A	신중도 B
클라이언트 ID	2 24 35 ...	5 12 20 ...



ID	혜택 선호도				
	혜택1	혜택2	혜택3	혜택4	혜택5
2	5	2	1	3	4
24	3	1	4	5	2
...	...	...	...	...	...
총합	134	245	78	210	129

1. 각 집단에 속하는 사람들을 선별

2. 혜택 순위를 기준으로  
점수를 매긴 후 점수를 모두 더한다



신중도	혜택1	혜택2	혜택3	혜택4	혜택5
신중도_A	0.13	0.25	0.21	0.14	0.27
신중도_B	0.11	0.21	0.23	0.17	0.29
신중도_C	0.10	0.24	0.19	0.16	0.3



신중도	혜택1	혜택2	혜택3	혜택4	혜택5
신중도_A	78	151	126	82	162
신중도_B	32	59	65	48	81
신중도_C	45	120	67	78	92

3. 혜택별 점수를 집단별 점수 총합으로 나누어  
집단에 따른 혜택별 선호도를 계산

# Business model

## Investigation process

신중도	혜택1	혜택2	혜택3	혜택4	혜택5
신중도_A	0.13	0.25	0.21	0.14	0.27
신중도_B	0.11	0.21	0.23	0.17	0.29
신중도_C	0.10	0.24	0.19	0.16	0.3

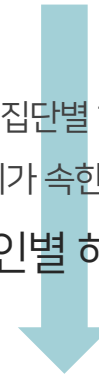
습관도	혜택1	혜택2	혜택3	혜택4	혜택5
습관도_A	0.13	0.25	0.21	0.14	0.27
습관도_B	0.11	0.23	0.22	0.14	0.29
습관도_C	0.10	0.24	0.19	0.16	0.3

충성도	혜택1	혜택2	혜택3	혜택4	혜택5
충성도_A	0.13	0.25	0.21	0.14	0.27
충성도_B	0.11	0.21	0.23	0.17	0.29
충성도_C	0.13	0.27	0.20	0.16	0.24

<클라이언트 ID = 24가 소속된 집단>

	혜택1	혜택2	혜택3	혜택4	혜택5
신중도_A	0.13	0.25	0.21	0.14	0.27
습관도_B	0.11	0.23	0.22	0.14	0.29
충성도_C	0.13	0.27	0.20	0.16	0.24

앞서 구한 집단별 혜택 선호 점수를  
클라이언트 아이디가 속한 집단의 값으로 가져와서  
이를 모두 합산하여 개인별 혜택 선호 점수를 계산한다.



클라이언트 ID	신중도	습관도	충성도	혜택별 점수 합				
				혜택1	혜택2	혜택3	혜택4	혜택5
24	신중도_A	습관도_B	충성도_C	0.38	0.75	0.63	0.43	0.79

<클러스터링이 완료된 L.Point 고객 정보>



# Appendix

## 쇼핑 성향에 따른 혜택 선호도 조사

안녕하세요. 설문에 참여해주셔서 감사합니다!

본 설문은 쇼핑 성향이 선호하는 혜택에 미치는 영향을 조사하기 위해 마련된 설문이며, 설문 내용은 연구 목적 이외에는 사용되지 않습니다. 감사합니다. 추천을 통해 스타벅스 기프티콘을 드릴 예정이니 많은 참여 부탁드립니다! 감사합니다!

\* 필수항목

홈페이지에 접속 후 물건을 구매하기까지 평균적으로 몇 분 정도가 소요되시나요? \*

- ☐ 10분 이내
- ☐ 10분 ~ 1시간
- ☐ 1시간 이상

구매 결정을 하기 전까지 평균적으로 몇 번의 뒤로가기를 누르시나요? (ex, 타 상품을 비교하기 위해서, 구매가 망설여져서) \*

- ☐ 거의 누르지 않는다
- ☐ 평균 1 ~ 2회
- ☐ 평균 10회 이상

물건을 구매할 때의 검색 스타일은 어떠한 편인가요? \*

- ☐ 검색어를 다양하게 바꿔가면서 검색해보는 스타일 ex) 패딩, 남자패딩, 노스페이스 패딩
- ☐ 하나의 검색어만 주로 사용하는 스타일 ex) 패딩, 우유, 옥수수
- ☐ 검색어를 사용하지 않은 스타일

한 쇼핑물에서 주기적으로 구입하는 상품이 있나요? (ex, 매주 이마트에서 물과 햇반을 주문) \*

- ☐ 예
- ☐ 아니오

주로 쇼핑을 하는 시간대가 언제인가요? \*

- ☐ 새벽 (12:00 ~ 6:00)
- ☐ 오전 (6:00 ~ 12:00)
- ☐ 오후 (12:00 ~ 6:00)
- ☐ 저녁 (6:00 ~ 12:00)
- ☐ 주로 쇼핑하는 시간대가 정해져있지 않음

가장 자주 방문하는 온/오프라인 쇼핑물을 일주일에 평균 몇 회 정도 방문하시나요? \*

- ☐ 자주 방문하는 쇼핑물이 없음
- ☐ 평균 1회 ~ 2회
- ☐ 평균 3 ~ 9회
- ☐ 평균 10회 이상

상품을 검색할 때 상품코드로 검색해본 적 있으신가요? \*

- ☐ 없다
- ☐ 종종 있다
- ☐ 자주 있다

가장 자주 방문하는 온/오프라인 쇼핑물에서 월 평균 얼마 정도의 쇼핑을 하시나요? \*

- ☐ 0 ~ 20만원
- ☐ 20만원 ~ 50만원
- ☐ 50만원 이상

# Appendix

## 쇼핑 성향에 따른 혜택 선호도 조사

\* 필수항목

### 혜택 선호도 조사

각각의 혜택의 선호에 맞게 순위를 매겨주시기 바랍니다. (1 ~ 5순위까지 존재합니다.) \*

	1순위	2순위	3순위	4순위	5순위
% 할인 쿠폰	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
마일리지 추가 적립	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
배송비 할인 쿠폰	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
타 쿠폰과 중복해서 사용할 수 있는 할인 쿠폰	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
구매 이후 해당 상품 재구매시 사용할 수 있는 쿠폰을 추가로 제공	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

뒤로

다음

Google 설문지를 통해 비밀번호를 제출하지 마세요.

## 쇼핑 성향에 따른 혜택 선호도 조사

응답이 기록되었습니다.

[다른 응답 제출](#)