

project

Data visualization

Team scalar





L.pay | L.POINT

롯데의 고객은 누구일까?

페르소나 찾기

L.POINT Big Data Competition “Be the L.BA” 데이터 분석

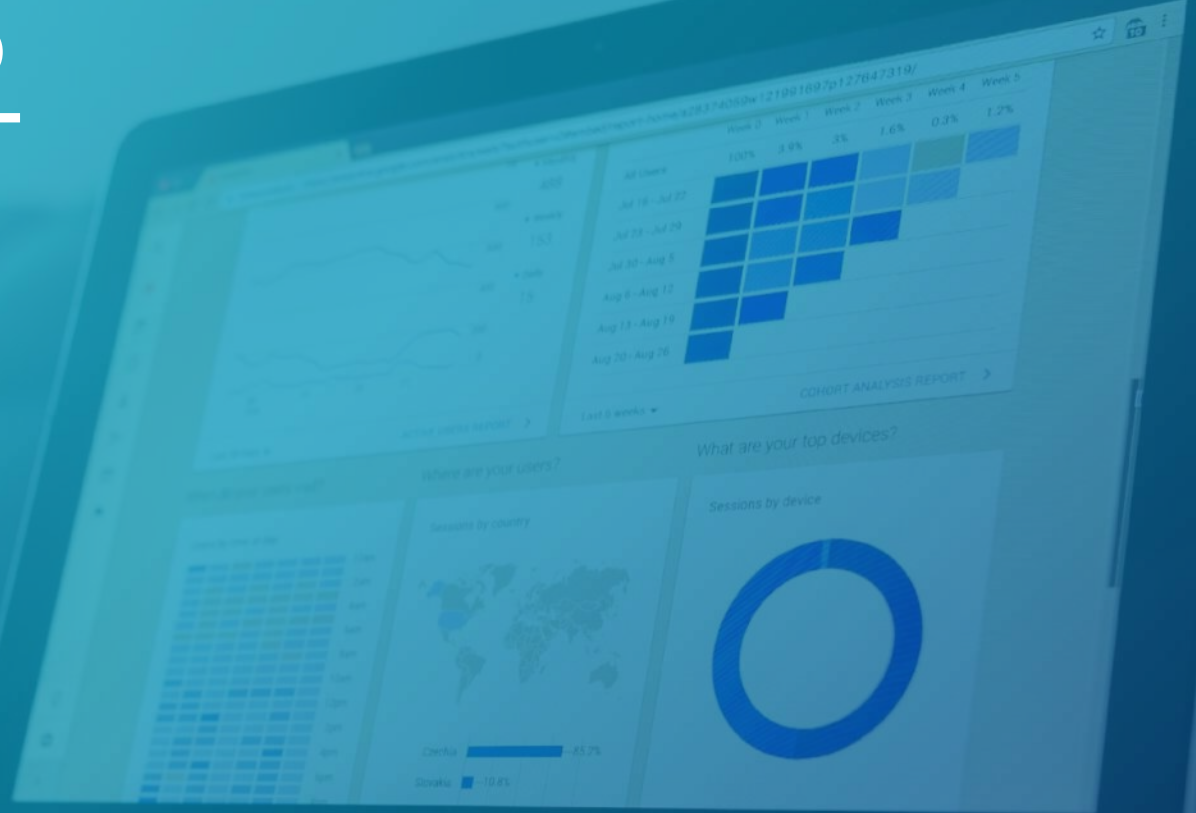
Contents



project

1. 프로젝트 개요

분석 배경
주제 / 목적 선정
진행 프로세스



2018년부터 롯데는, 계열사(롯데마트, 롯데홈쇼핑 등)를 하나의 플랫폼으로 통합 시도!

롯데百貨의 실험 "모든 쇼핑을 모바일로"

"앞으로 모바일로 할 수 없는 일은 하지 않을 겁니다."

강 대표는 이날 롯데백화점 혁신을 키워드로 '디지털 트랜스포메이션(DT)'을 꼽았다. 롯데백화점은 DT를 통해 궁극적으로 고객의 쇼핑 여정을 100% 디지털화하는 것을 목표로 하고 있다. 그는 "40년 역사를 가진 오프라인 유통업체다 보니 고질적으로 시대정신에 뒤떨어지는 부분이 있었다"면서도 "앞으로는 처음부터 끝까지 모든 것을 DT화하고 모바일로 할 수 없는 것은 하지 않을 것"이라고 말했다. 그는 이어 "현재 전체 업무 중 80%를 모두 리빌딩하고 있다"고 덧붙였다.

롯데 e커머스 출범... 대표에 '온라인통' 김
경호 롯데닷컴 대표  본문듣기 · 설정

과연, 롯데는
온라인 쇼핑 계열사*를
통합하면 시너지 효과가 날까?

* 롯데쇼핑 7개사:
백화점·마트·홈쇼핑·하이마트·롭스·닷컴

프로젝트 주제 / 목적 롯데 온라인 고객 분석

우리의 고객은 어떤 사람일까?

온라인 행동 데이터 기반 구매자 특성 분류

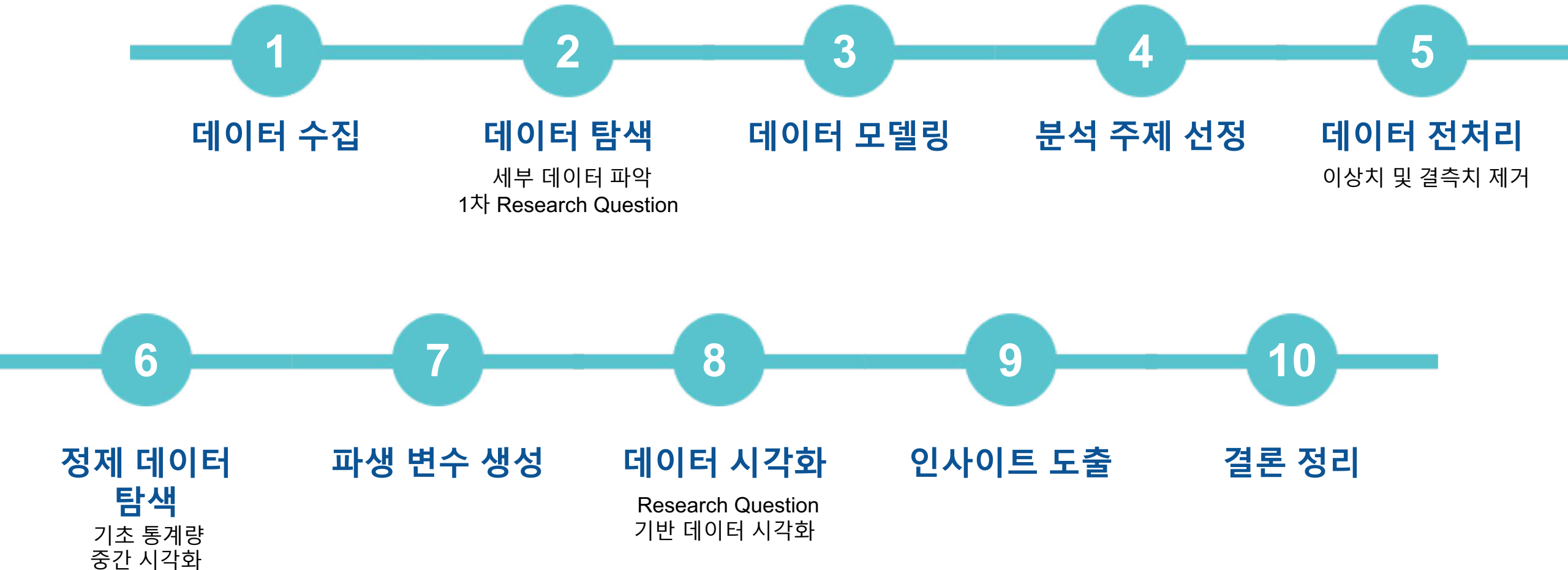
롯데 그룹 온라인 계열사의 페르소나 정의

왜?

롯데의 핵심 타겟층의 니즈와 특성을 파악해,

향후 마케팅 전략에 반영하려 함.

진행 프로세스 데이터 분석 및 시각화



2. 데이터 탐색

데이터 소개

데이터 모델링

데이터 프레임별 기초 탐색 및 통계량



데이터 소개 롯데 온라인 행동 데이터 및 상품 분류 데이터

데이터 수집 시기 2018.4 ~ 2018.9

데이터 범위:

임의 선정된 롯데 그룹 온라인
계열사의 구매 및 이용 이력

분석 테이블

고객 정보

고객의 인구통계학적 특징

세션

고객 온라인 행동

주문

상품 특징별 주문 금액, 횟수

상품 정보

상품 분류별 고객 니즈

검색어

시기별 검색량
검색 & 구매 상관관계
검색 키워드

Customer	Session	Order	Master	Search1	Search2
클라이언트 ID	클라이언트 ID	클라이언트 ID	상품코드	클라이언트 ID	검색일자
고객성별	세션 ID	세션 ID	상품명	세션 ID	키워드
고객나이	방문횟수	HIT 횟수	대분류	키워드	검색횟수
	방문날짜	상품코드	중분류	검색횟수	
	조회페이지수	상품추가정보	소분류		
	총세션시간	상품 브랜드			
	디바이스	개당 가격			
	지역대분류	구매수량			
	지역중분류				

데이터 구조 사용자 시나리오

사용자 시나리오 관점에서의 데이터 구조

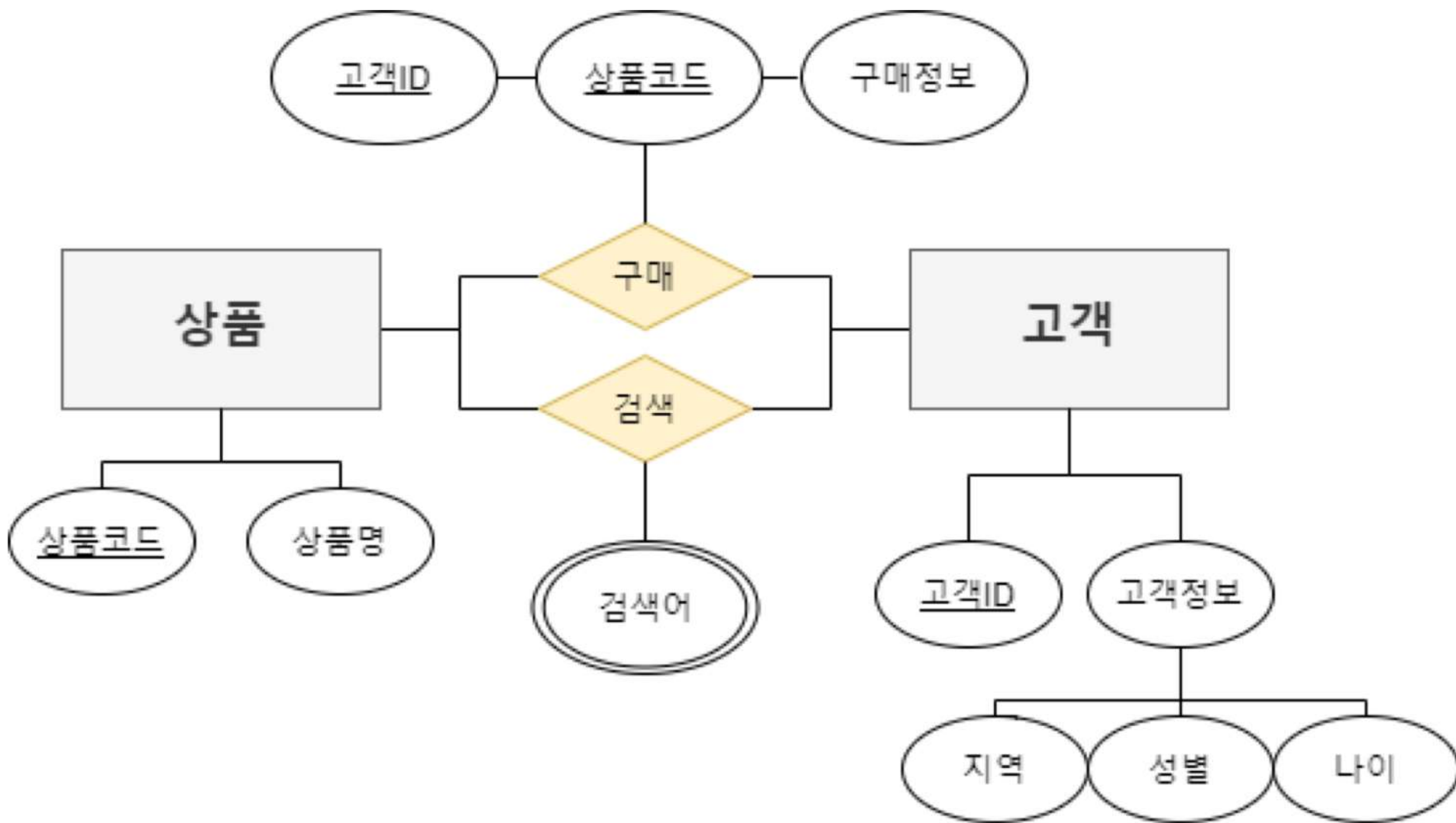


search2

session
search1

order
customer
master

데이터 모델링 개념적 데이터 모델링(피터첸 표기법)



데이터 모델링 개념적 데이터 모델링(IE 표기법)



데이터 모델링 논리적 모델링(데이터 스키마)

[개체]	롯데 계열사	Master (<u>상품코드</u> , 상품명, 상품 대분류명, 상품 중분류명, 상품 소분류명)
	고객	Customer (<u>클라이언트ID</u> , 성별, 연령대)
		Session (<u>클라이언트ID</u> , <u>세션ID</u> , 세션일련번호, 세션일자, 총페이지조회건수, 총세션시간값,기기유형, 지역대분류, 지역중분류)
[관계]	주문	Product (<u>클라이언트ID</u> , <u>세션ID</u> , <u>히트일련번호</u> , <u>상품코드</u> , <u>상품추가정보</u> , 상품브랜드, 단일상품금액, 구매건수, 총금액)
	검색	Search1 (<u>클라이언트ID</u> , <u>세션ID</u> , <u>검색키워드명</u> , 검색건수)
		Search2 (<u>세션일자</u> , <u>검색키워드명</u> , 검색건수)

데이터 기초 탐색 및 통계량 데이터 프레임별

Custom 구매고객 성별, 연령 데이터

```
df_cus['CLNT_AGE'].value_counts()
```

```
30    313600
40    216217
20     95297
50     36624
60     5256
10     3807
70      846
80       32
Name: CLNT_AGE, dtype: int64
```

```
df_cus['CLNT_GENDER'].value_counts()
```

```
F    570616
M    101063
Name: CLNT_GENDER, dtype: int64
```

```
df_ages = df_cus.groupby(['CLNT_AGE', 'CLNT_GENDER'])['CLNT_GENDER'].count()
df_ages
```

```
CLNT_AGE  CLNT_GENDER
10         F           3007
          M            800
20         F          82269
          M         13028
30         F         271712
          M         41888
40         F         180398
          M         35819
50         F         28881
          M          7743
60         F         3784
          M         1472
70         F          542
          M          304
80         F           23
          M            9
Name: CLNT_GENDER, dtype: int64
```

기초 탐색 및 통계 결과

1. 각 데이터의 client id 개수가 다르다
2. 이상치와 결측치가 있다

3. 데이터 탐색

고객 세그멘테이션



고객 세그멘테이션 롯데 고객 그룹 분류

“

롯데의 핵심 고객, 즉 구매력이 높은 고객은 누구일까?
기준은 어떻게 잡아야 할까?

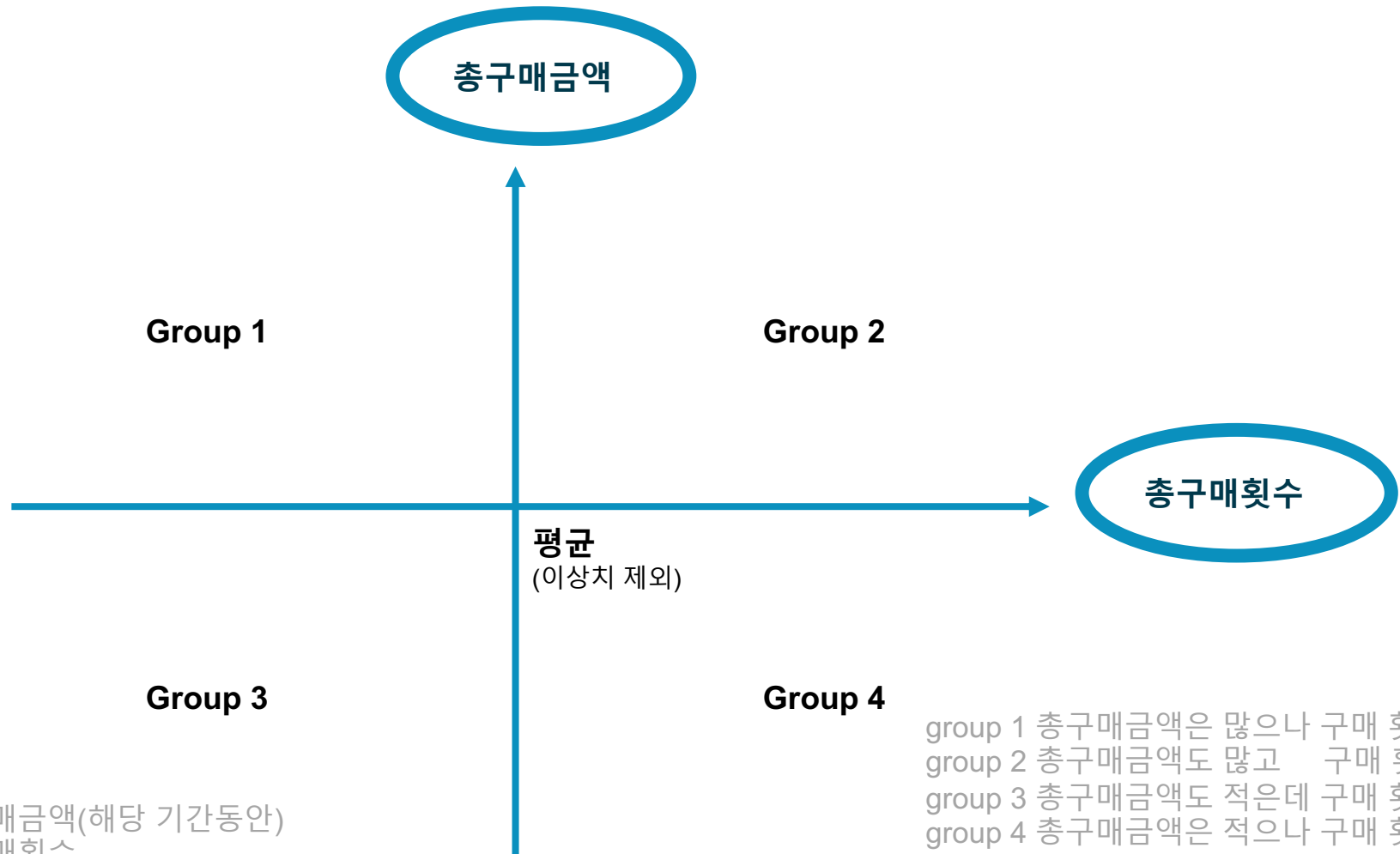
”

구매력이 높은 충성고객 정의

총구매금액과 구매횟수 기준으로 4개의 그룹으로 분류

고객 세그멘테이션 롯데 고객 그룹 분류

롯데의 온라인 고객 세그멘테이션



총구매금액 : 클라이언트ID 당 총 구매금액(해당 기간동안)
총구매횟수 : 클라이언트ID 당 총 구매횟수

group 1 총구매금액은 많으나 구매 횟수가 적은 그룹
group 2 총구매금액도 많고 구매 횟수도 많은 그룹
group 3 총구매금액도 적는데 구매 횟수도 적은 그룹
group 4 총구매금액은 적으나 구매 횟수가 많은 그룹

그룹별 기초 통계 분석

고객 세그멘테이션 데이터

segs 구매고객 성별, 연령, 총구매금액, 총구매횟수 데이터

가장 적은 금액을 소비한 고객은 누구일까?

```
df_segs[df_segs['PD_BUY_TOT'] == df_segs['PD_BUY_TOT'].min()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
416307	5234821	20	2	3
624782	479115	20	2	3
728198	1350746	20	2	3
856845	3561700	20	2	3

20원이라니..(그것도 2회에)

가장 적은 횟수로 롯데 온라인 플랫폼을 이용한 고객은 과연 누구일까?

```
df_segs[df_segs['BUY_COUNT'] == df_segs['BUY_COUNT'].min()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
94	3385915	39900	1	3
115	3823306	39800	1	3
256	296056	39900	1	3
686	6096697	39000	1	3
852	1852714	39800	1	3
...
898324	1582601	56700	1	3
898343	5560617	53000	1	3
898357	357205	55000	1	3
898398	844375	59000	1	3
898406	720069	32000	1	3

19114 rows × 4 columns

Group 1 기초 탐색 및 통계량

Group 1 총 구매횟수는 평균보다 적지만 총구매금액이 평균보다 높게 형성된 고객 그룹

그룹1의 킹스러운 소비요정은 누구일까?

```
Agroup[Agrou[ 'PD_BUY_TOT' ] == Agroup[ 'PD_BUY_TOT' ].max( ) ]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
149473	923121	37935000	3	1

3천7백9십만원을 단 3회 구매로 소비하신

아름다운 고객님

진정한 플렉서를 찾아보자! (한 번에 통 크게 쓰신 분 후즈댓)

```
Thanksbro = Agroup[Agrou[ 'BUY_COUNT' ] == Agroup[ 'BUY_COUNT' ].min( ) ]  
Thanksbro[Thanksbro[ 'PD_BUY_TOT' ] == Thanksbro[ 'PD_BUY_TOT' ].max( ) ]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
513705	4802049	5621440	1	1

한 번에 5백6십2만원 Thanks bro!

Group 2 기초 탐색 및 통계량

Group 2 총 구매횟수가 평균 이상이고 총구매금액도 평균 이상인 매우 이상적인 고객 그룹

그룹 내 제일 많은 소비를 한 고객 찾기 VVIP 플렉스 소비왕

```
df_segs[df_segs['PD_BUY_TOT'] == df_segs['PD_BUY_TOT'].max()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
787566	1466175	634150900	2194	2

6개월 동안 총 6억3천4백만원 FLEX! 사랑합니다 고객님

롯데 온라인 쇼핑 중독자님 누구신가요

```
df_segs[df_segs['BUY_COUNT'] == df_segs['BUY_COUNT'].max()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
591397	4736937	183191200	3154	2

3154번에 1억8천3백만원 감사합니다

(6개월 동안 평균 하루 17번을 구매.. 가능?)

Group 3 기초 탐색 및 통계량

Group 3 총구매횟수가 평균 이하이고 총구매금액도 평균 이하인, 롯데 입장에서 매우 슝스한 고객 그룹

가장 적은 금액을 소비한 고객은 누구일까?

```
Cgroup[Cgroup[ 'PD_BUY_TOT' ] == Cgroup[ 'PD_BUY_TOT' ].min()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
416307	5234821	20	2	3
624782	479115	20	2	3
728198	1350746	20	2	3
856845	3561700	20	2	3

배송비가 2,500원 아닌가요

가장 적은 횟수로 롯데 온라인 쇼핑을 한 고객 알아보기!

```
bunbalhera = Cgroup[Cgroup[ 'BUY_COUNT' ] == Cgroup[ 'BUY_COUNT' ].min()]  
bunbalhera[bunbalhera[ 'PD_BUY_TOT' ] == bunbalhera[ 'PD_BUY_TOT' ].min()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
896902	1929700	100	1	3

Group 4 기초 탐색 및 통계량

Group 4 총구매횟수가 평균 이상인데 총구매금액은 평균 이하인 고객 그룹

적은 횟수로 적은 금액을 소비한 분발이 필요한 고객 찾기

```
bunbalhera = Dgroup[Dgroup['BUY_COUNT'] == Dgroup['BUY_COUNT'].min()]
bunbalhera[bunbalhera['PD_BUY_TOT'] == bunbalhera['PD_BUY_TOT'].min()]
```

	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
	261	5746813	400	4
	156411	2672577	400	4
	259941	1581391	400	4
	467701	4293188	400	4
	519609	892189	400	4

그룹 내 제일 구매 횟수가 많은 고객은? feat.프로자잘소비러

```
Dgroup[Dgroup['BUY_COUNT'] == Dgroup['BUY_COUNT'].max()]
```

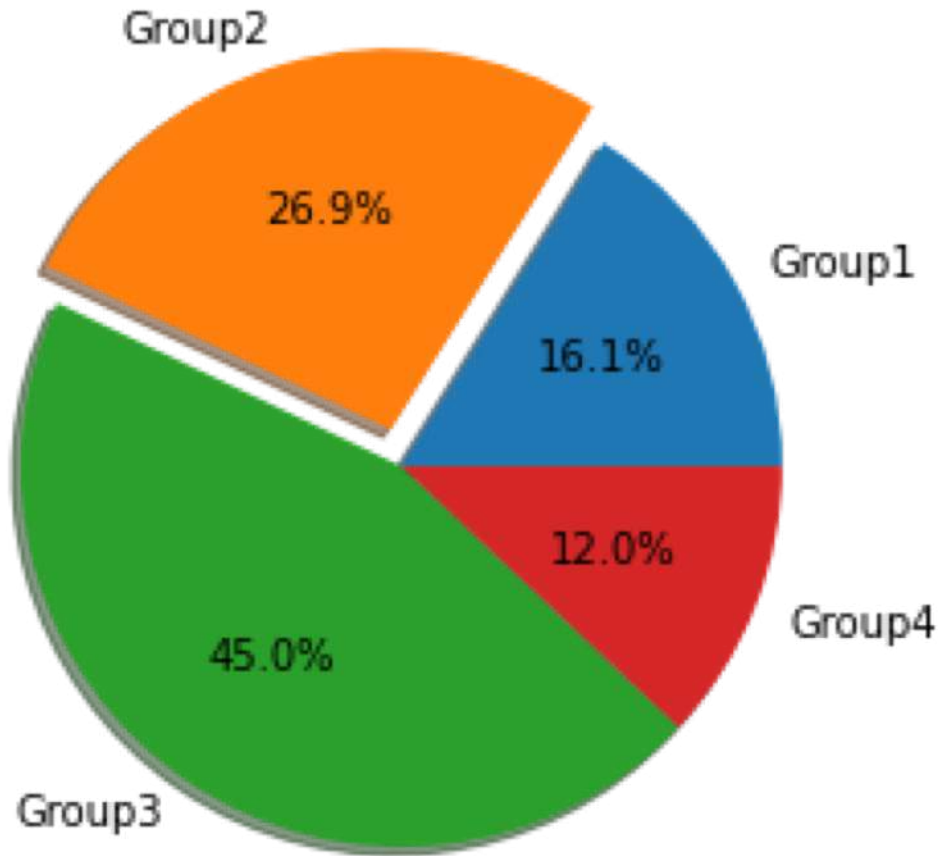
	CLNT_ID	PD_BUY_TOT	BUY_COUNT	CUS_SEG
	574888	2759633	125700	58

4. 데이터 시각화 & 인사이트 도출



데이터 시각화 & 인사이트 고객파악하기

< 그룹별 비중 파이 차트 (%) >

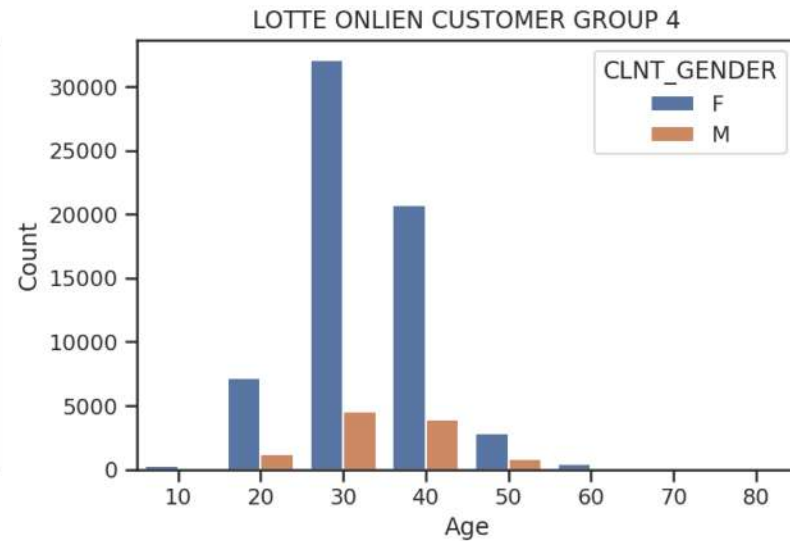
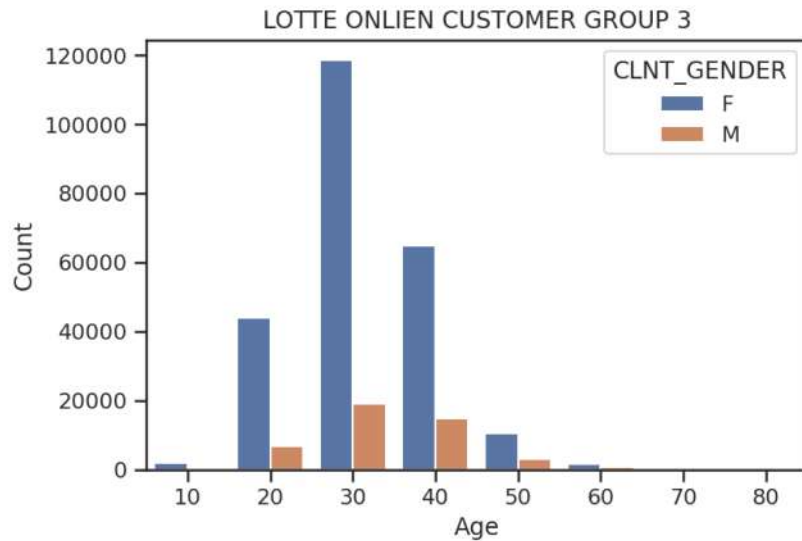
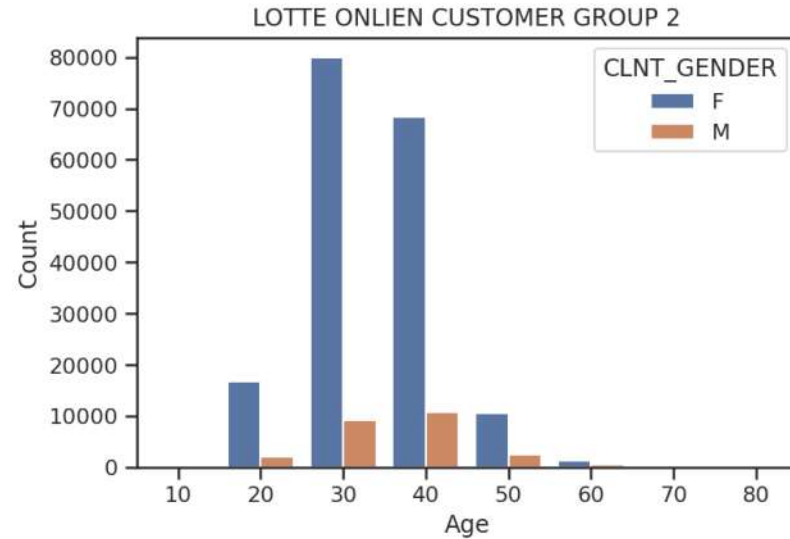
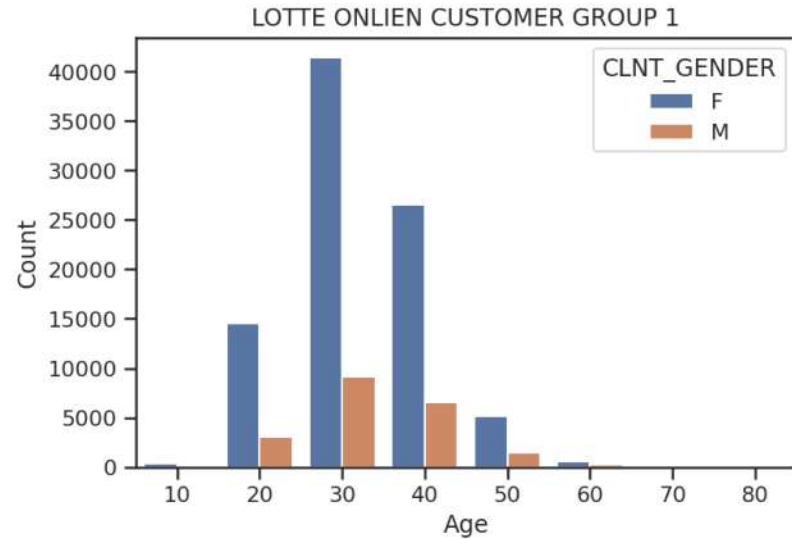


구매력으로 본 롯데의 고객은
어떻게 구성되어 있을까?

구매력이 낮은 Group3이 거의 절반을 차지함!

데이터 시각화 & 인사이트 고객파악하기

< 그룹별 연령x성별 그룹바 차트 (명) >



롯데의 주 고객층은
어떤 사람일까?

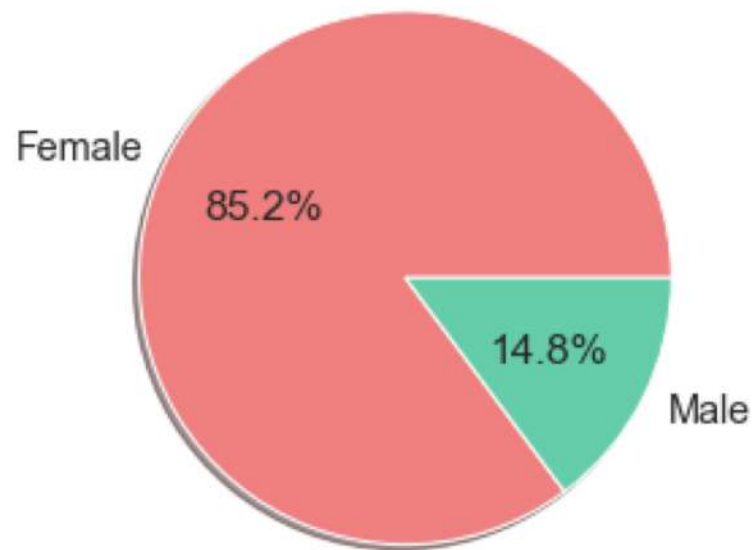
모든 그룹에서 30대 여성,
40대 여성, 20대 여성 순으로
구매자 수가 가장 많았다

데이터 시각화 & 인사이트 고객파악하기

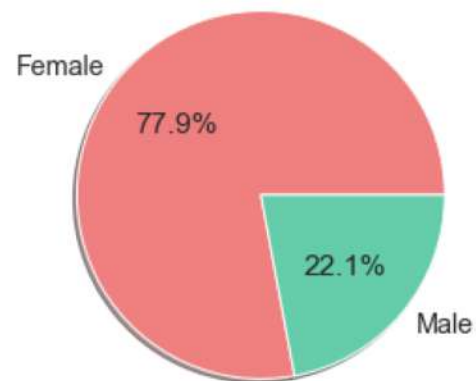
총구매금액 기준, 구매력이 있는 고객은 남성일까 여성일까?

< 그룹별 성별 비중, 파이 차트 (%) >

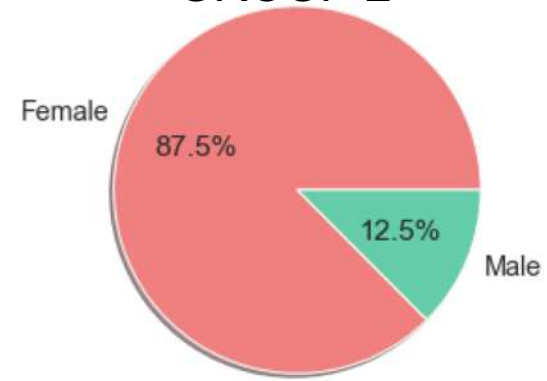
<고객 전체 평균>



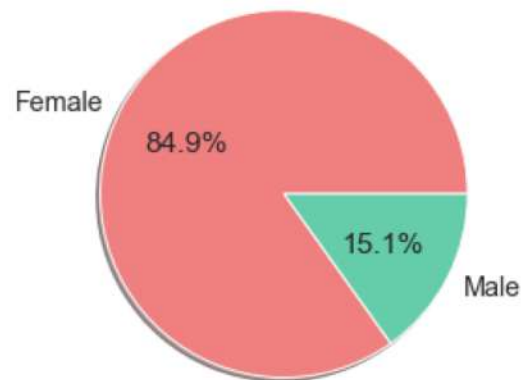
GROUP 1



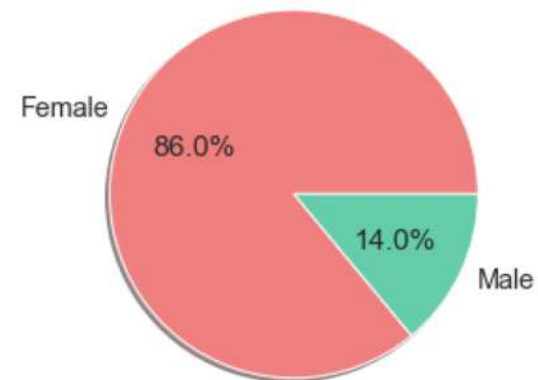
GROUP 2



GROUP 3



GROUP 4

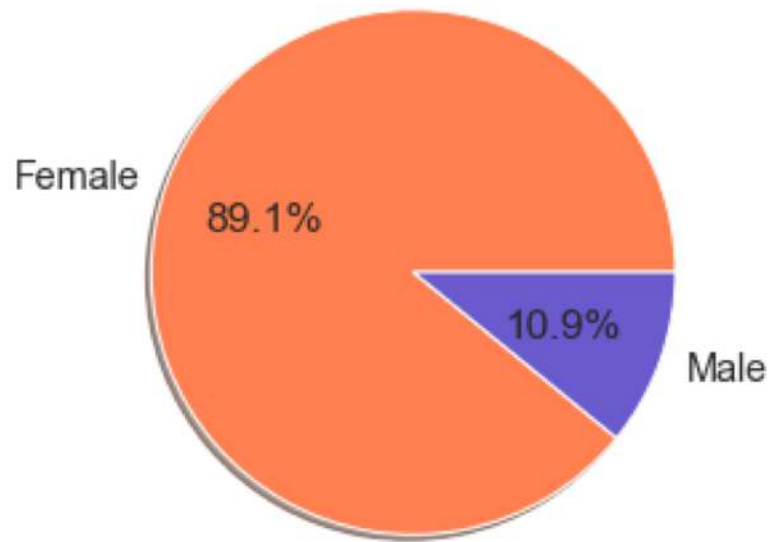


데이터 시각화 & 인사이트 고객파악하기

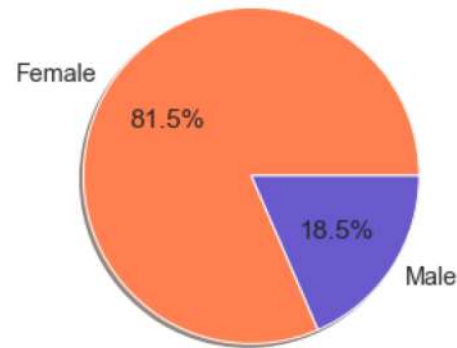
구매횟수 기준, 구매력이 있는 고객은 남성일까 여성일까?

< 그룹별 성별 비중, 파이 차트 (%) >

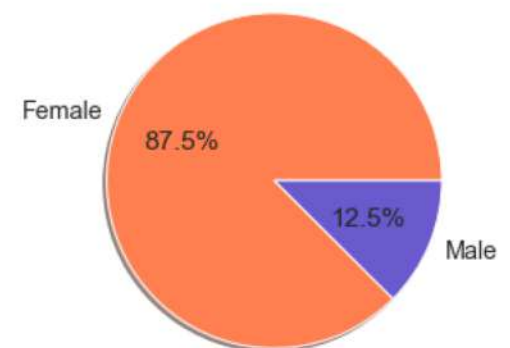
<고객 전체 평균>



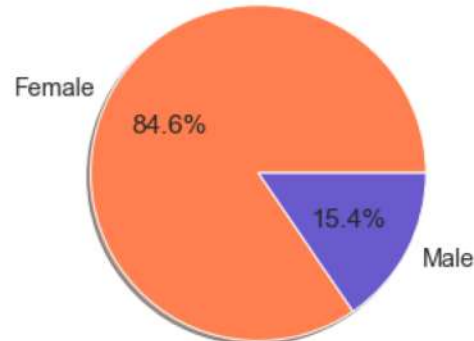
GROUP 1



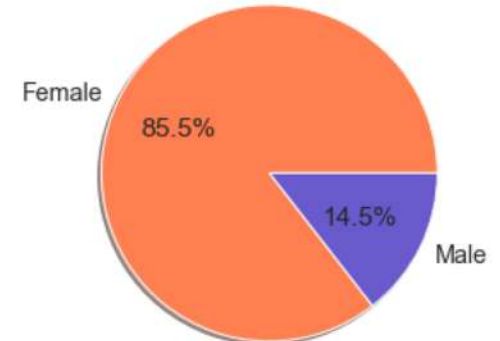
GROUP 2



GROUP 3



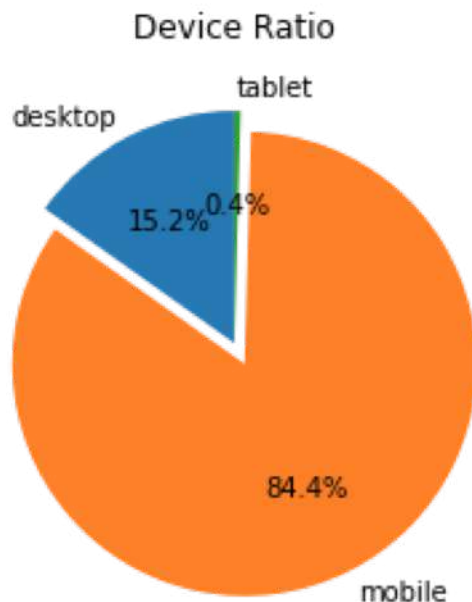
GROUP 4



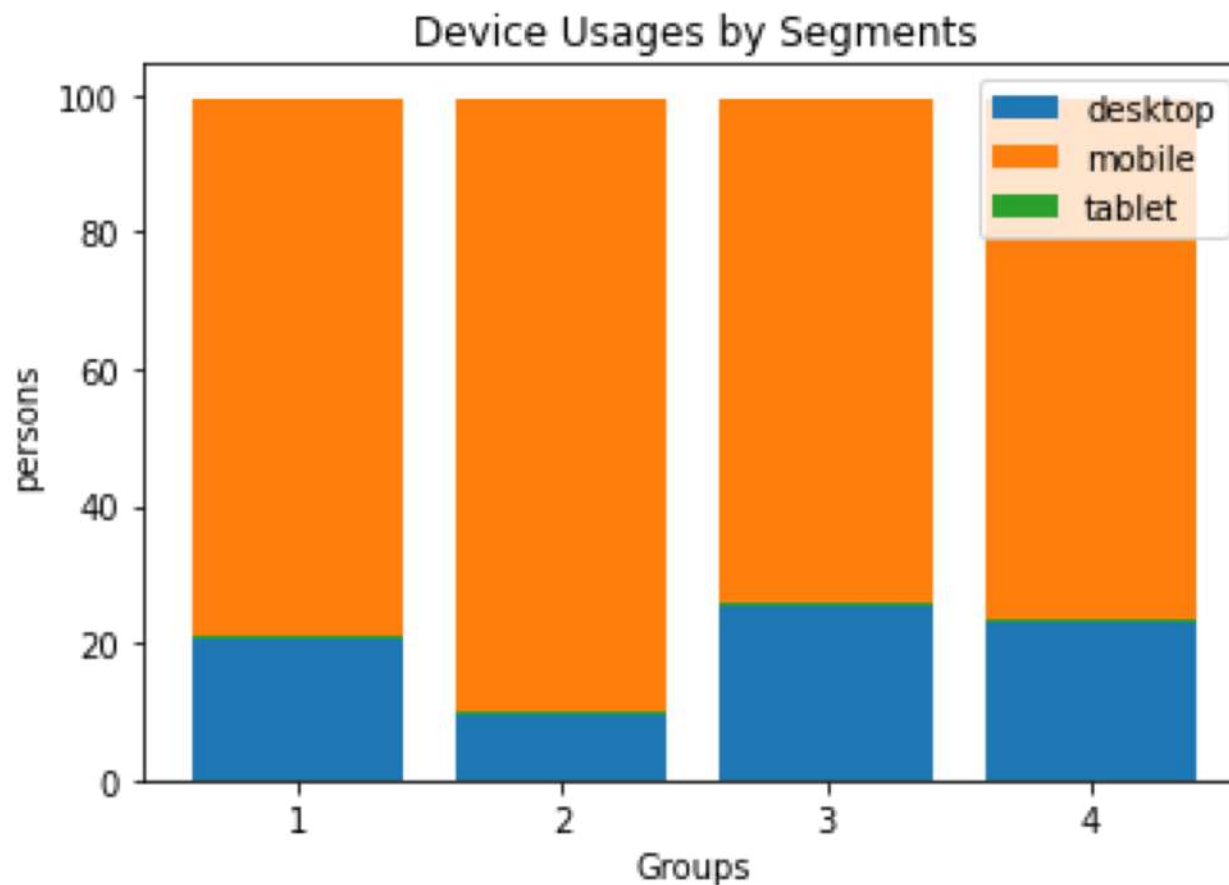
데이터 시각화 & 인사이트 고객파악하기

고객들은 어떤 디바이스로 롯데 온라인 쇼핑을 즐길까?

<고객 전체 평균>



<그룹별 디바이스 사용 비중, 100% 누적바 차트 (%)>



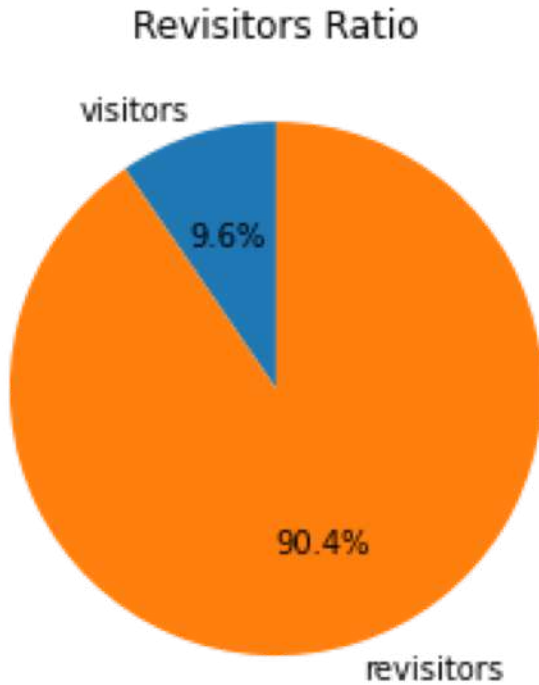
데이터 시각화 & 인사이트 고객파악하기

재구매율은 얼마나 될까?

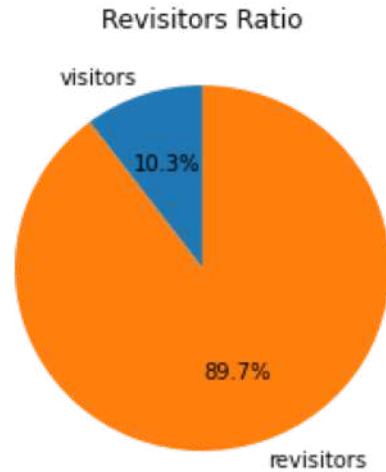
Group2의 재구매율이 96%로 매우 높게 나타났다.

< 그룹별 재구매율 파이 차트 (%) >

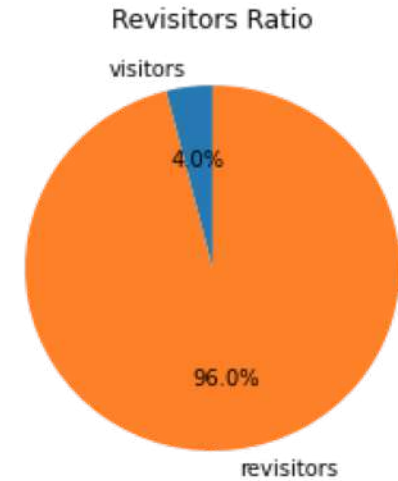
<고객 전체 평균>



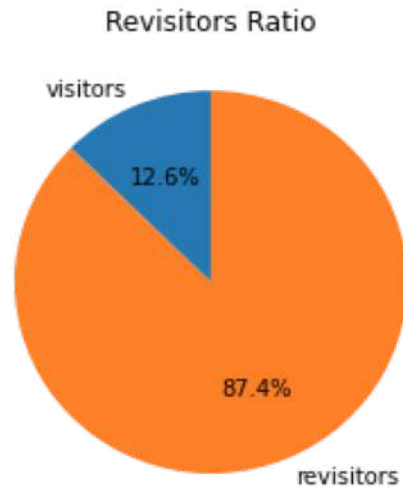
GROUP 1



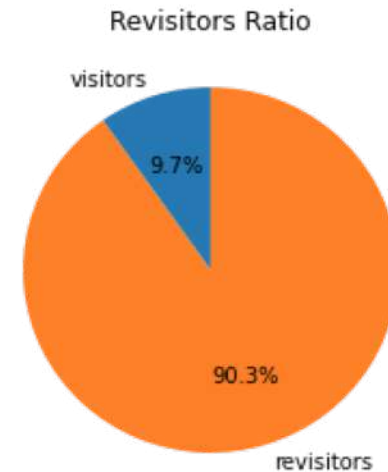
GROUP 2



GROUP 3

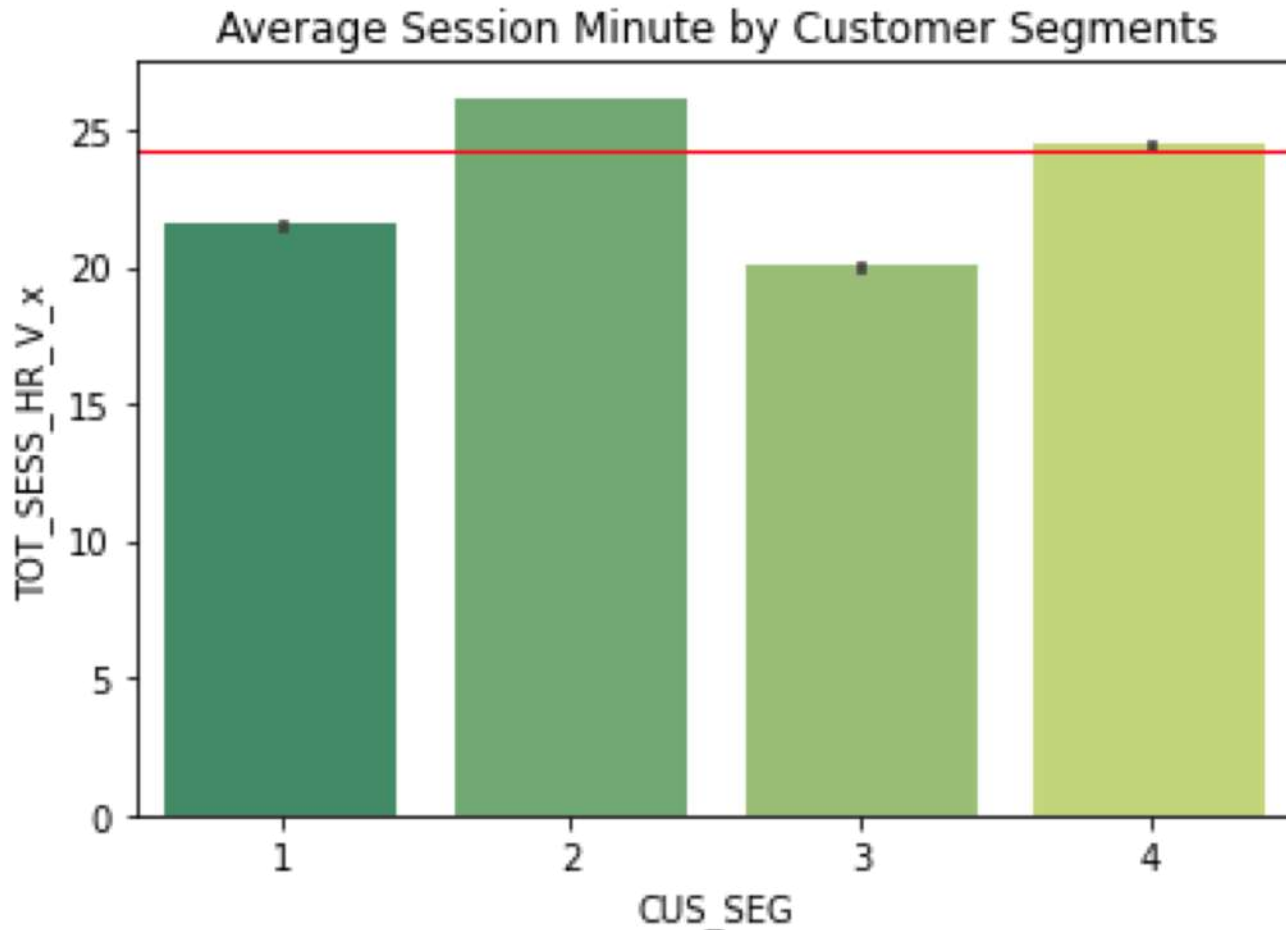


GROUP 4



데이터 시각화 & 인사이트 구매 행동 패턴 발견하기

< 그룹별 평균 세션시간, 막대그래프 (분) >

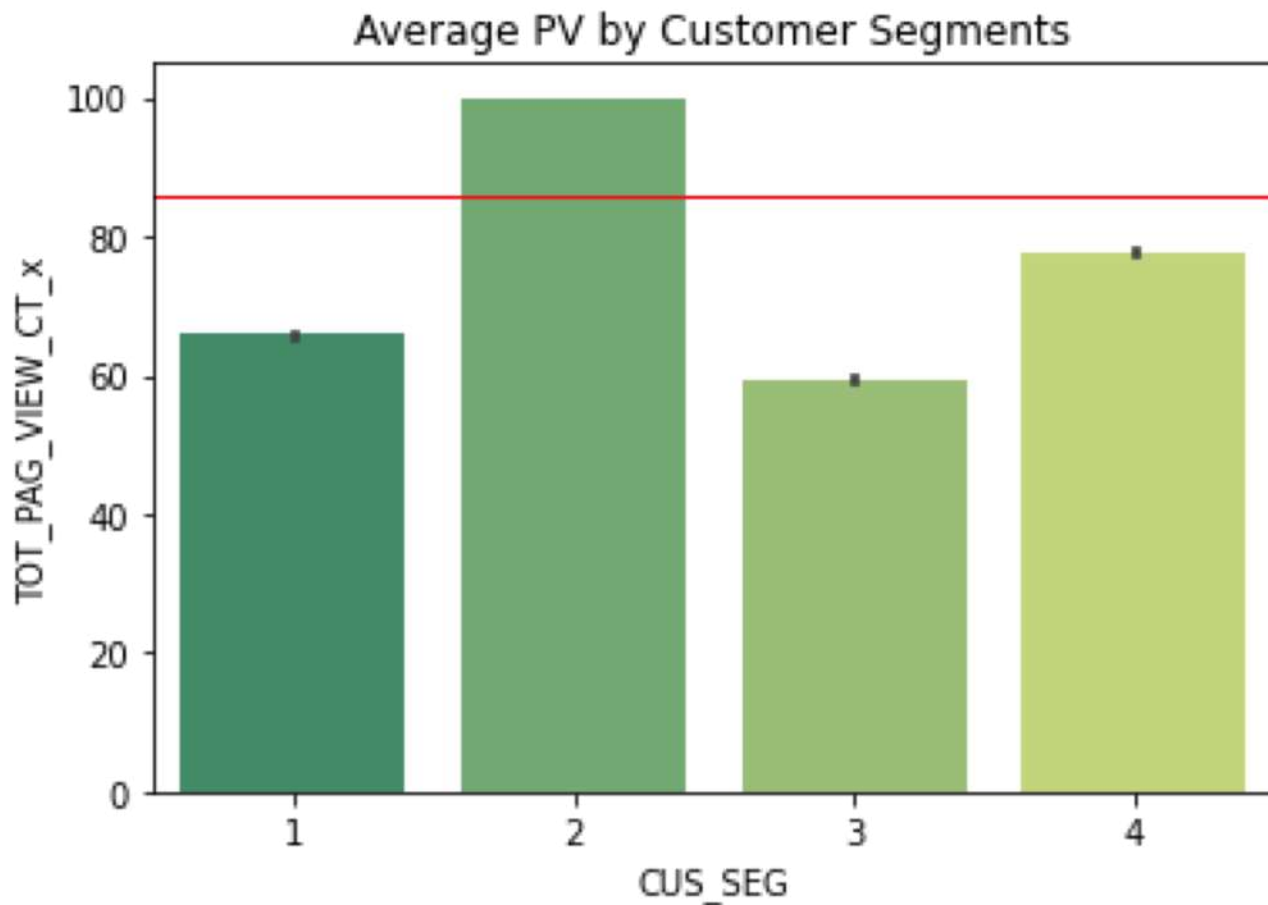


고객들은 한 번 구매할 때 얼마나 오랫동안 쇼핑할까?

전체 평균은 24.19분이었고, 구매횟수가 높은 그룹이 평균 이상으로 나와 세션 시간 또한 구매횟수가 높은 그룹과 관련있어 보인다

데이터 시각화 & 인사이트 구매 행동 패턴 발견하기

< 그룹별 평균 페이지뷰수, 막대그래프 (페이지) >

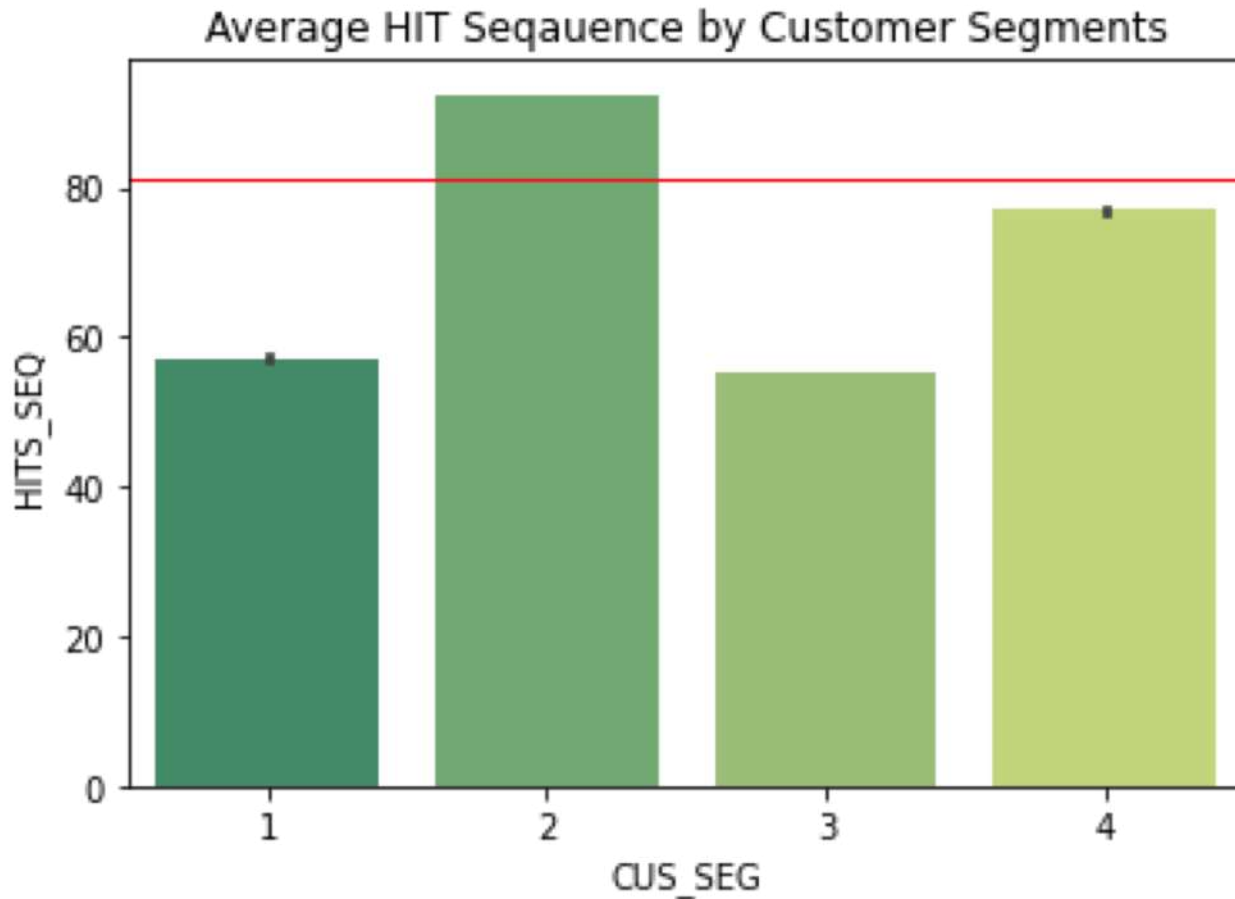


구매자는 평균 몇 페이지를
탐색할까?

전체 평균 85.6 페이지 뷰 수
구매력이 가장 높은 Group2만 평균
이상으로 나와 전체 평균을 높인것을
확인할 수 있었다

데이터 시각화 & 인사이트 구매 행동 패턴 발견하기

< 그룹별 세션시간, 평균 히트 시퀀스 (회) >



구매자는 몇 번의 클릭을 했을까?

Group 1과 Group3 은 상대적으로 낮지만 Group2는 평균 이상으로 클릭했다는 것을 확인할 수 있었다

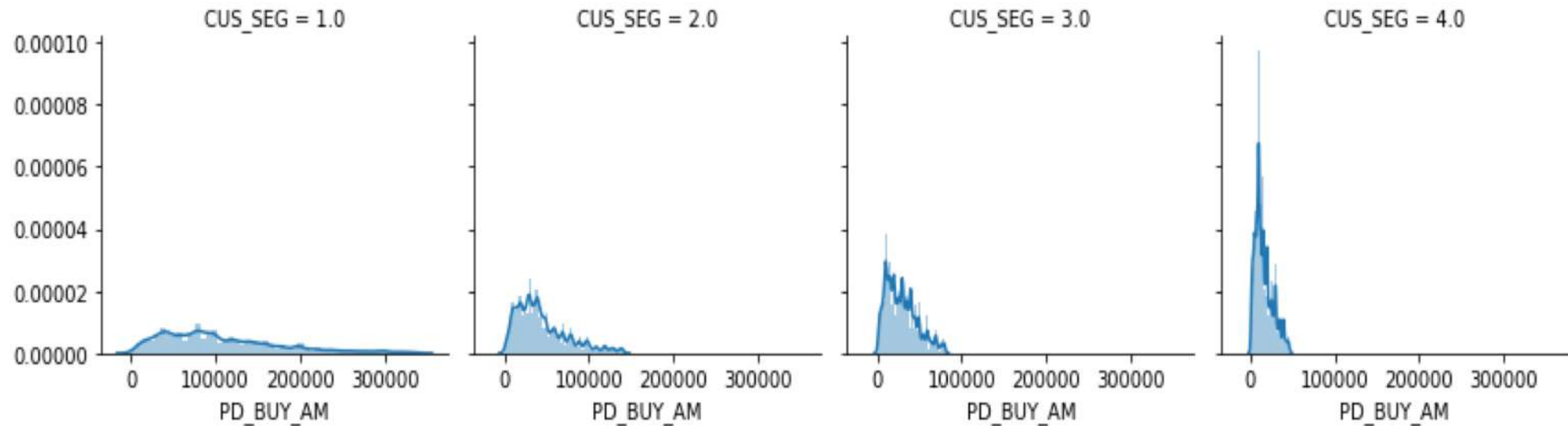
데이터 시각화 & 인사이트

고객 그룹별 1회 구매 금액

고객들은 한 번 구매할 때 얼마의 소비를 할까?

Group 1 > Group 2 > Group 3 > Group 4 순으로 고객의 소비 범위가 넓게 분포
Group 4의 경우 저가 제품 구매를 많이 하였음을 확인 할 수 있다

< 그룹별 1회 구매금액(이상치 제거), 히스토그램 >



가장 많이 구매한 제품의 금액은?

그룹별 최빈치: 79,000원 / 39,000원 / 39,000원 / 9,900원

데이터 시각화 & 인사이트

제품 카테고리 구분

고객들은 어떤 제품을 구매할까?

고객의 구매 패턴을 이해하기 위해 상품 분류

내구재 내구재는 **1년 이상** 사용이 가능하고 주로 **고가의 상품**으로 승용차, 가전제품, 가구 등이 해당

준내구재 준내구재는 **1년 이상** 사용이 가능하나 주로 **저가인 상품**으로 의복, 신발, 가방, 운동 및 오락용품 등이 해당

비내구재 비내구재는 주로 **1년 미만** 사용되는 상품으로 음식료품, 화장품, 서적 및 문구 등이 해당

데이터 시각화 & 인사이트

제품 카테고리 구분

롯데에서 판매하는 제품을 상품 타입별로 그룹핑, 파생변수 생성

- 총 37개의 상품 대분류 기준을 통해 내구재/준내구재/비내구재를 구분

- 내구재: D(Durable), 12개

가구	D
계절가전	D
냉장/세탁가전	D
모바일	D
생활/주방가전	D
식기/조리기구	D
영상/음향가전	D
인테리어/조명	D
자동차용품	D
주방잡화	D
침구/수예	D
컴퓨터	D

- 비내구재: N(Non-durable), 13개

건강식품	N
과일	N
냉동식품	N
냉장식품	N
문구/사무용품	N
세제/위생	N
완구	N
원예/애완	N
음료	N
청소/세탁/욕실용품	N
축산물	N
출산/육아용품	N
화장품/뷰티케어	N

- 준내구재: S(Semi-durable), 12개

구기/필드스포츠	S
남성의류	S
상품권	S
속옷/양말/홈웨어	S
스포츠패션	S
시즌스포츠	S
아웃도어/레저	S
여성의류	S
유아동의류	S
패션잡화	S
퍼스널케어	S
헬스/피트니스	S

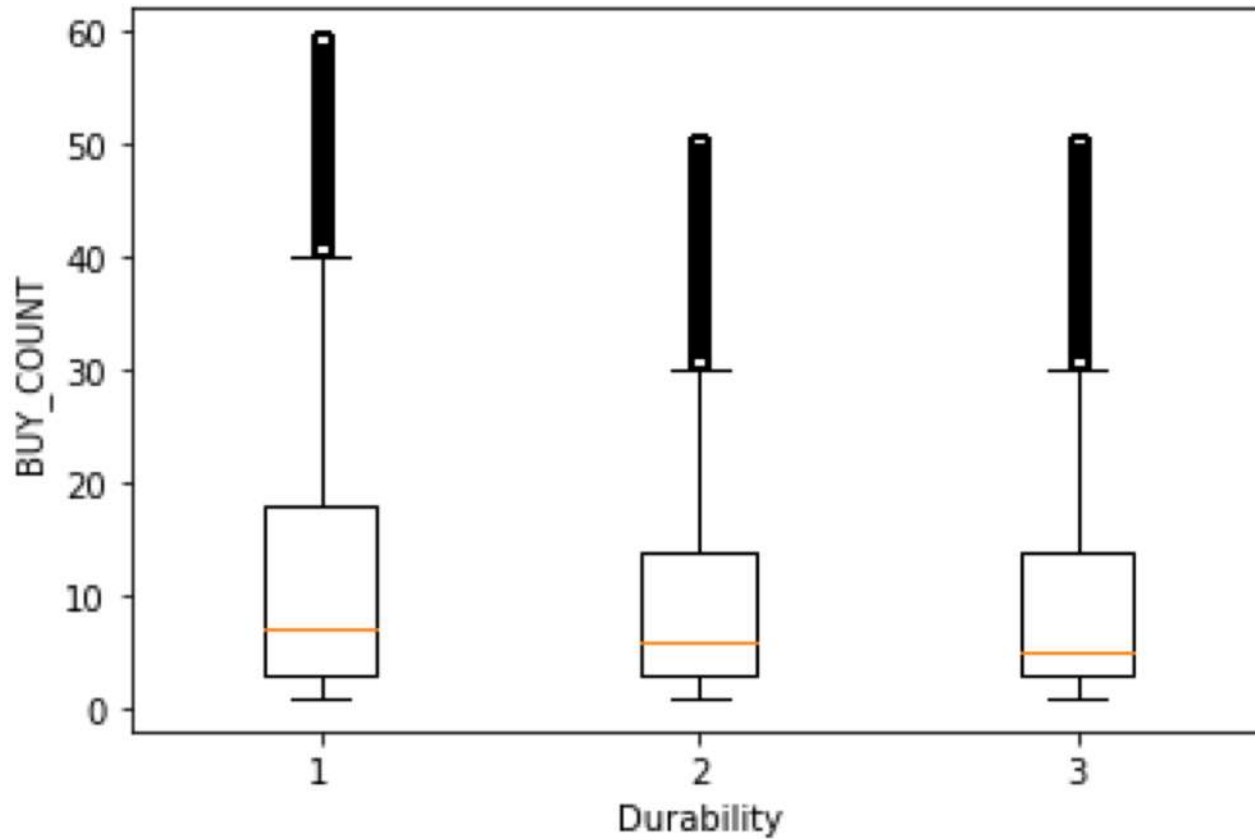
데이터 시각화 & 인사이트

상품 타입별 구매 횟수 차이

상품 타입별로 구매 횟수가 다르지 않을까?

상품 타입별 차이는 거의 나타나지 않았지만 준내구재(1)에서 구매횟수 값의 범위가 좀 더 넓었다
각 그룹에서 어느 제품을 사는지 분석하는 것이 좋을 듯!

< 상품 타입별 총 구매횟수(이상치 제거), 박스플롯 (회) >



1 = 준내구재(S)
2 = 비내구재(N)
3 = 내구재(D)

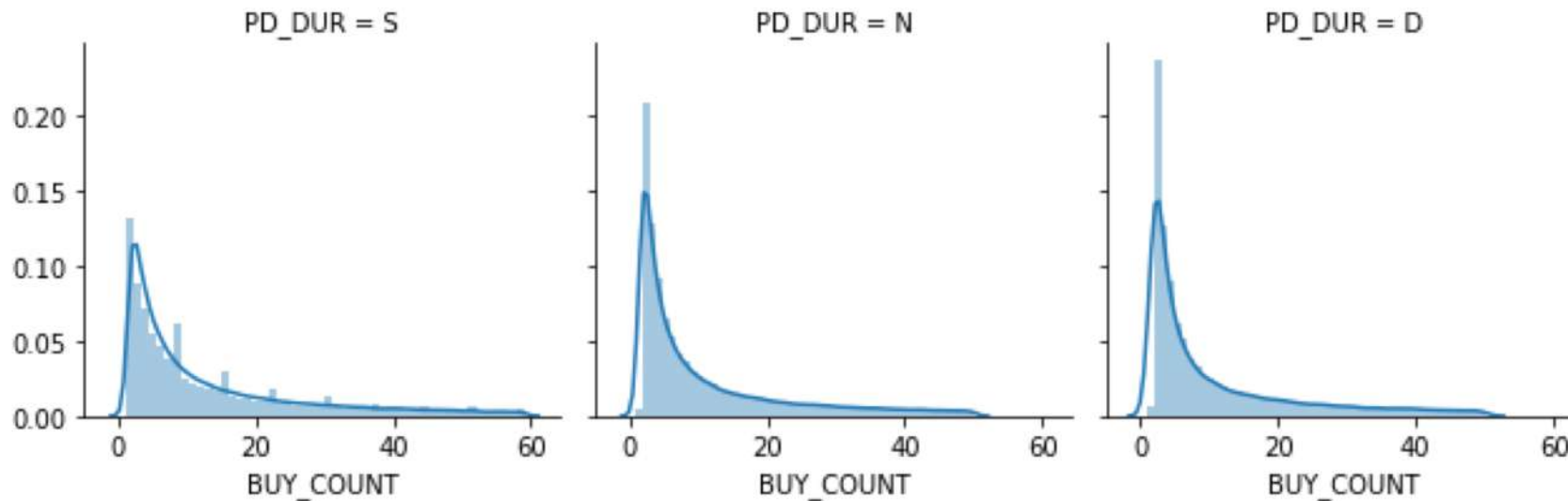
데이터 시각화 & 인사이트

상품 타입별 구매 횟수 차이 2

상품 타입별로 구매 횟수가 다르지 않을까?

히스토그램으로 보니깐 준내구재의 분포가 넓음을 확인!
최빈치를 보아 가장 많이 나타난 구매 횟수는 2회로 크게 다르지 않았다

< 상품 타입별 총 구매횟수(이상치 제거), 히스토그램 >



1 = 준내구재(S)
2 = 비내구재(N)
3 = 내구재(D)

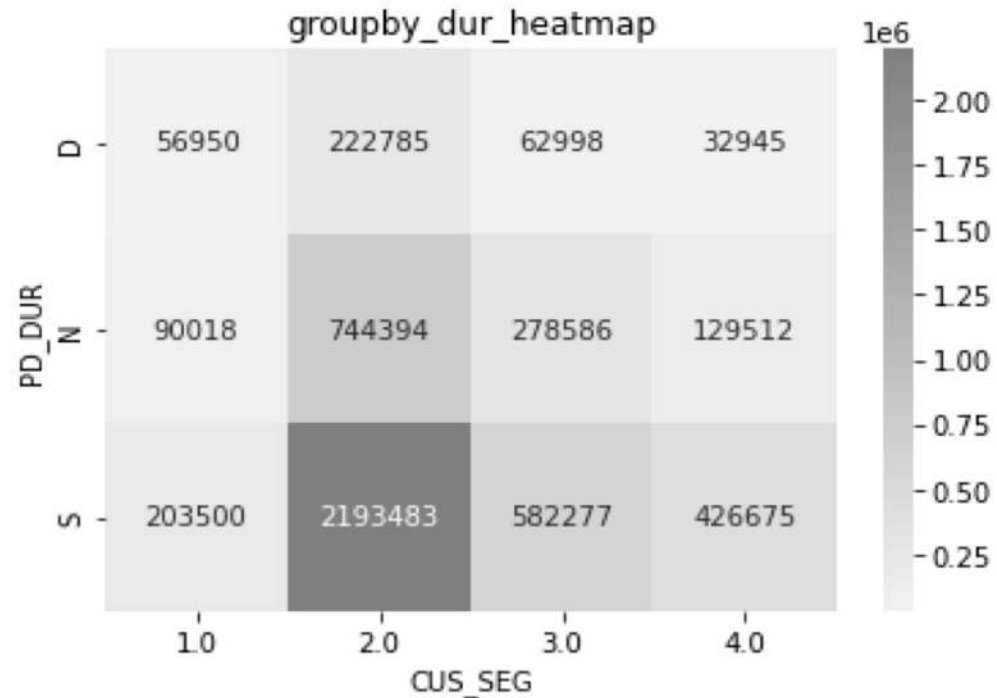
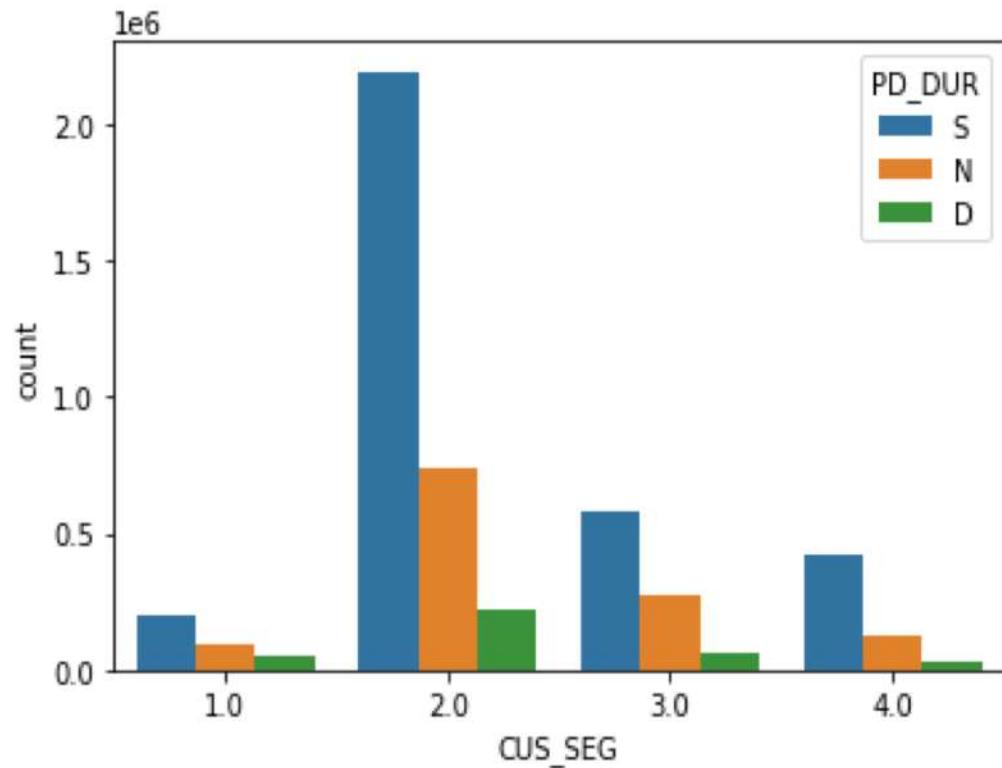
데이터 시각화 & 인사이트

상품 타입별 구매 횟수 차이 - 그룹별

고객 그룹별 상품 타입에 따른 구매 횟수는 어떨까?

각 그룹별 준내구재 > 비내구재 > 내구재 순으로 구매 비율이 높았다

<고객 그룹별 상품 타입에 따른 구매횟수, 막대 그래프&히트맵 (회)>



1 = 준내구재(S)
2 = 비내구재(N)
3 = 내구재(D)

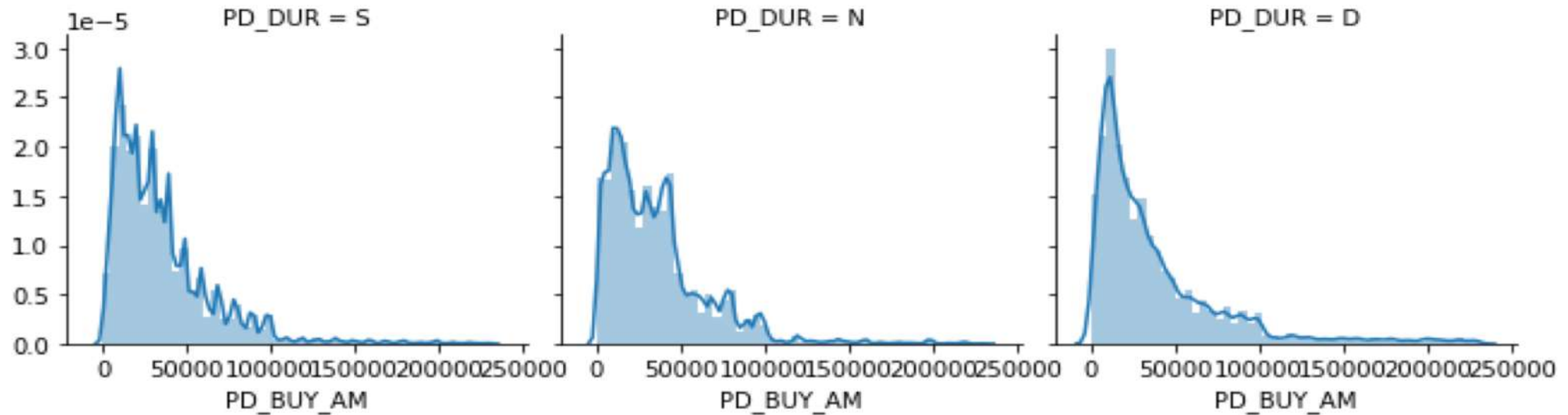
데이터 시각화 & 인사이트

상품 타입별 구매 금액 차이

상품 타입별로 구매 금액이 다르지 않을까?

상품 타입별 차이는 거의 없었지만 최빈치는 모두 39,000원으로
롯데 고객층은 주로 중저가 상품을 구매한다는 것을 확인하였다

< 상품 타입별 1회 구매금액(이상치 제거), 히스토그램 >



1 = 준내구재(S)
2 = 비내구재(N)
3 = 내구재(D)

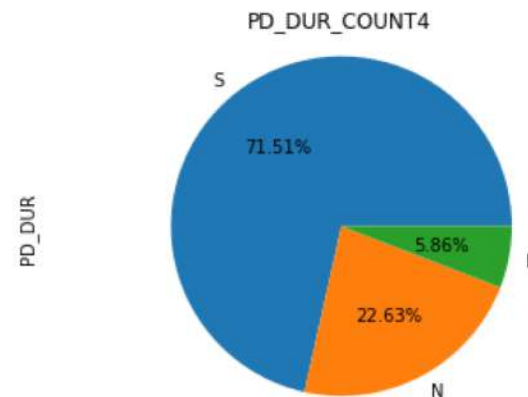
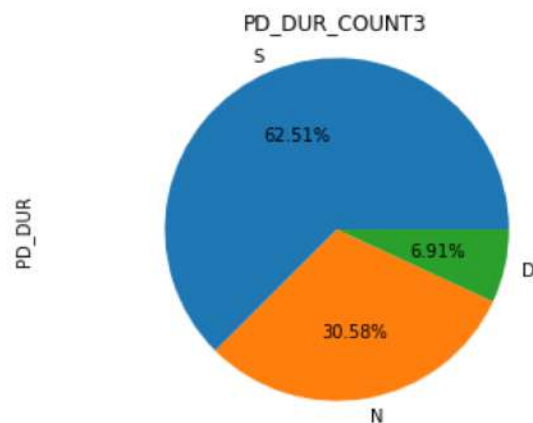
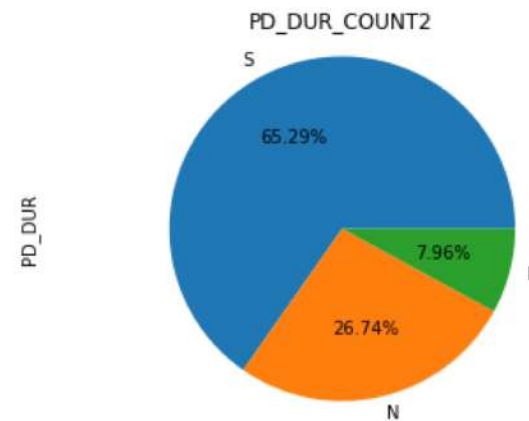
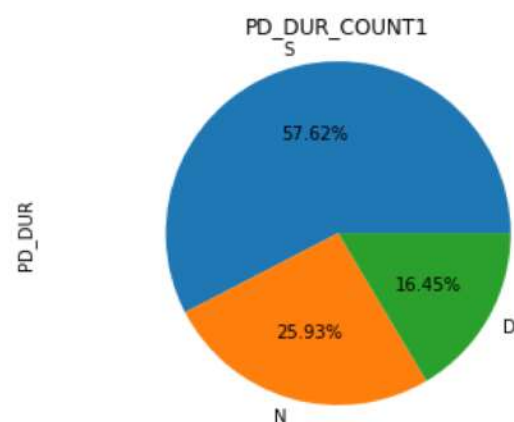
데이터 시각화 & 인사이트

상품 타입별 구매 금액 차이 - 그룹별

고객 그룹별 상품 타입에 따른 구매 비율은 그럼 어떨까?

각 그룹별 준내구재 > 비내구재 > 내구재 순으로 구매 비율이 높았다

< 고객 그룹별 상품 타입에 따른 총 구매비율(이상치 제거), 파이차트 (%) >



1 = 준내구재(S)
2 = 비내구재(N)
3 = 내구재(D)

데이터 시각화 & 인사이트

상품 타입별 구매횟수 및 금액 차이 - 그룹별

상품 타입으로 본 그룹별 특징을 유추하자면...

Group1 가격대가 높은 상품군을 조금만 사는 그룹

Group2 준내구재 비율과 횟수가 가장 높은 그룹. 롯데에서 가장 사랑받는 제품군은 준내구재라고 볼 수 있다.

Group3 비내구재 비율이 높은 편, 저렴한 제품을 많이 사는 그룹

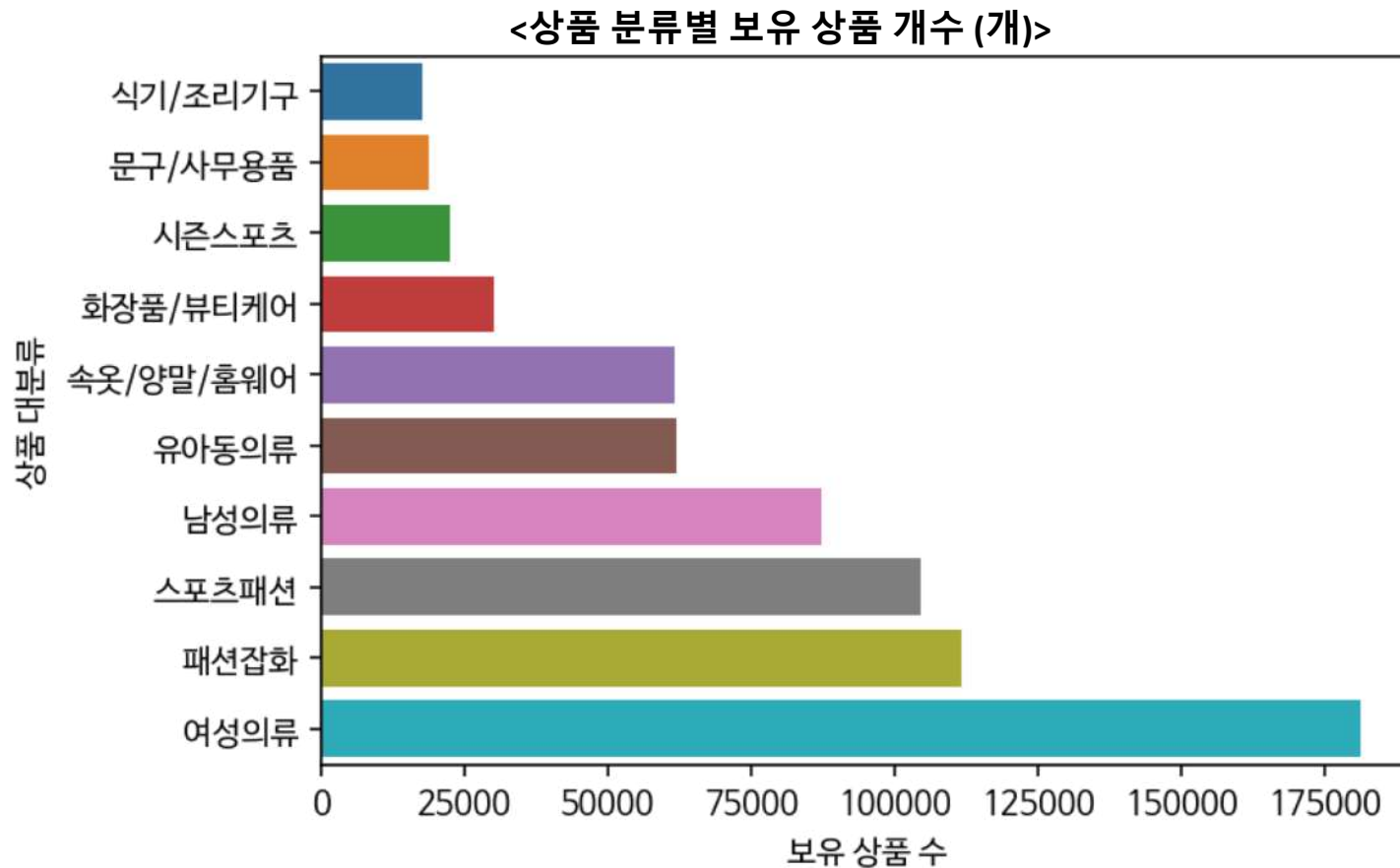
Group4 내구재의 구매 비율이 가장 낮아, 가성비를 추구하는 그룹

데이터 시각화 & 인사이트

상품 대분류별 보유 상품 수량

롯데 온라인 몰에서 가장 많이 보유하고 있는 상품은 어떤 것일까?

여성의류와 패션잡화, 스포츠패션 순으로 많이 보유하고 있음을 확인할 수 있었다

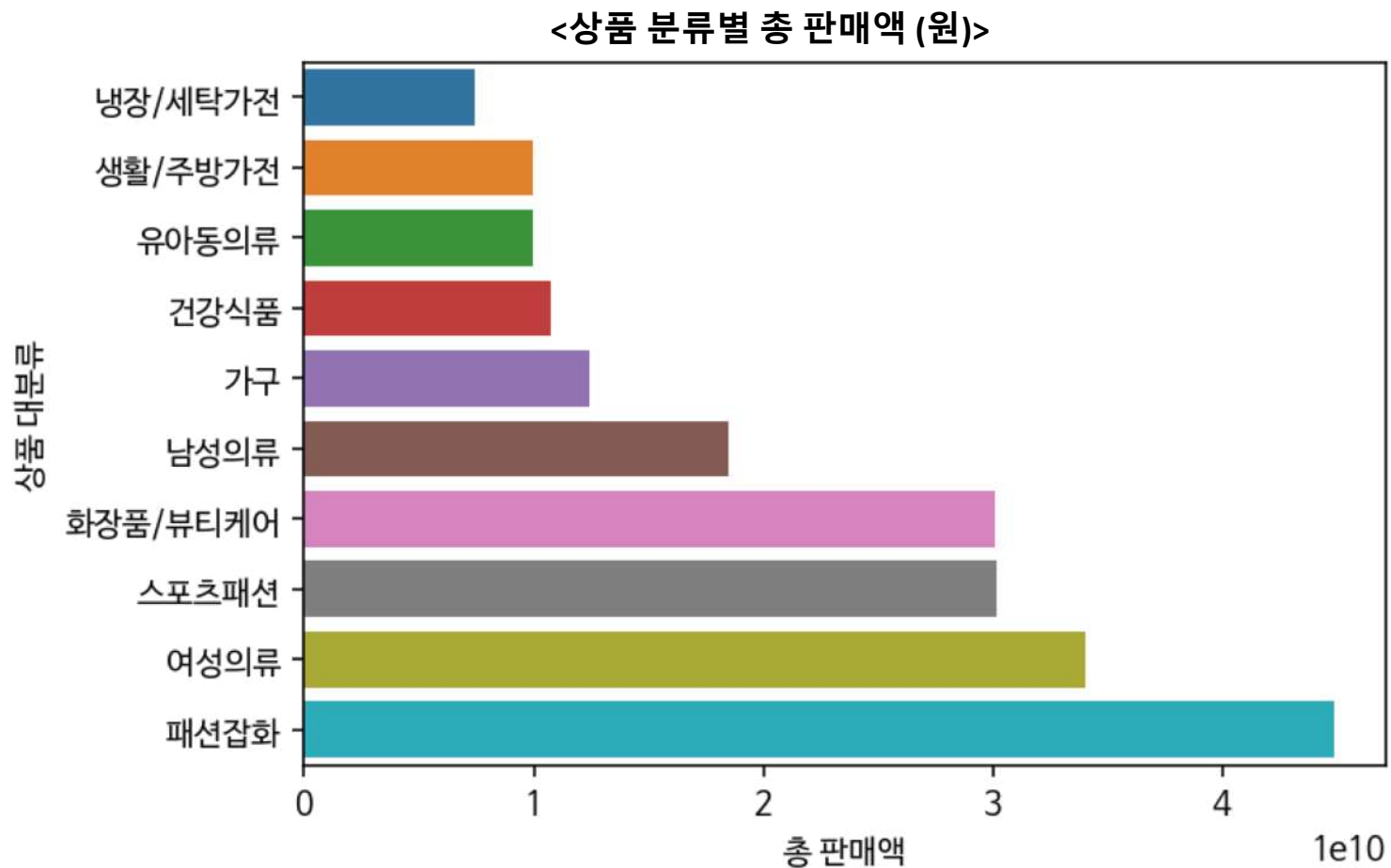


데이터 시각화 & 인사이트

상품 대분류별 보유 상품 판매액

롯데 온라인 몰에서 가장 판매액이 높은 상품은?

상품 보유 수량 TOP 3와 동일하게 패션잡화, 여성의류, 스포츠패션 순으로 매출이 발생하였다

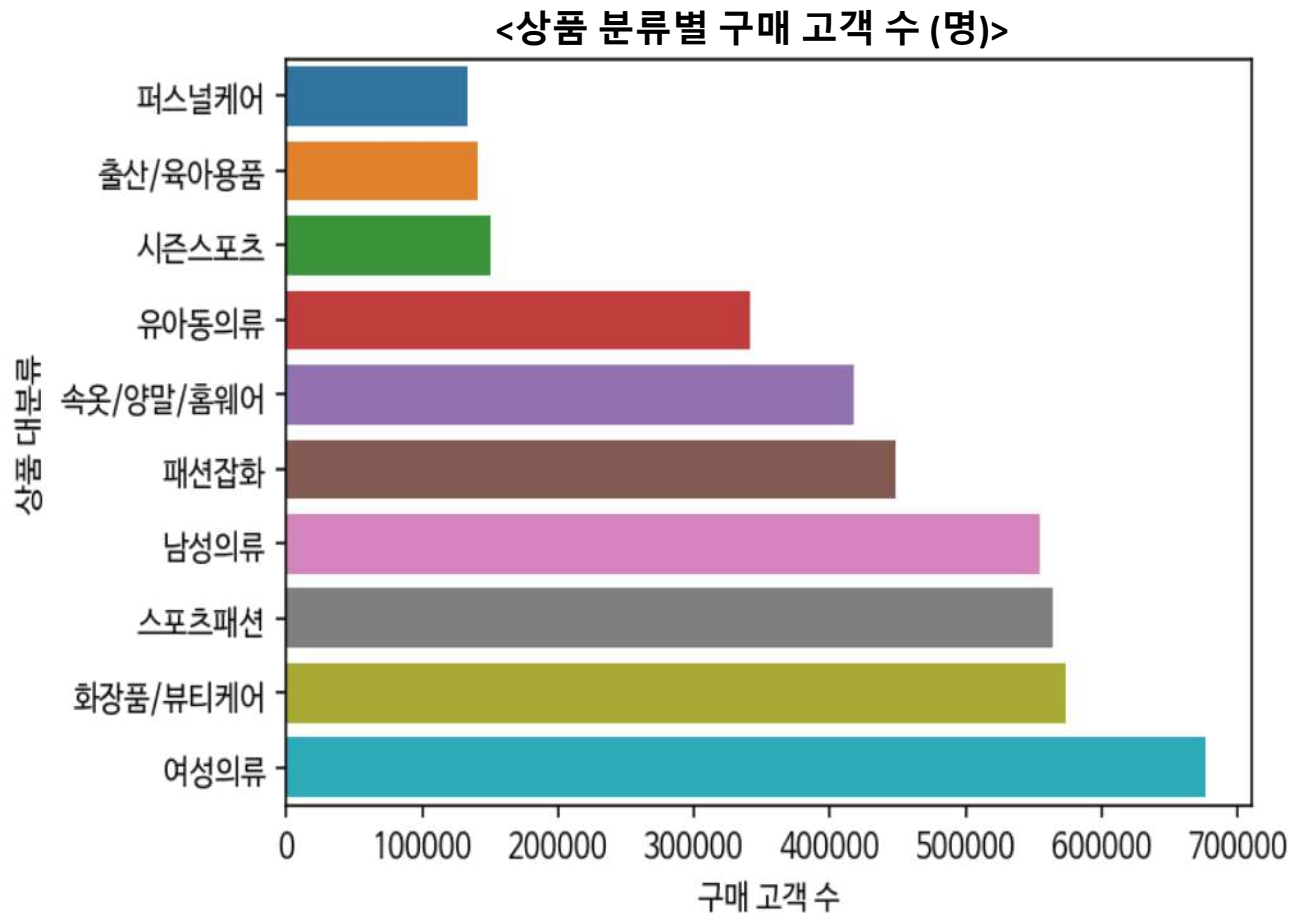


데이터 시각화 & 인사이트

상품 대분류별 구매고객수

가장 구매고객이 많은 상품은?

여성의류, 화장품/뷰티케어, 스포츠패션 순으로 구매 고객의 수를 확인할 수 있었다

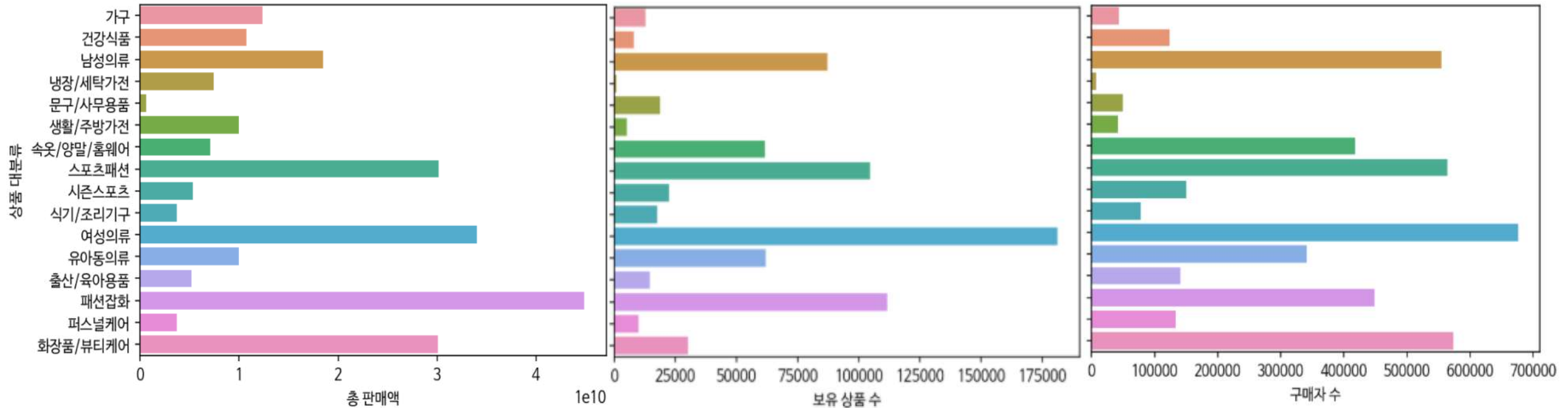


데이터 시각화 & 인사이트 상품

판매액, 상품 수, 구매자 수는 어떻게 다를까?

세가지 특징 비교를 통한 카테고리 추출

<상품 분류별 총 판매액, 보유 상품 개수, 구매자 수 - 동일 카테고리 비교>

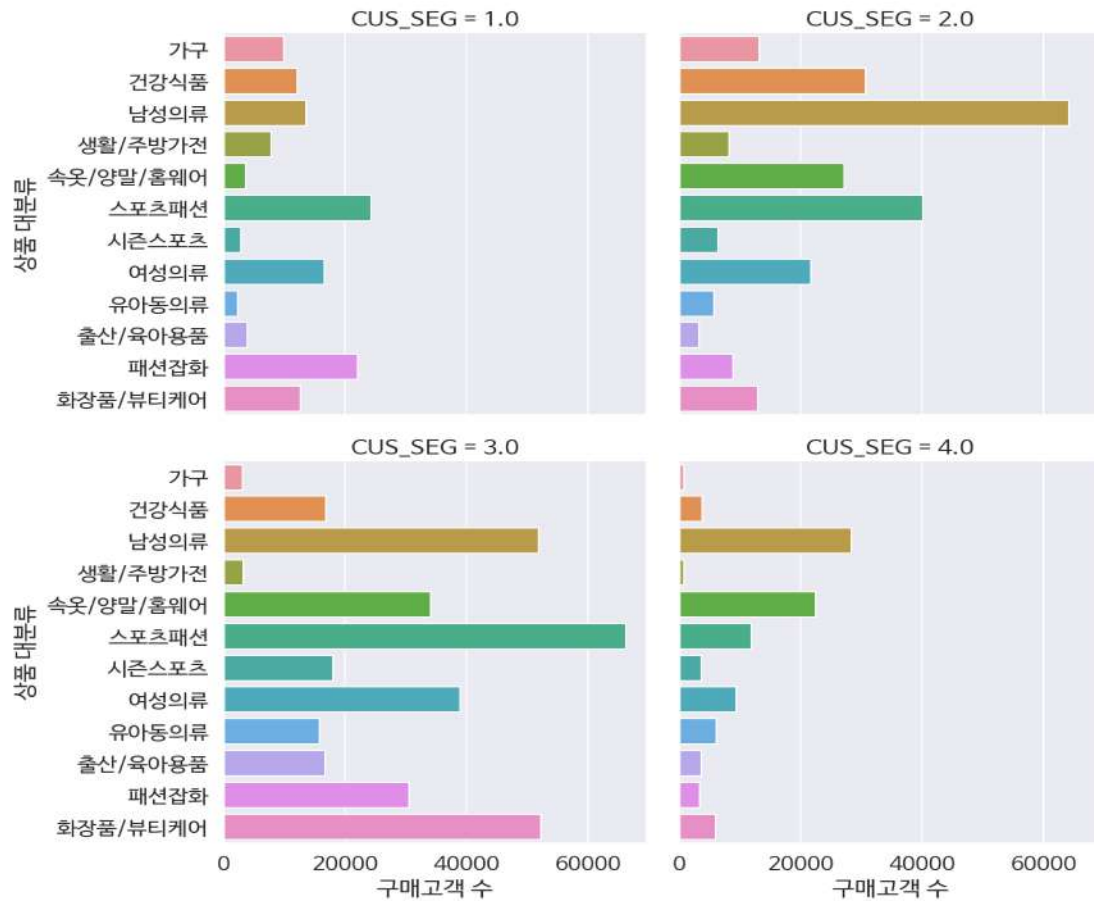


데이터 시각화 & 인사이트 상품

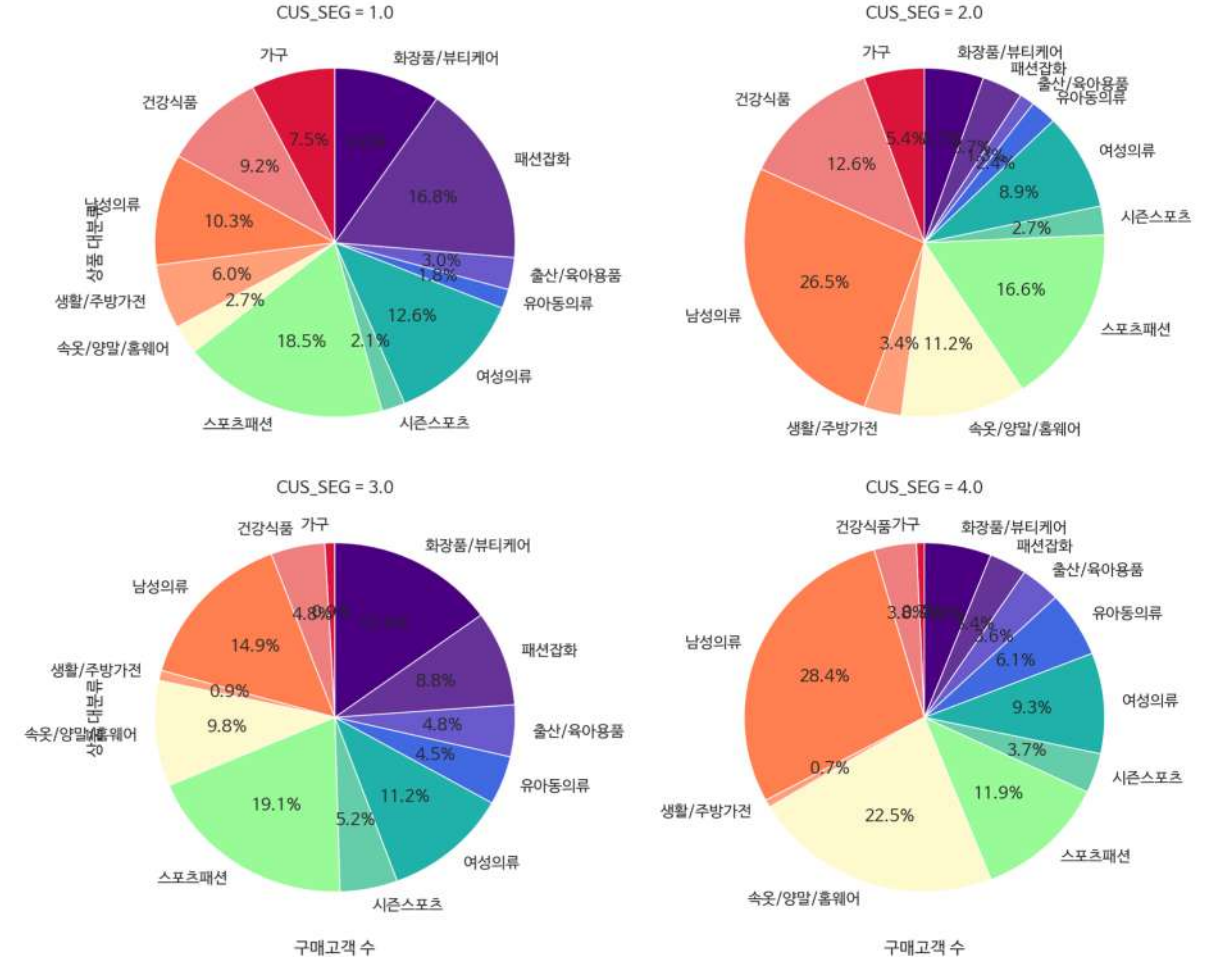
그룹별 구매고객수는 어떻게 다를까?

그룹별 주요 상품군 파악

<상품 분류별 구매고객 수 (명)>



<상품 분류별 구매고객 수 비율 (%)>

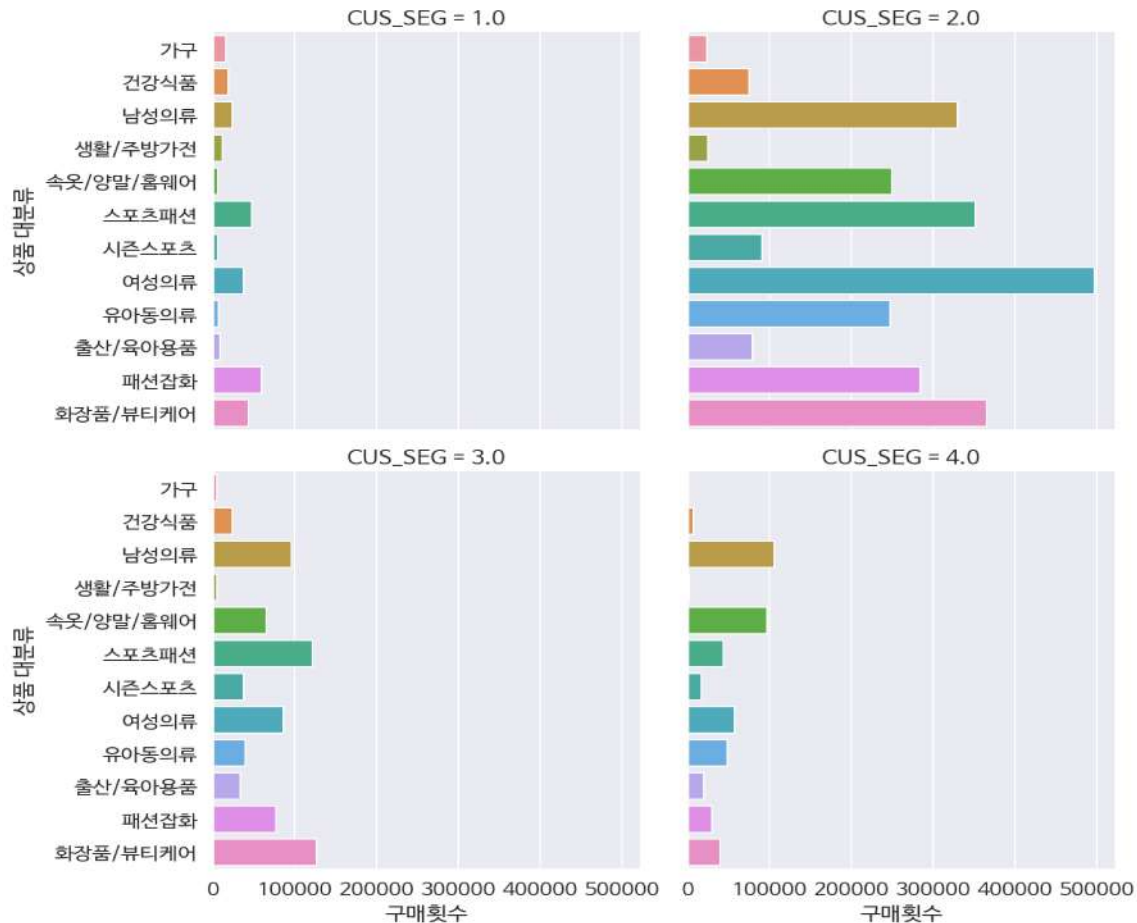


데이터 시각화 & 인사이트 상품

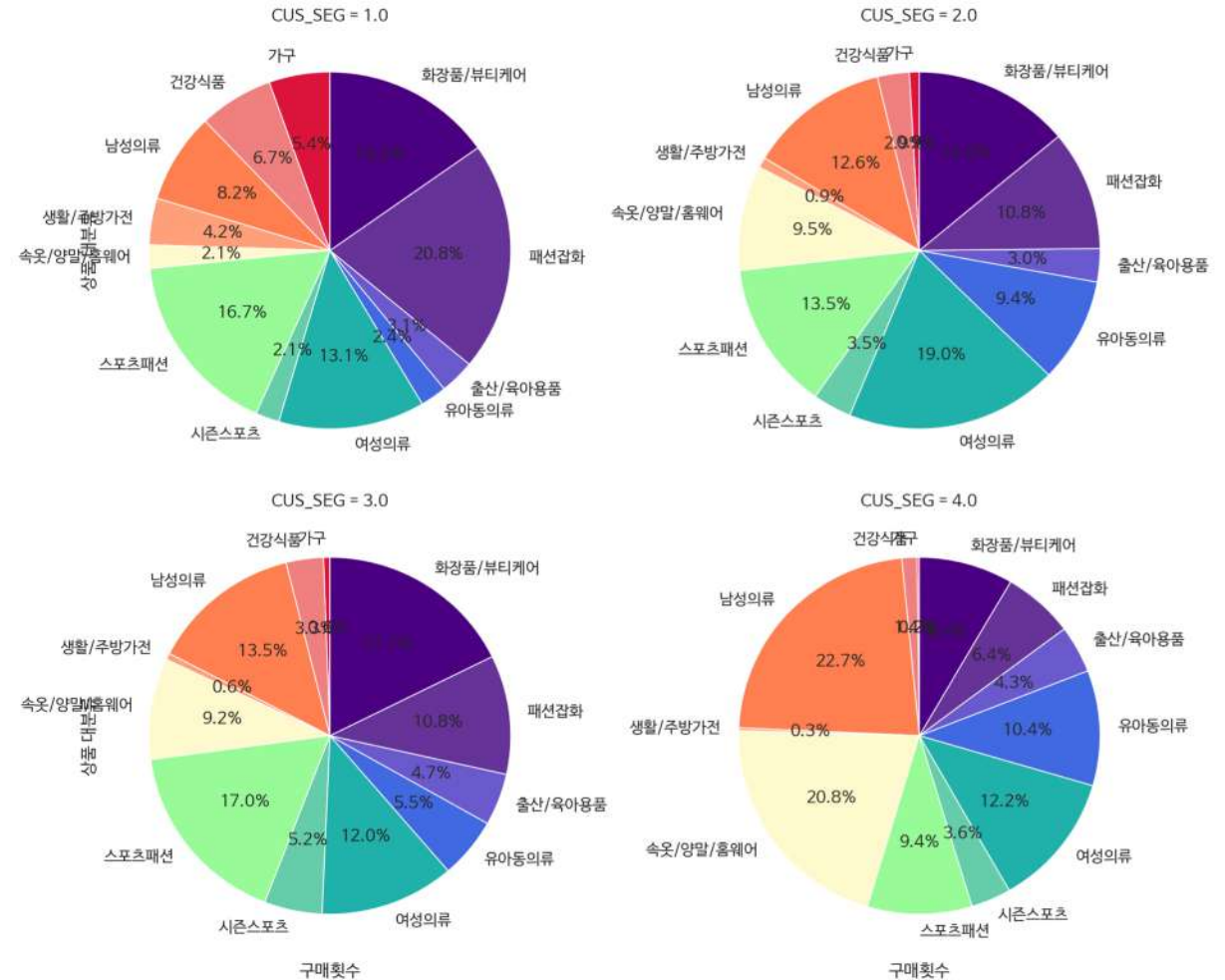
그룹별 구매횟수는 어떻게 다를까?

그룹별 높은 구매빈도의 상품군 파악

<상품 분류별 구매횟수 (번)>



<상품 분류별 구매횟수 비율(%)>

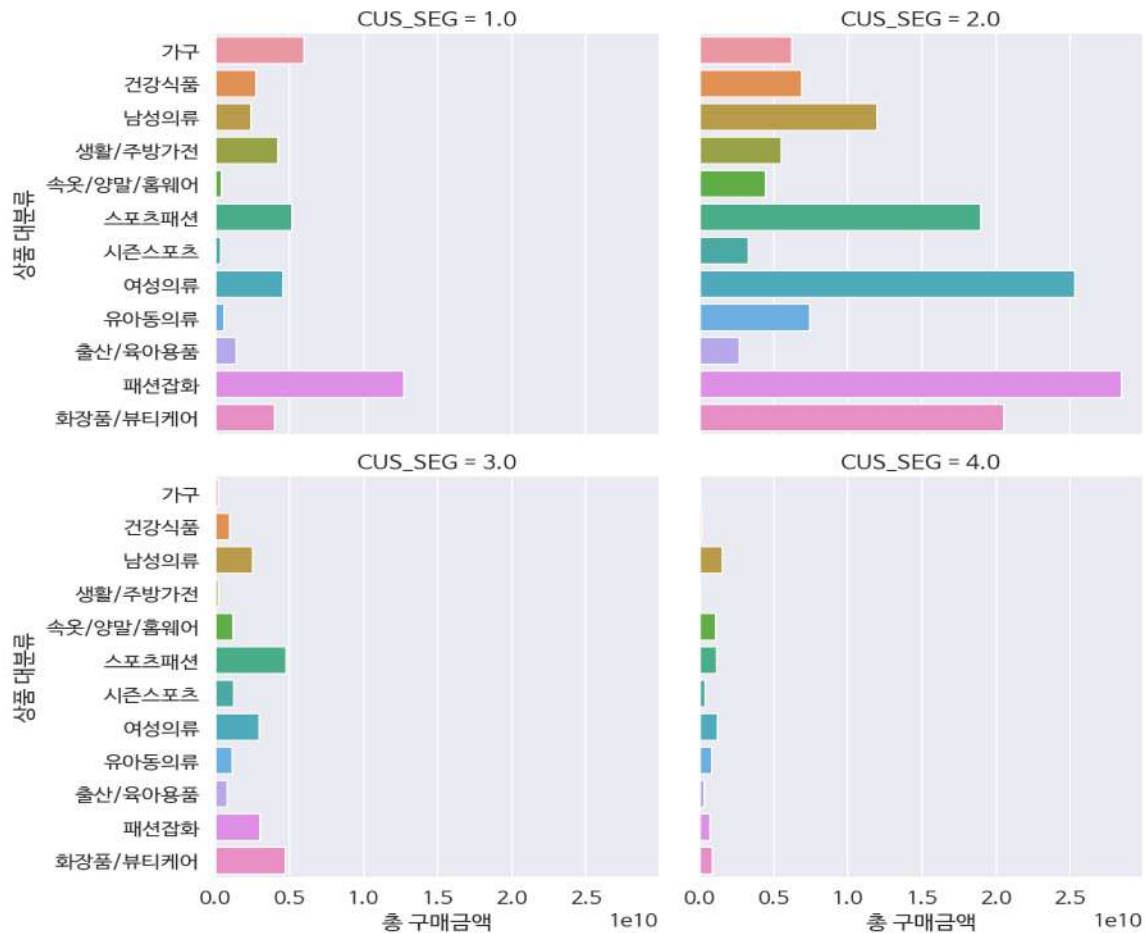


데이터 시각화 & 인사이트 상품

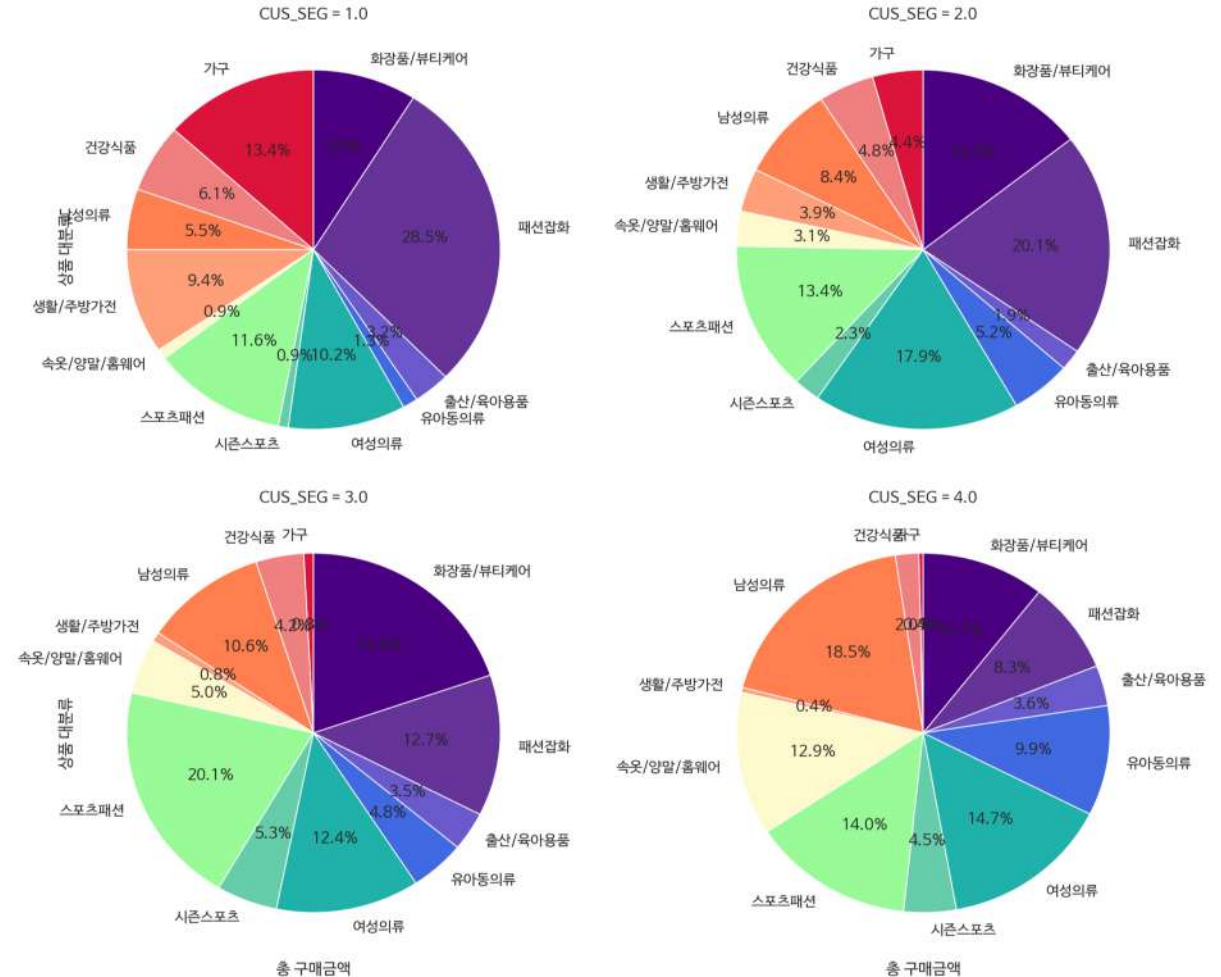
그룹별 구매금액 비율은 어떻게 다를까?

그룹별 높은 구매금액의 상품군 파악

<상품 분류별 구매금액 (원)>



<상품 분류별 구매금액 비율(%)>

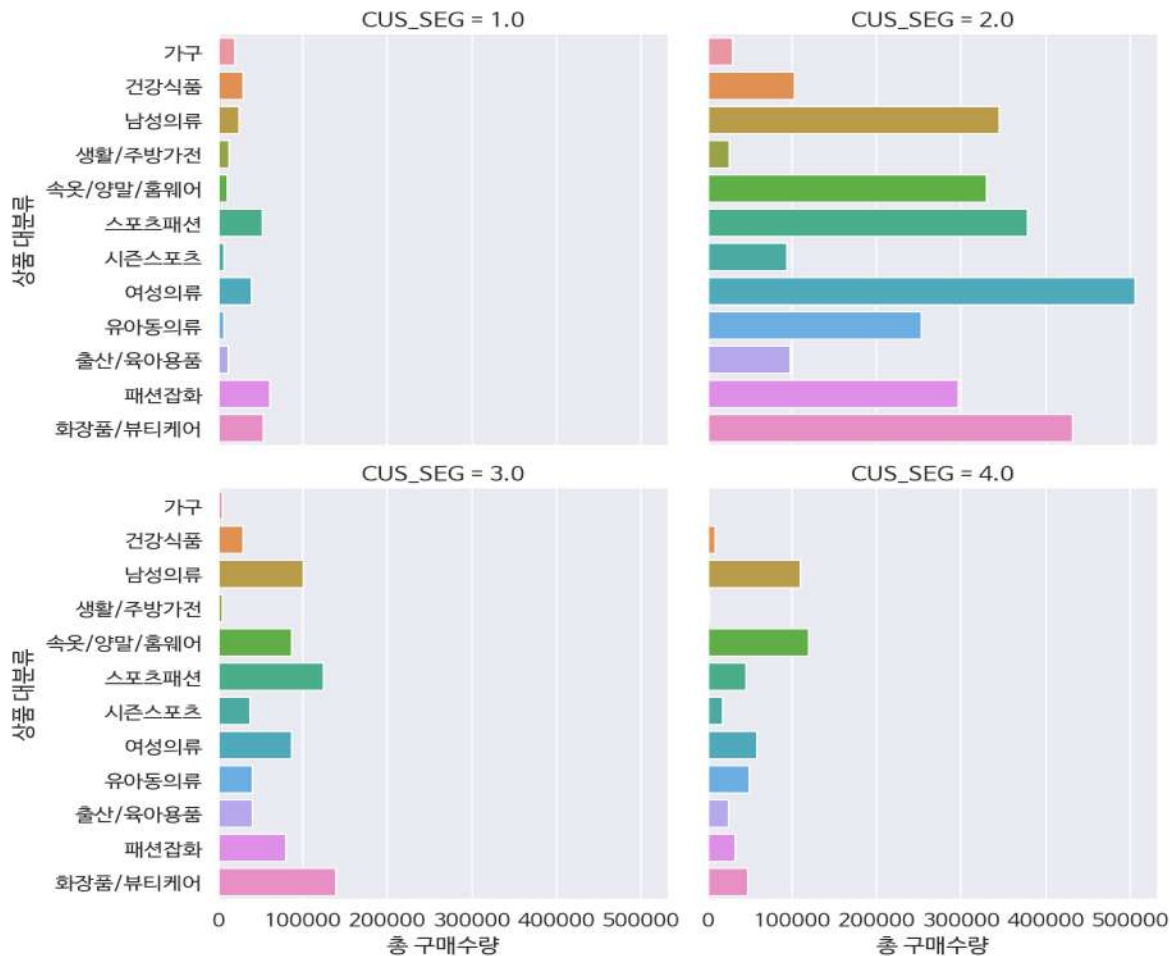


데이터 시각화 & 인사이트 상품

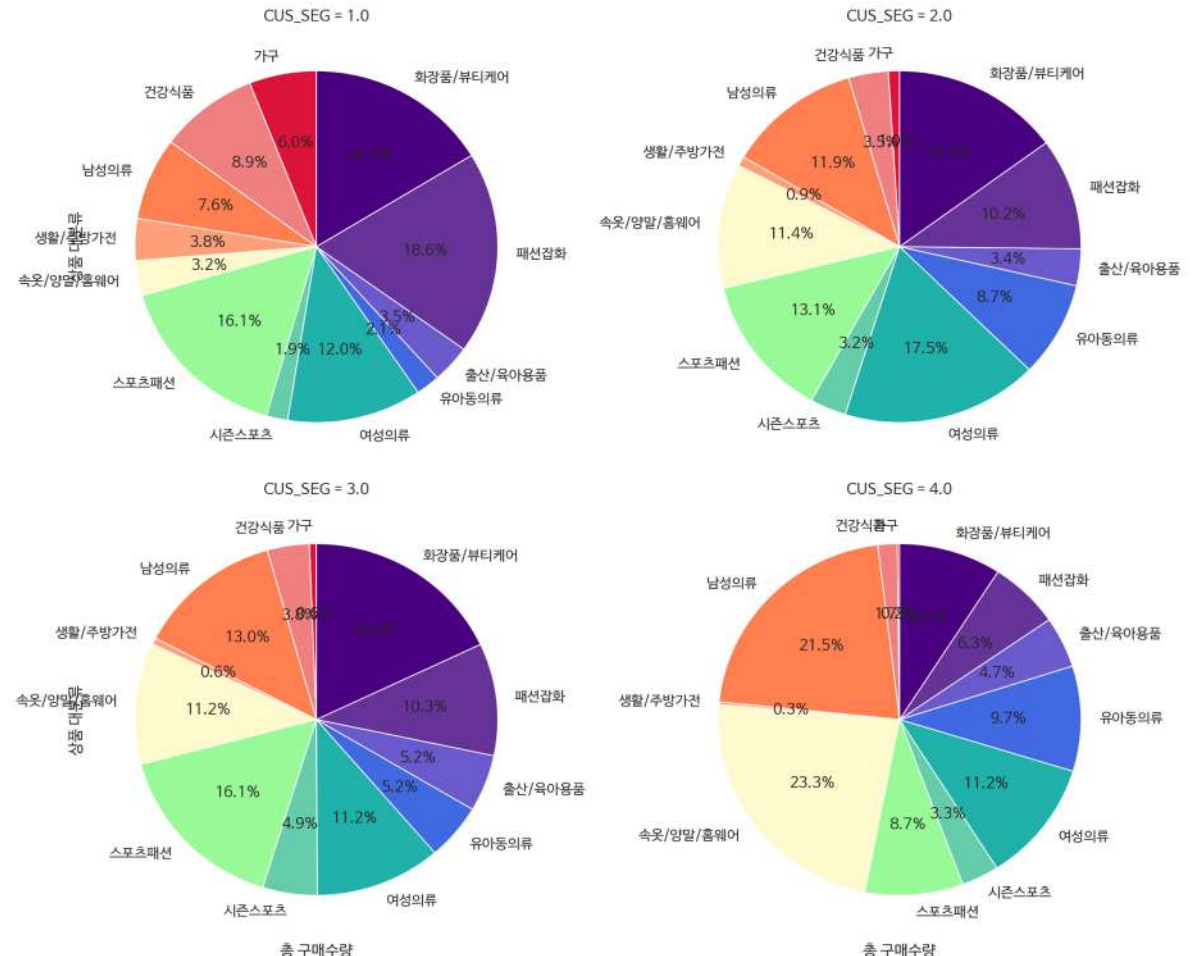
그룹별 구매수량은 어떻게 다를까?

그룹별 많은 구매수량의 상품군 파악

<상품 분류별 구매수량 (개)>



<상품 분류별 구매수량 비율(%)>



Group 1의 주력 상품군

<div>기준</div> <div>등수</div>	구매고객수	구매횟수	구매금액	구매수량
1위	스포츠패션	패션잡화	패션잡화	패션잡화
2위	패션잡화	스포츠패션	가구	화장품/뷰티케어
3위	여성의류	화장품/뷰티케어	스포츠패션	스포츠패션

Group 2의 주력 상품군

<div>기준</div> <div>등수</div>	구매고객수	구매횟수	구매금액	구매수량
1위	남성의류	여성의류	패션잡화	여성의류
2위	스포츠패션	화장품/뷰티케어	여성의류	화장품/뷰티케어
3위	건강식품	스포츠패션	화장품/뷰티케어	스포츠패션

Group 3의 주력 상품군

<div>기준</div> <div>등수</div>	구매고객수	구매횟수	구매금액	구매수량
1위	스포츠패션	화장품/뷰티케어	스포츠패션	화장품/뷰티케어
2위	화장품/뷰티케어	스포츠패션	화장품/뷰티케어	스포츠패션
3위	남성의류	남성의류	패션잡화	남성의류

Group 4의 주력 상품군

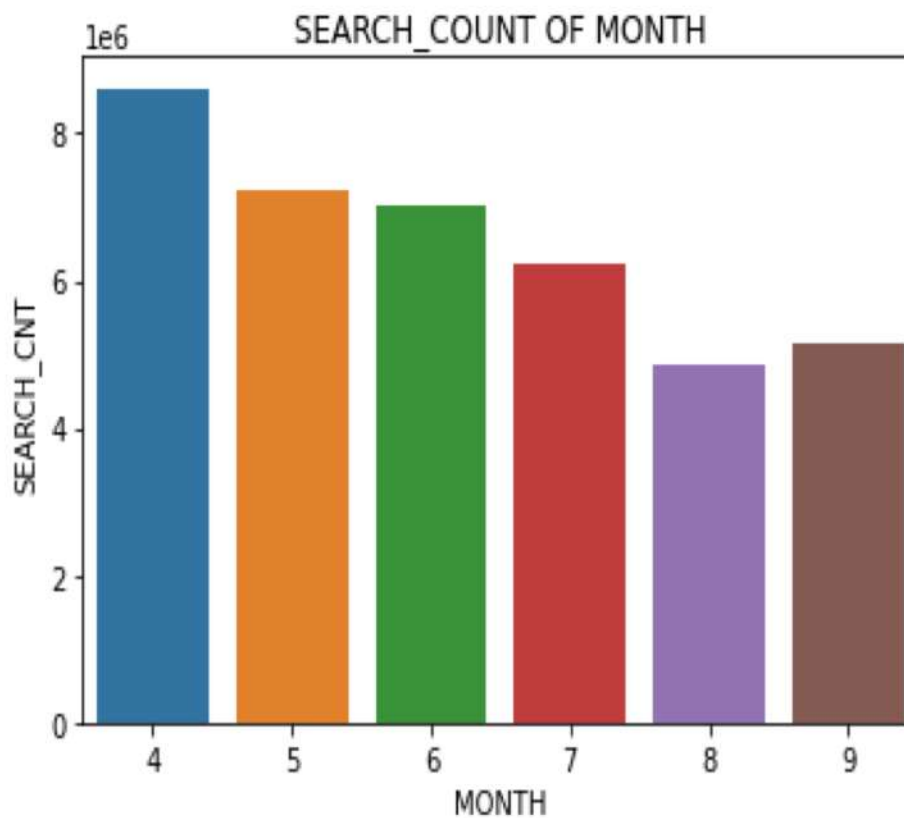
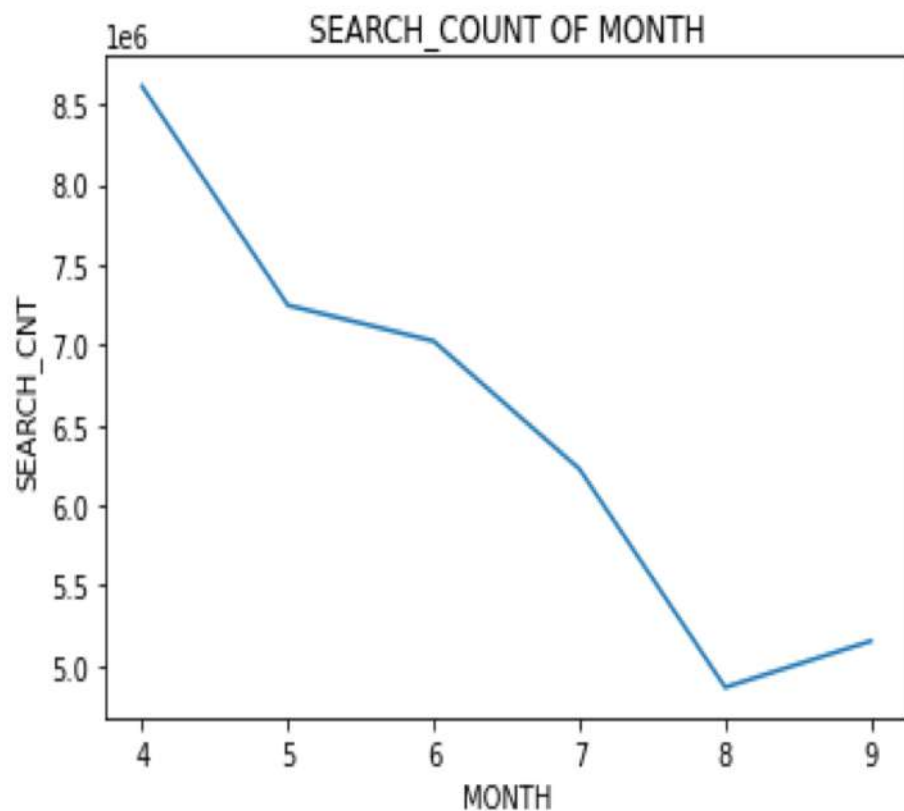
<div>기준</div> <div>등수</div>	구매고객수	구매횟수	구매금액	구매수량
1위	남성의류	남성의류	남성의류	속옷/양말/홈웨어
2위	속옷/양말/홈웨어	속옷/양말/홈웨어	여성의류	남성의류
3위	스포츠패션	여성의류	스포츠패션	여성의류

데이터 시각화 & 인사이트 검색

월별 총 검색량은 어떨까?

시간이 지날수록 검색량 감소
시기와 사용자 감소 측면에서 예측

<월별 총 검색량 (번)>

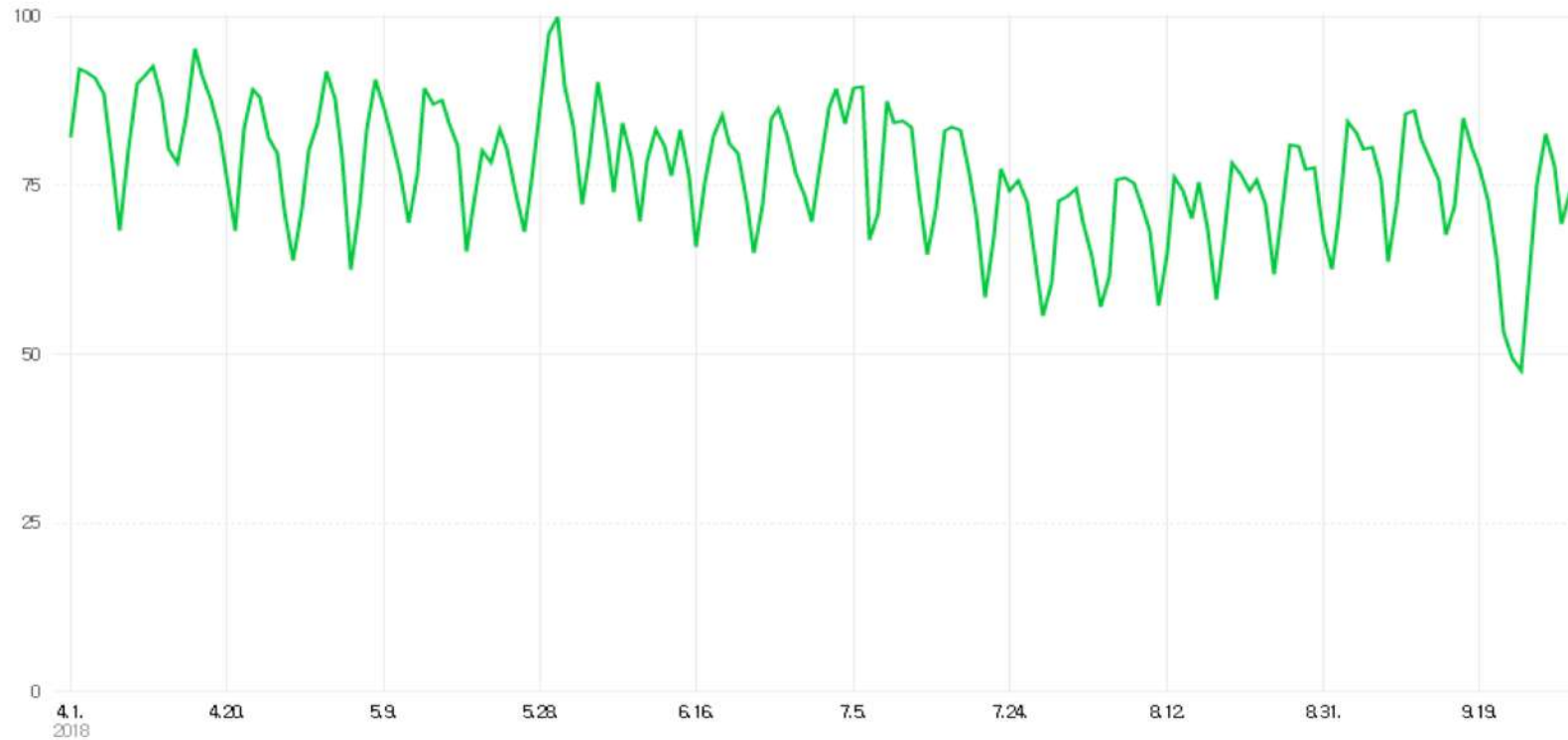


데이터 시각화 & 인사이트 검색

다른 플랫폼의 시기별 검색량은?

시기와 사용자 감소 측면에서 예측

< 네이버 쇼핑 관련 검색 데이터 >



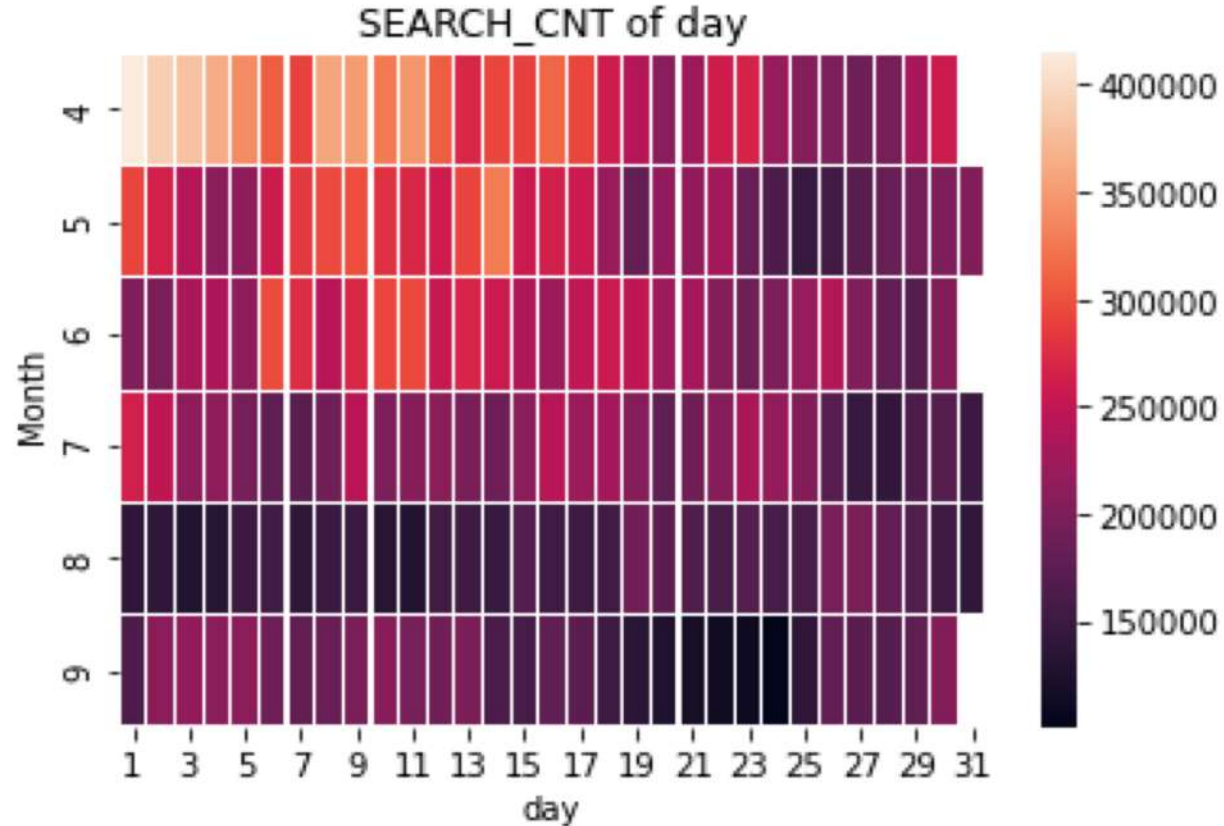
● 쇼핑 | 가구,건강식품,남성의류,냉장,세탁가전,문구,사무용품,생활,주방가전,속옷,홈웨어,스포츠패션,식기,조리기구,여성의류,유아동의류,출산,육아용품,패션잡화,화장품

데이터 시각화 & 인사이트 검색

검색량이 많은 시기는?

월 초의 검색량이 높은 경향

<일별 검색량 (번)>



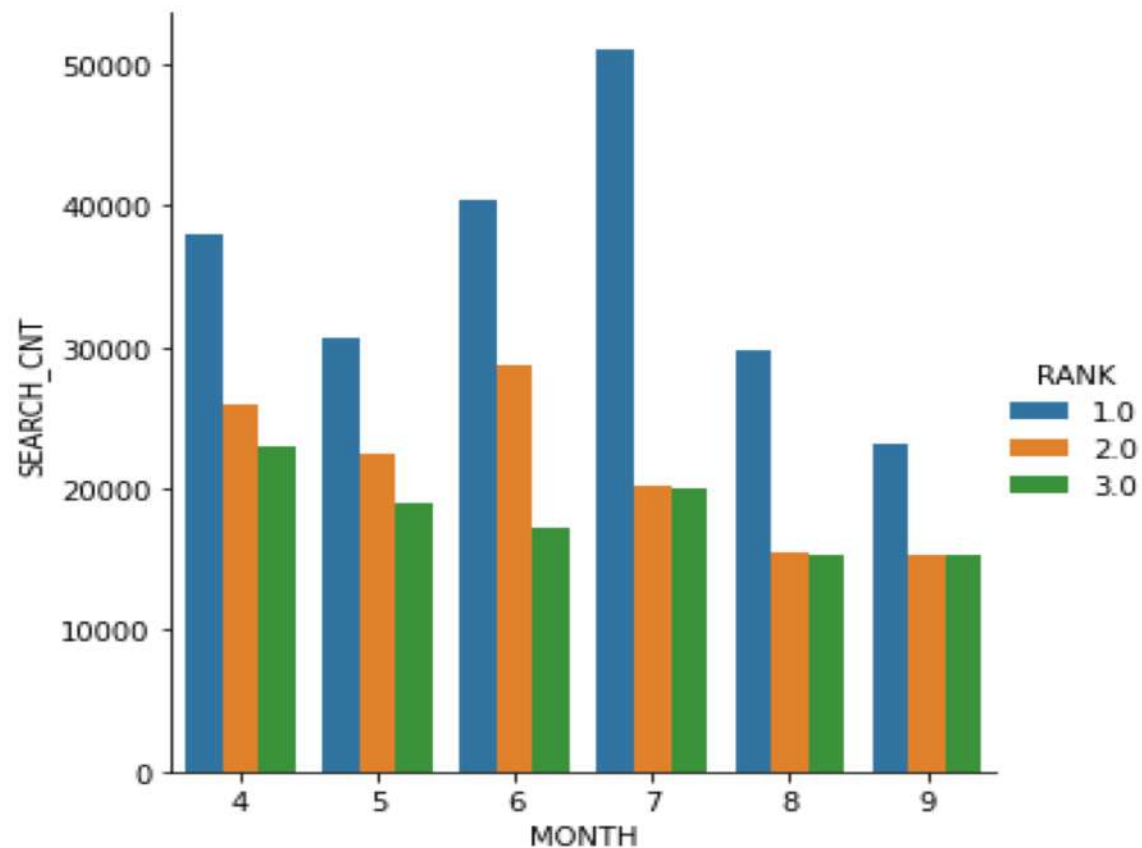
데이터 시각화 & 인사이트 검색

월별 Top3 검색어는?

꾸준한 1위 원피스

	SEARCH_CNT	NAME	RANK	MONTH
KWD_NM				
원피스	37950	원피스	1.0	4
블라우스	25926	블라우스	2.0	4
써스데이아일랜드	23015	써스데이아일랜드	3.0	4
원피스	30561	원피스	1.0	5
에이치커넥트	22524	에이치커넥트	2.0	5
블라우스	19053	블라우스	3.0	5
원피스	40391	원피스	1.0	6
여성샌들	28643	여성샌들	2.0	6
에이치커넥트	17265	에이치커넥트	3.0	6
원피스	51038	원피스	1.0	7
여성샌들	20240	여성샌들	2.0	7
양산	19967	양산	3.0	7
원피스	29739	원피스	1.0	8
에고이스트	15430	에고이스트	2.0	8
블라우스	15270	블라우스	3.0	8
원피스	23227	원피스	1.0	9
에고이스트	15379	에고이스트	2.0	9
헤지스레이디스	15280	헤지스레이디스	3.0	9

<1, 2, 3위 검색어의 검색량 (번)>



데이터 시각화 & 인사이트 검색

구매 고객들이 가장 많이 검색한 단어는 무엇일까?

검색량이 높은 단어들의 특징 도출

< 롯데 고객들의 검색 키워드 워드 클라우드 >



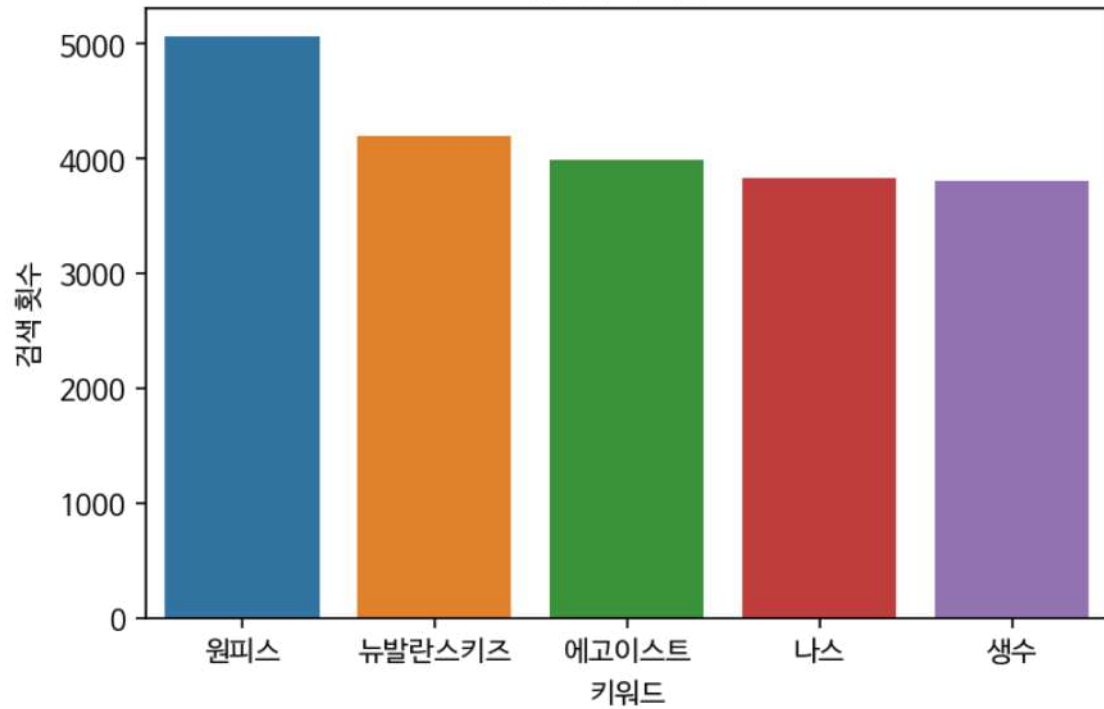
데이터 시각화 & 인사이트 검색

여성, 남성의 검색 키워드는 어떻게 다를까?

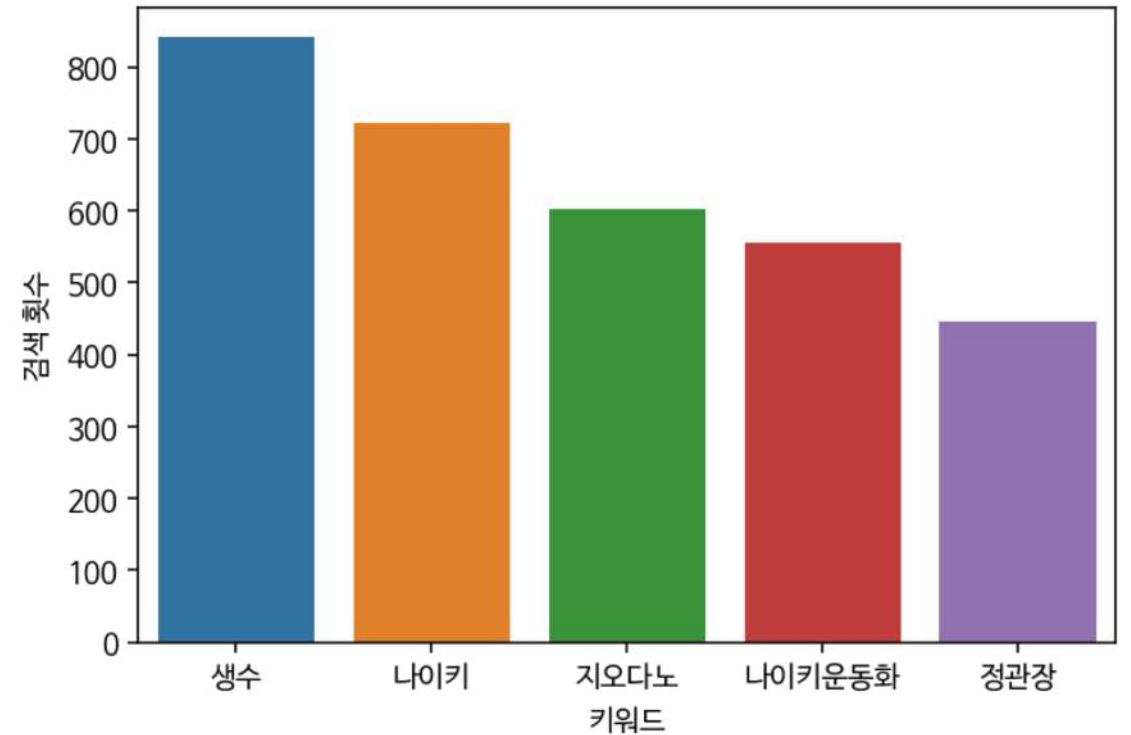
성별에 따라 검색 키워드가 확실히 다름

< 성별 검색 키워드 막대 그래프(회) >

여성 검색 키워드



남성 검색 키워드

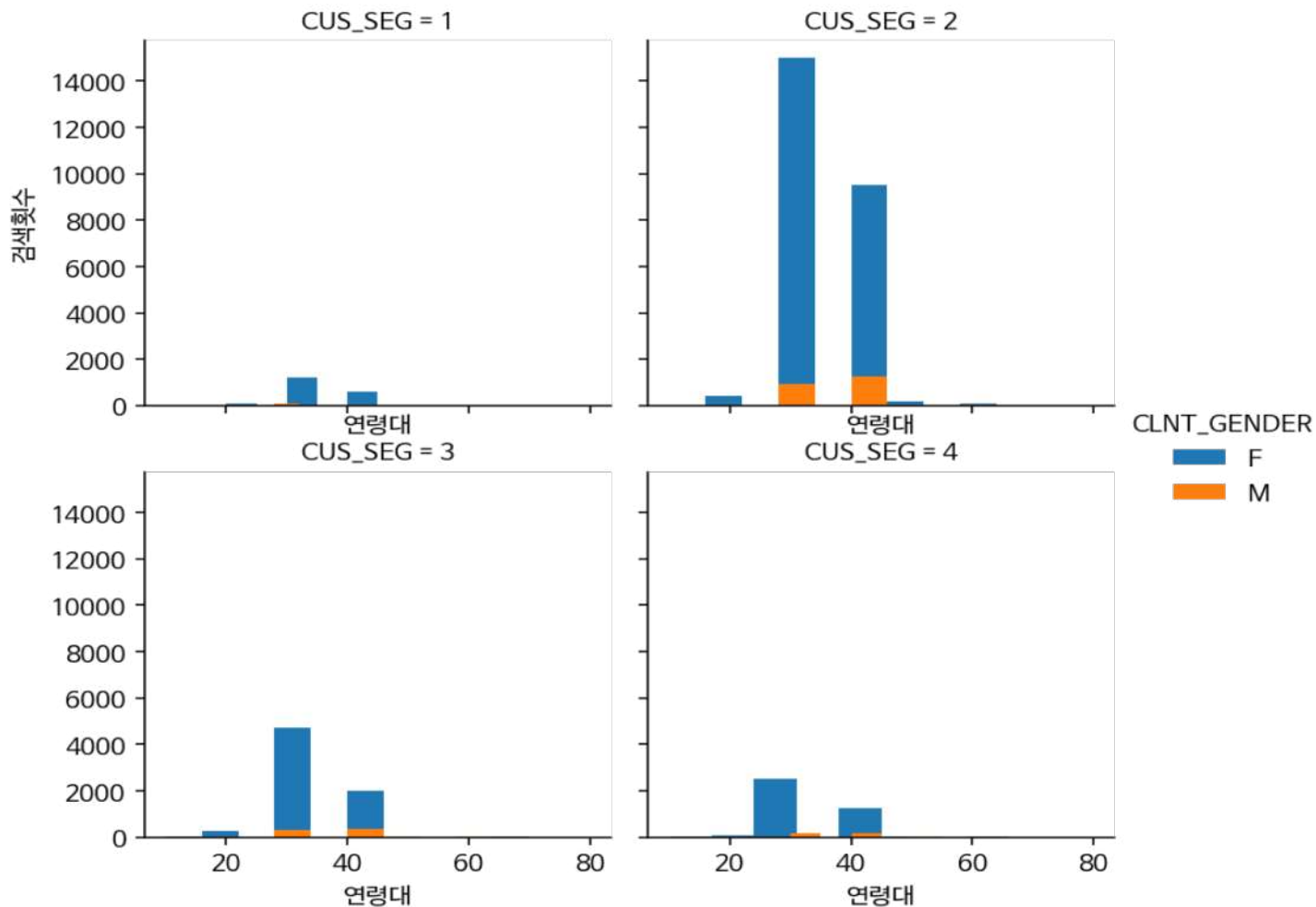


데이터 시각화 & 인사이트 검색

‘키즈’ 키워드 검색량은 그룹별에 따라 어떻게 다를까?

‘키즈’ 검색은 여성이 독보적

< 그룹별 성별 검색 횟수 막대 그래프 (회) >

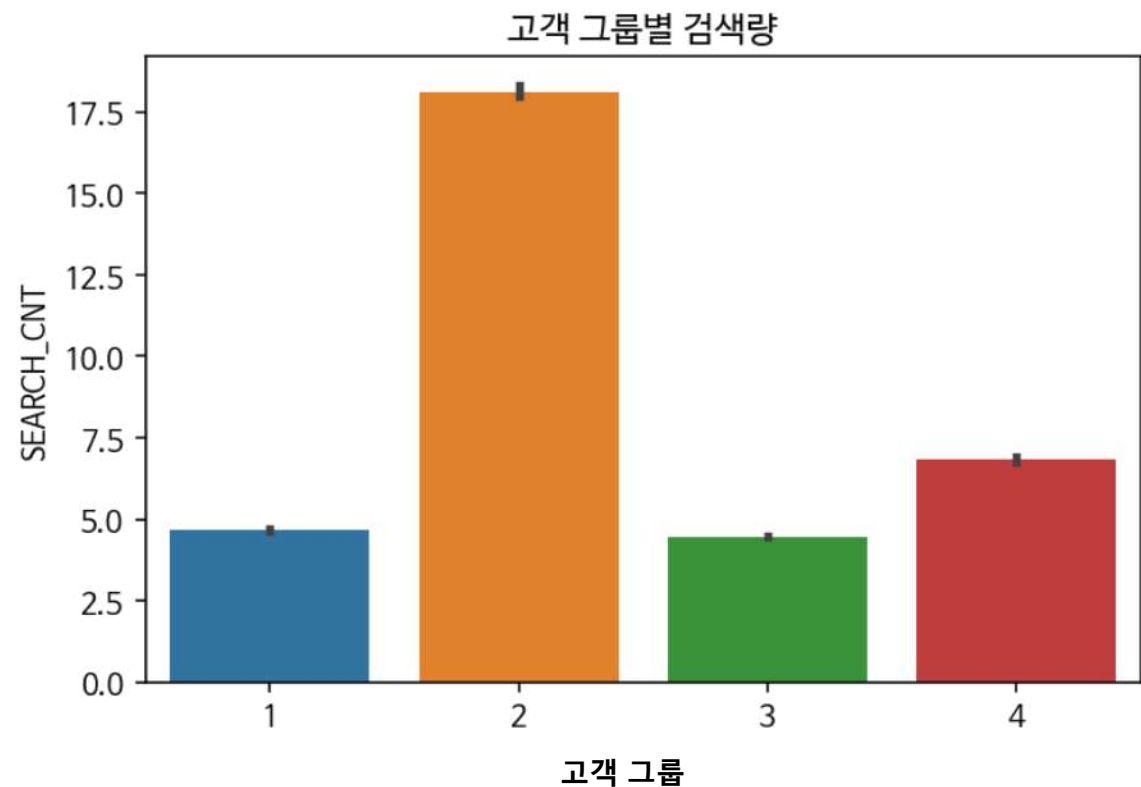
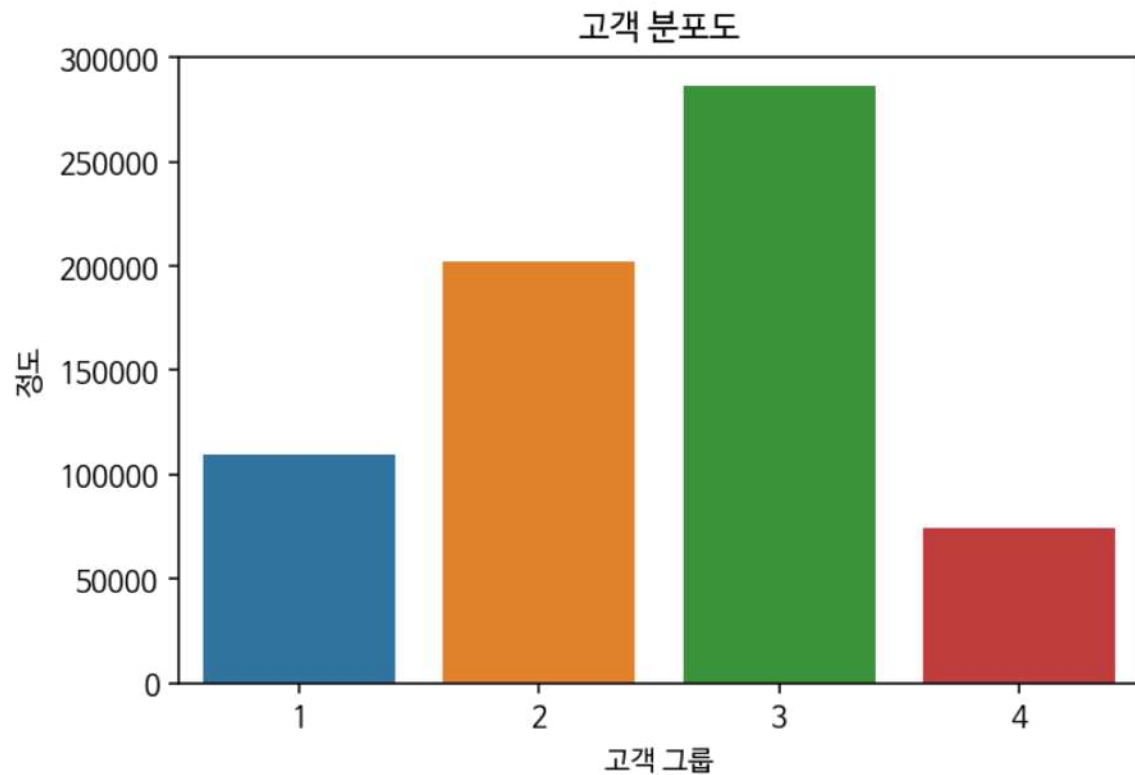


데이터 시각화 & 인사이트 검색

그룹별 고객수와 검색량은 어떻게 다를까?

그룹 고객수와 검색량은 비례하지 않음

< 그룹별 고객 분포, 검색량 막대 그래프 (명, 회) >

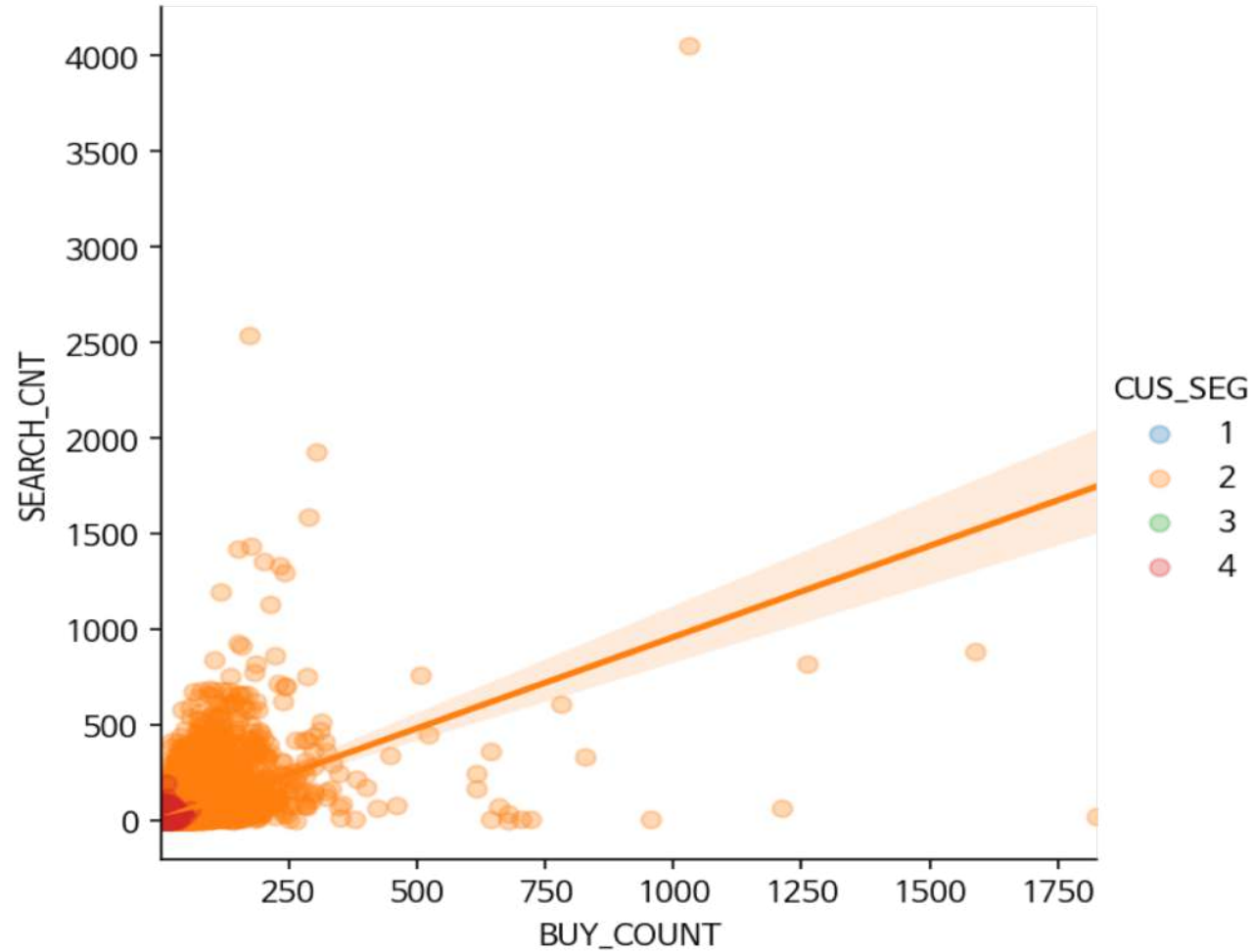


데이터 시각화 & 인사이트 검색

그룹별 구매횟수와 검색횟수는 어떻게 다를까?

회귀선으로 선형적 관계 확인

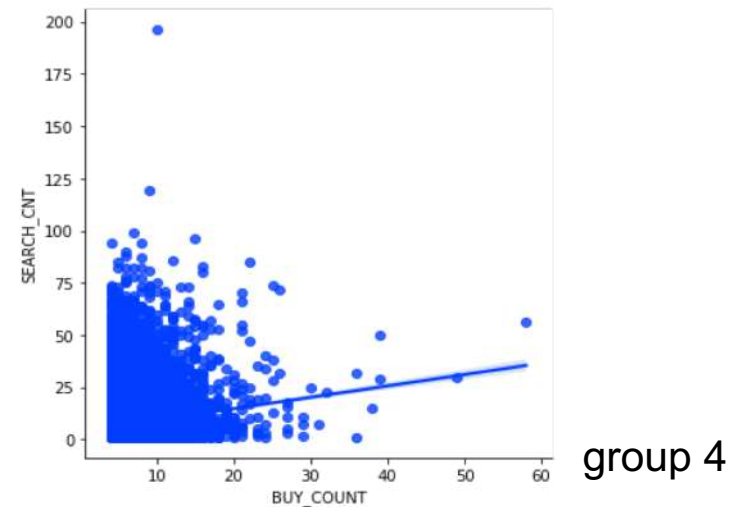
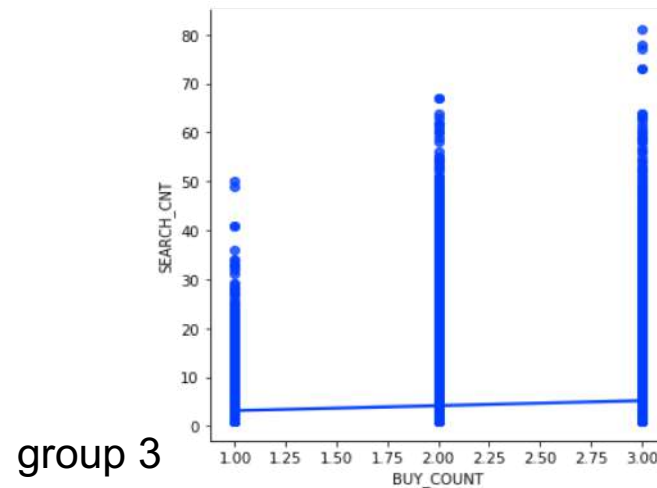
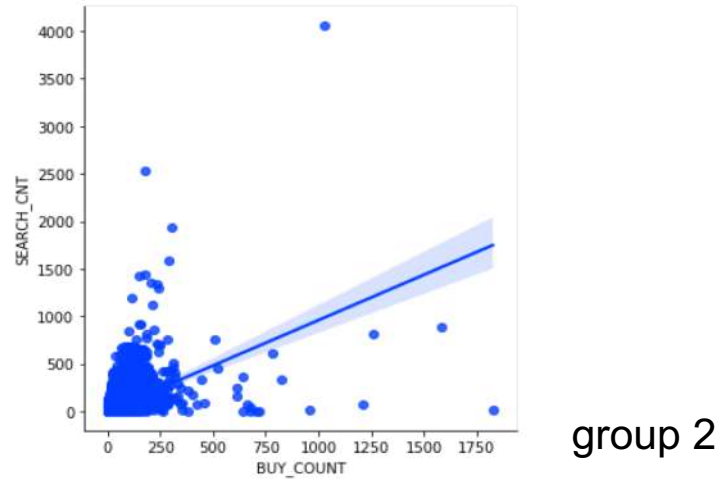
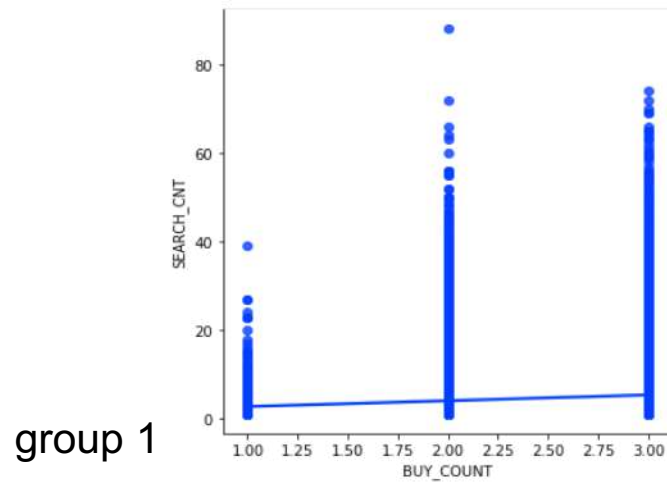
< 그룹별 구매 횟수와 검색 횟수 산점도 (회) >



데이터 시각화 & 인사이트 검색

그룹별 구매횟수와 검색횟수는 어떻게 다를까?

< 그룹별 구매 횟수와 검색 횟수 산점도 (회) >



5. 분석 결과

그룹별 특징 및 정리





L.pay | L.POINT

롯데의 고객은 누구일까?

3040 여성이 전체 고객의 약 68% 차지

클릭 수가 많을수록 구매량은 감소

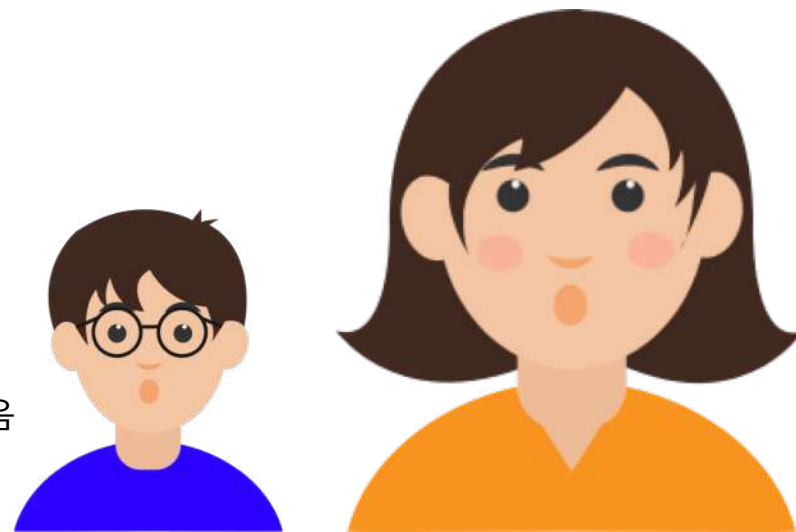
준내구재를 가장 **많이** 구매 *1년 이상 사용 가능/저가 제품

내구재를 가장 **적게** 구매 *1년 이상 사용 가능/고가 제품

패션잡화, 여성의류, 스포츠 패션 상품은 스테디 셀러

남성의류 구매 고객은 가성비를 중시 *구매 고객은 많은데, 총 판매액은 적음

화장품/뷰티 케어 카테고리는 상품 추가 필요 *상품 수는 적는데, 구매고객/판매액은 많음



“프로 아이쇼핑러” 35세 지훈씨

Group 1 구매 횟수는 적지만, 구매 금액이 높은 그룹

30대 남성의 비율이 높음

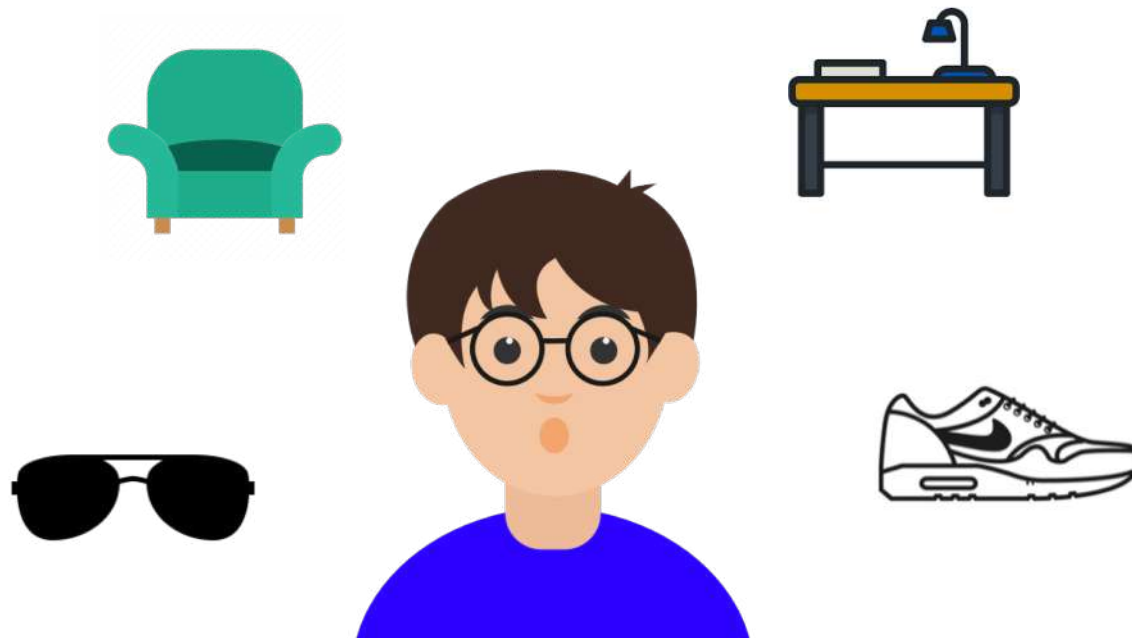
적극적이지 않은 온라인 행동

*방문 횟수, 클릭 수, 상품 탐색 시간 적음

내구재 구매 비율 높음 → 가구의 주 고객

최다 주문 금액이 4 그룹 중 가장 높음 *79,000원

패션잡화, 스포츠 패션 구매 많음



김지훈 (35)

“나를 위한 소비, 가족을 위한 소비” 33세 은정씨

Group2 구매 횟수도 높고, 구매 금액도 높은 그룹

30, 40대 여성

적극적인 온라인 행동 *모바일 사용, 방문 횟수, 클릭 수, 상품 탐색 시간 가장 높음

준내구재 구매 비중이 가장 높음

최다 주문 금액이 4 그룹 중 중간 *39,000원

가격대가 높은 **여성의류, 화장품/뷰티 케어** 제품을 주로 구매

남성 의류를 구매하는 고객이 여성 의류를 구매하는 고객 수의 4배

But 구매한 남성 의류의 가격대는 낮게 형성되어 있음

유,아동 의류를 자주, 많이 구매하고 구매량 또한 높음



박은정 (38)

“힙스터 워너비” 25세 수빈씨

Group3 구매 횟수도 적고 구매 금액도 적은 그룹

고객수가 가장 많은 그룹

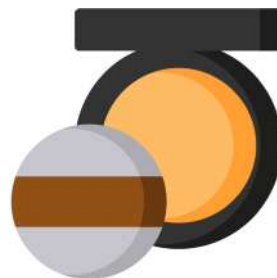
20대 여성의 비율이 높음

재구매율이 가장 낮음

가장 소극적인 온라인 행동

화장품/뷰티 케어, 스포츠 패션, 남성의류 주로 구매

최다 주문 금액이 39,000원으로 형성

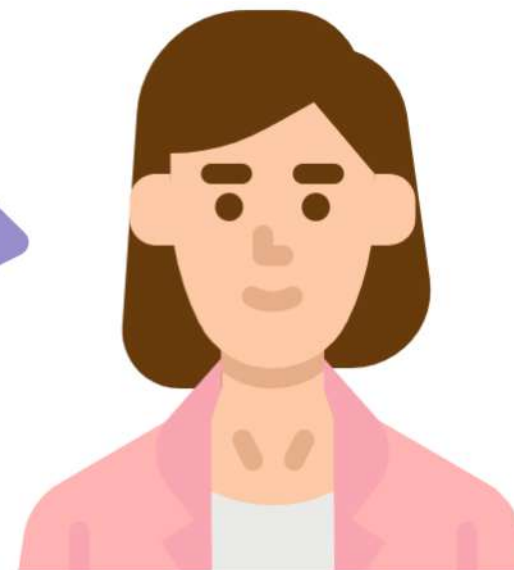


이수빈 (25)

“알뜰 살뜰 가성비왕” 36세 미영씨

Group 4 구매 횟수는 높지만 구매 금액이 낮은 그룹

고객수가 제일 적고, 30대 여성이 주 고객
상품 탐색 시간이 긴 편
준내구재 구매 비율 1위
남성의류, 속옷/양말/홈웨어 가장 많이 구매
최다 주문 금액 9,900원



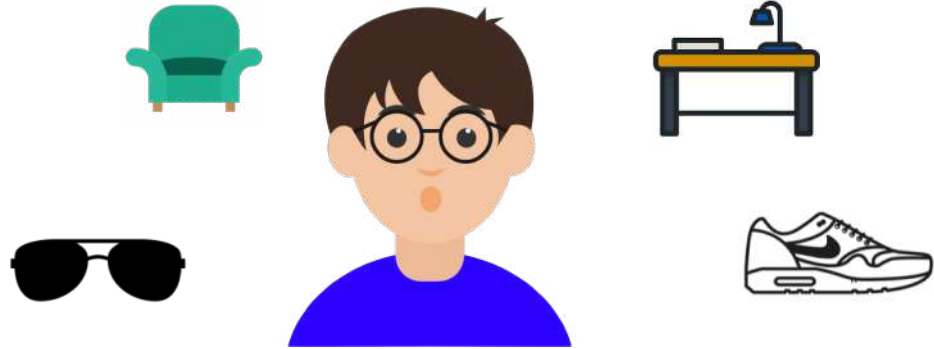
정미영 (36)

6. 결론

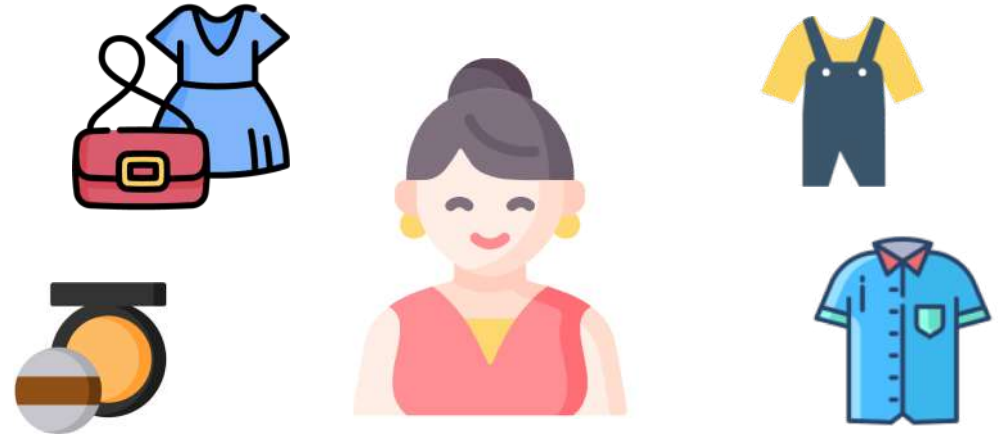
페르소나 선정



“프로 아이쇼핑러” 35세 지훈씨



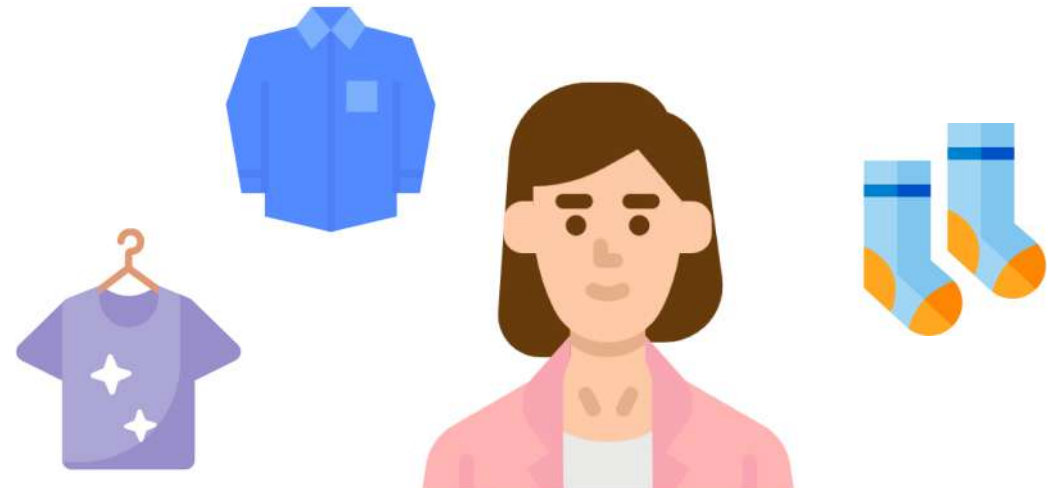
“나를 위한 소비, 가족을 위한 소비” 38세 은정씨



“힙스터 워너비” 25세 수빈씨



“알뜰살뜰 가성비 왕” 36세 미영씨



project

Thanks !

Team scalar

