

롯데 온라인 쇼핑몰 소비패턴 분석을 통한 개인화 서비스 제공

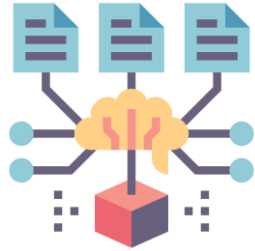
올림푸스 | 김수빈 김아름 박대한
안영은 유찬희 이동일 이용빈



CONTENTS



1. 분석 배경



2. 분석 데이터 탐색



3. 문제 원인 파악



4. 고급 분석 및 추천전략



5. 결론

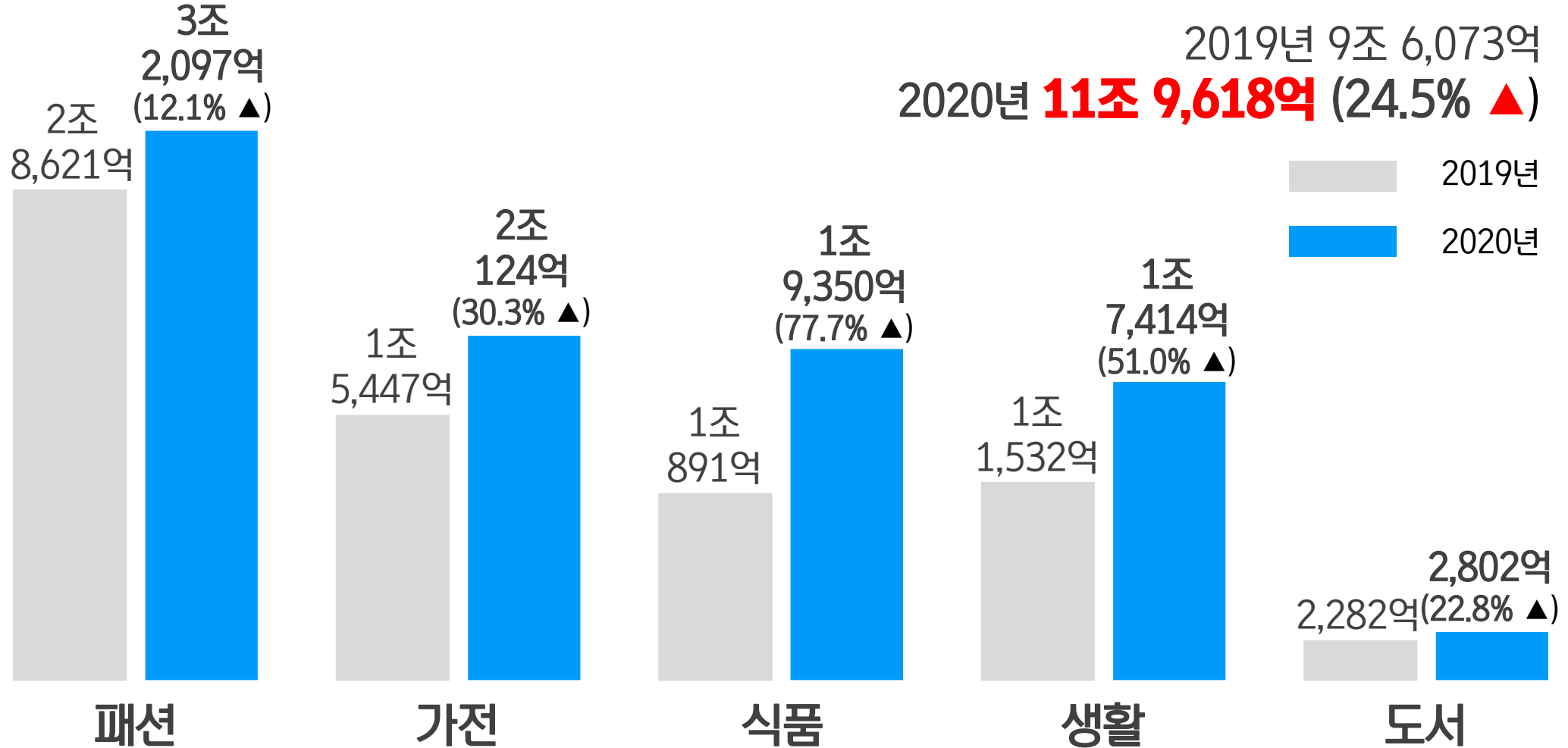


01 분석 배경

온라인 쇼핑 동향
롯데 구매 추이 및 전망 예측
분석 목표 설정

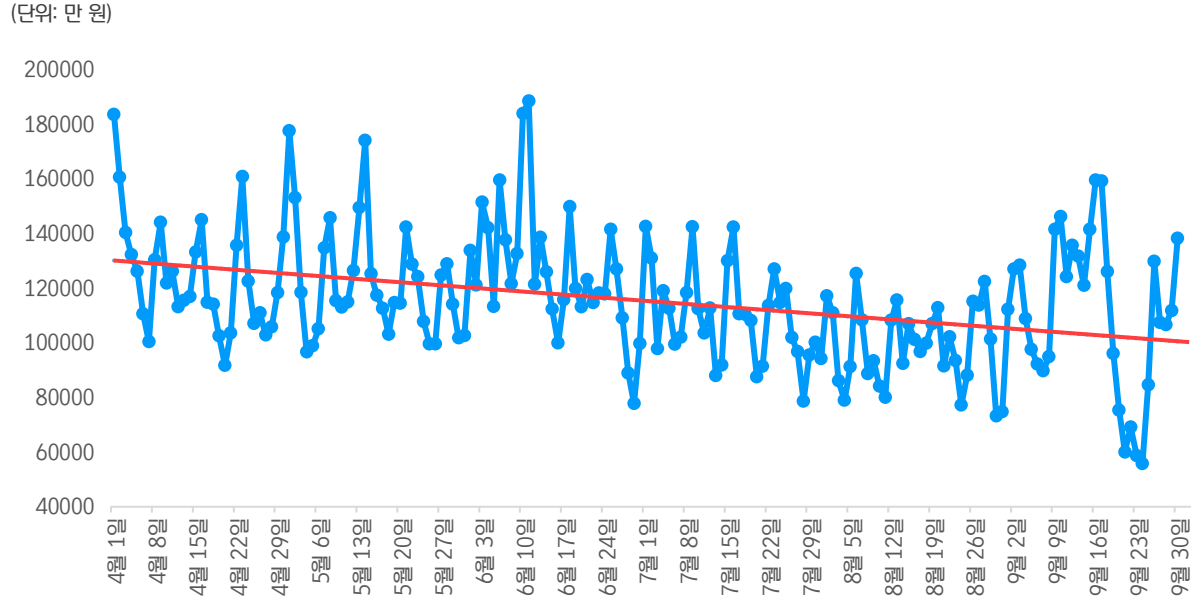
온라인 쇼핑 동향

온라인 쇼핑 시장 규모 (단위: 원)

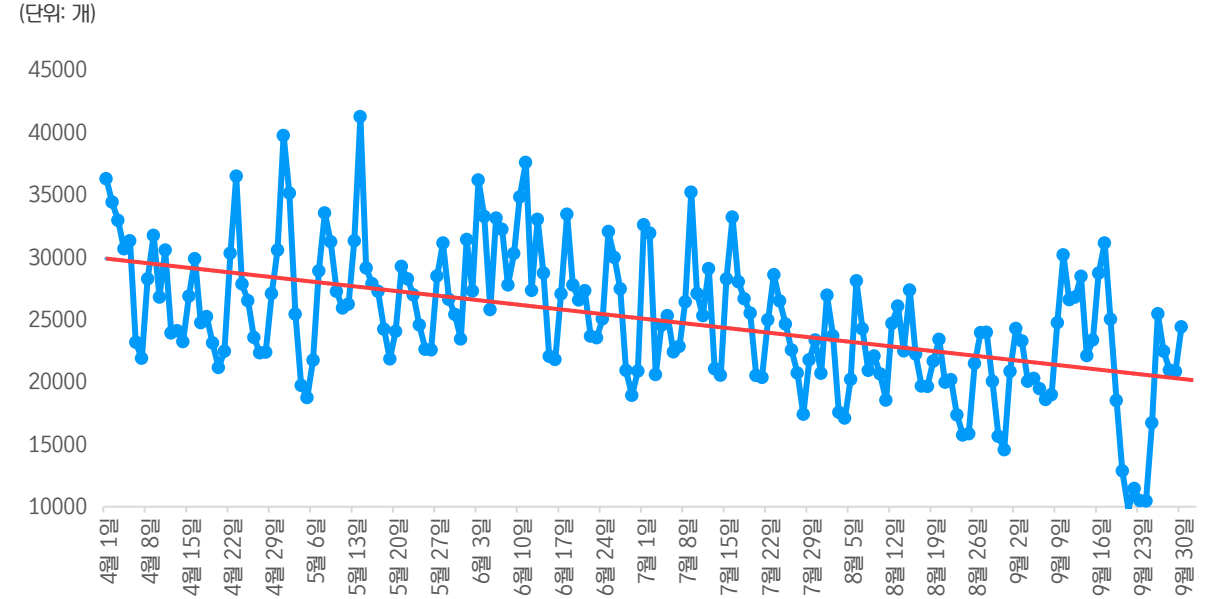


롯데 온라인 쇼핑몰 구매 추이 및 전망 예측

일자별 총 구매액



일자별 총 구매량

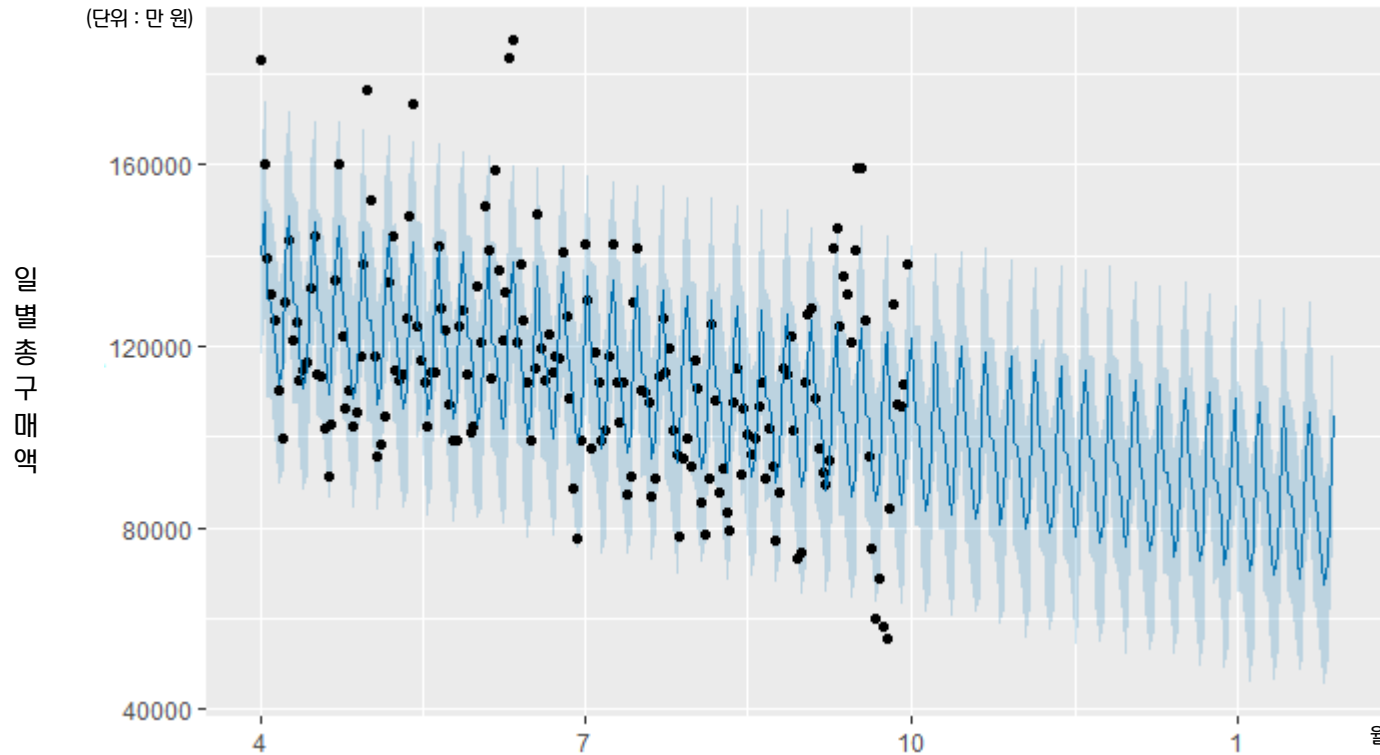


온라인 쇼핑 시장의 성장세에도 불구하고

롯데 온라인 쇼핑몰의 구매액 및 구매량은 감소 추세를 보이고 있음

롯데 온라인 쇼핑몰 구매 추이 및 전망 예측

일별 구매액 추이 및 전망 예측



Facebook. Prophet을 이용한 추세선 그래프를 통해
큰 이슈가 없는 한 향후 3개월간의 매출이 하락할 것을 예측함

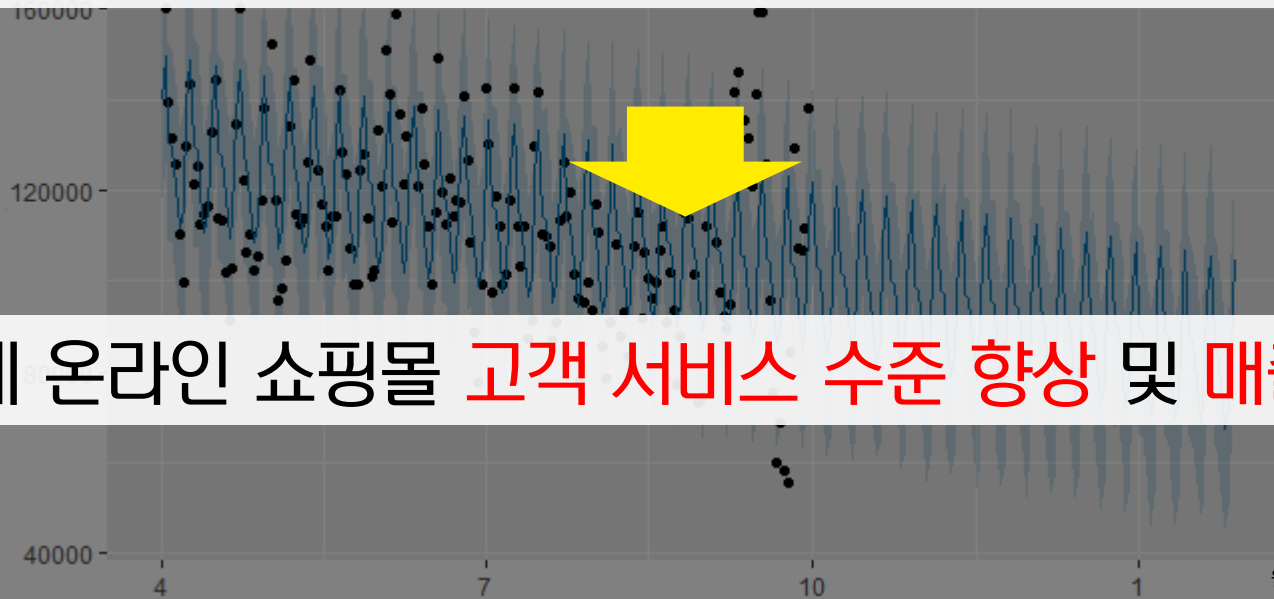
분석 목표 설정

일별 구매액 추이 및 전망 예측

(단위: 만 원)

소비패턴 분석을 통한 **개인화 서비스** 제공

일
별
총
구
매
액



“롯데 온라인 쇼핑몰 **고객 서비스 수준 향상 및 매출 향상**”

Facebook. Prophet을 이용한 추세선 그래프를 통해
큰 이슈가 없는 한 향후 3개월간의 매출이 하락할 것을 예측함



02

분석 데이터 탐색

분석 툴 및 데이터 소개

데이터 전처리

기초 데이터 탐색

분석 툴 및 데이터 소개

OS



DB



ORACLE®

언어



분석 틀 및 데이터 소개

구분		내용
인간	구매(Order)	구매된 상품명, 상품금액/건수 등 방문자의 상품구매 정보
	검색어(Search)	상품을 구매한 방문자(Visitors)가 검색창에 입력한 검색어에 대한 검색량
행동	일자별 검색어 (Search Date)	분석데이터 기간 내 일별/검색어별 검색량
데이터	고객(Client)	방문자(Visitors)의 성별/연령 정보
	세션(Session)	상품구매한 방문자(Visitors)의 세션 접속일자, 총 세션시간, 세션기준 발생 지역, 기기유형 등 세션 정보
메타	상품(Master)	상품별 상품명, 대/중/소 분류명 정보

데이터 전처리

① 기초 데이터 전처리

Data Frame	변수명	Cleansing Rule
구매 (Order)	CLNT_ID	고객(Client) 테이블에 고객정보 없음
	SESS_ID	세션(Session) 테이블에 세션정보 없음
	HITS_SEQ	HITS_SEQ = 1
	PD_BRA_NM	불규칙한 브랜드명 표기 통일
	PD_BUY_AM	PD_BUY_AM = 0
	PD_BUY_CT	PD_BUY_CT = 0
검색어 (Search)	CLNT_ID	고객(Client) 테이블에 고객정보 없음
	SESS_ID	세션(Session) 테이블에 세션정보 없음
세션 (Session)	SESS_SEQ	SESS_SEQ = 0
	TOT_SESS_HR_V	TOT_SESS_HR_V IS NULL
	TOT_PAG_VIEW_CT	TOT_PAG_VIEW_CT = 0

데이터 전처리

② 제품 대분류 통합 및 파생변수 생성

새로운 대분류	기존 대분류	구매 횟수	구매 비율
전자제품 (C1)	계절가전, 냉장/세탁가전, 모바일, 생활/주방가전, 영상/음향가전, 컴퓨터	C1_CT	전자제품
리빙 (C2)	세제/위생, 식기/조리기구, 청소/세탁/욕실용품, 가구, 인테리어/조명, 침구/수예, 주방잡화	C2_CT	리빙
식품 (C3)	건강식품, 과일, 냉동식품, 냉장식품, 음료, 축산물	C3_CT	식품
스포츠 (C4)	구기/필드스포츠, 시즌스포츠, 아웃도어/레저, 헬스/피트니스	C4_CT	스포츠
패션 (C5)	남성의류, 속옷/양말/홈웨어, 스포츠패션, 여성의류, 패션잡화	C5_CT	패션
뷰티 (C6)	퍼스널케어, 화장품/뷰티케어	C6_CT	뷰티
유아용품 (C7)	완구, 유아동의류, 출산/육아용품	C7_CT	유아용품
취미용품 (C8)	자동차용품, 원예/애완	C8_CT	취미용품
사무용품 (C9)	문구/사무용품	C9_CT	사무용품
상품권 (C10)	상품권	C10_CT	상품권

데이터 전처리

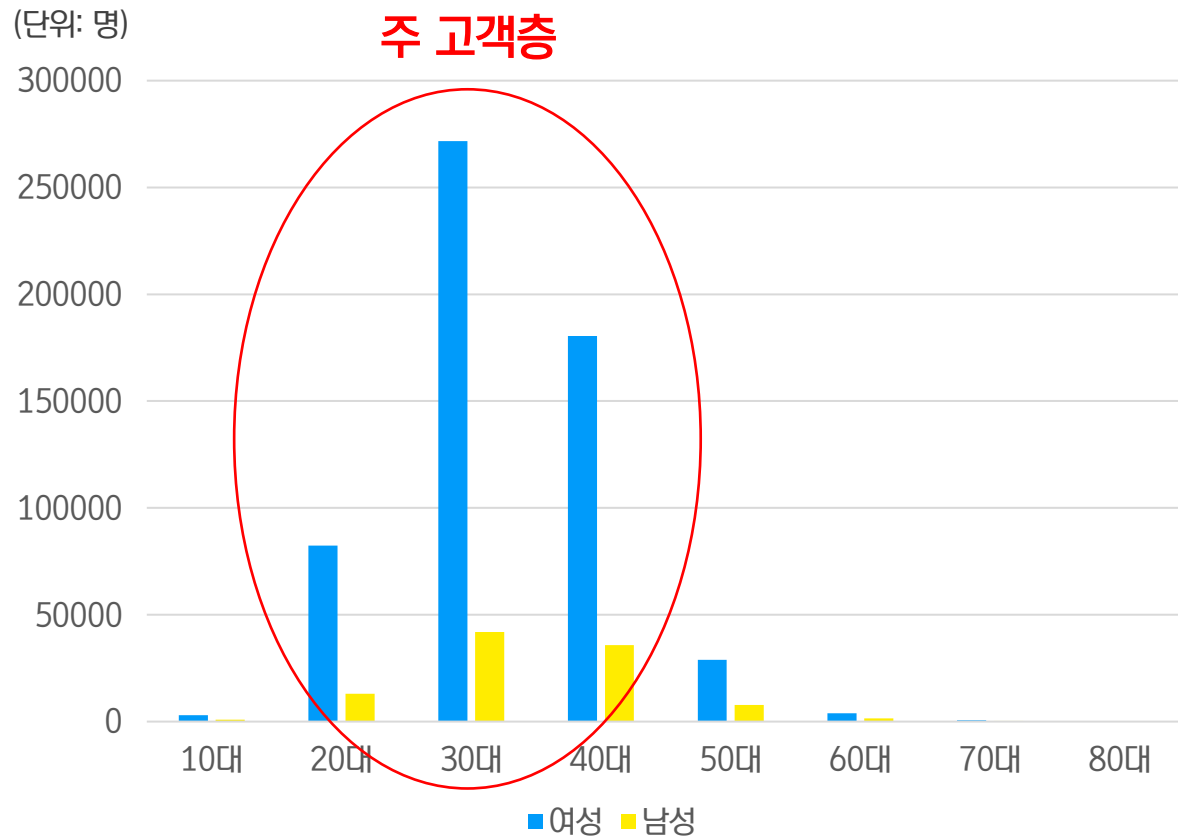
③ 파생변수 생성

Data Frame	변수명	한글명	내용
고객 (Client)	HIT_AVG	세션 평균 히트수	CLNT_ID(클라이언트ID)당 히트 수 평균
	PAG_VIEW_AVG	세션 평균 페이지뷰	CLNT_ID(클라이언트ID)당 페이지 뷰 평균
	SESS_CT	세션 수	CLNT_ID(클라이언트ID)당 생성 세션 수
	CLNT_BUY_CT	총 구매 상품 수	CLNT_ID(클라이언트ID)당 구매 수량
	CLNT_BUY_TOT	총 구매액	CLNT_ID(클라이언트ID)당 구매한 금액
	CLNT_CL	고객 분류 번호	고객 분류 (1: 10, 20대 남 / 2: 10대 여 / 3: 20대 여 / 4: 30대 남 / 5: 30대 여 / 6: 40, 50대 남 / 7: 40, 50대 여 / 8: 60, 70, 80대 남 / 8: 60, 70, 80대 여)
	Cluster_num	클러스터링 결과	패션형, 뷰티형, 생활형
상품 (Master)	CN_our	대분류	대분류 코드(C1~C10)
	CN_Name	대분류명	대분류명
세션 (Session)	DAY_OF_WEEK	구매 요일	구매 일자에 해당하는 요일
주문 (Order)	ORDER_ID	주문 ID	하나의 주문에 대한 고유한 ID
	PD_BUY_TOT	총 금액	상품 단위 당 금액 x 구매 개수 (원)

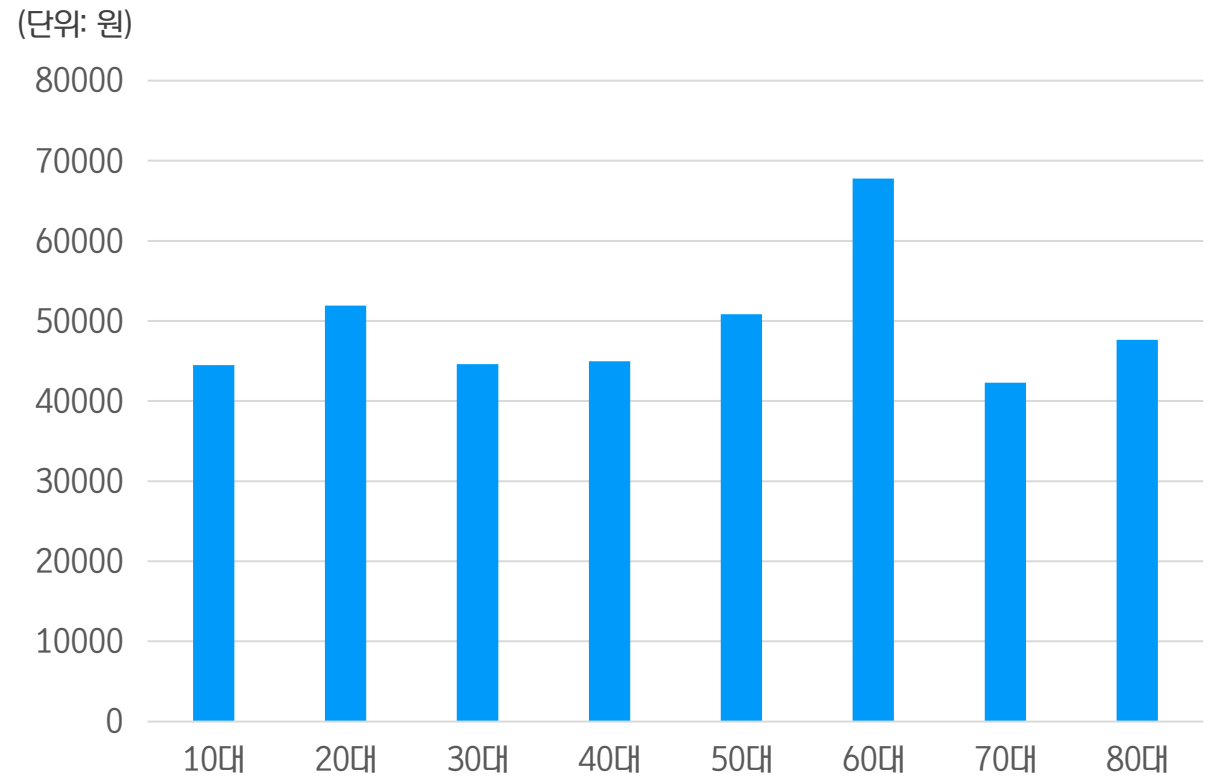
기초 데이터 탐색

① 고객 데이터 탐색

성별/연령별 고객 비율



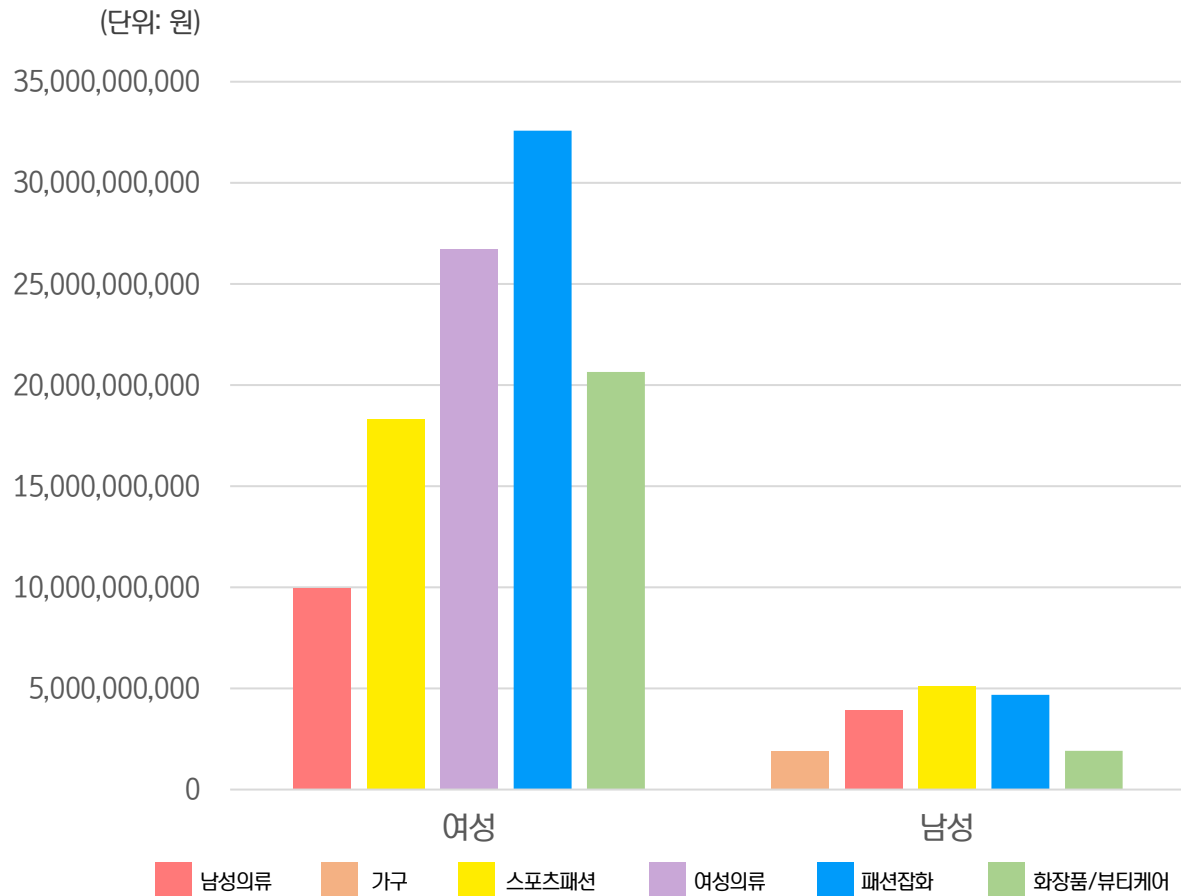
연령별 평균 매출액



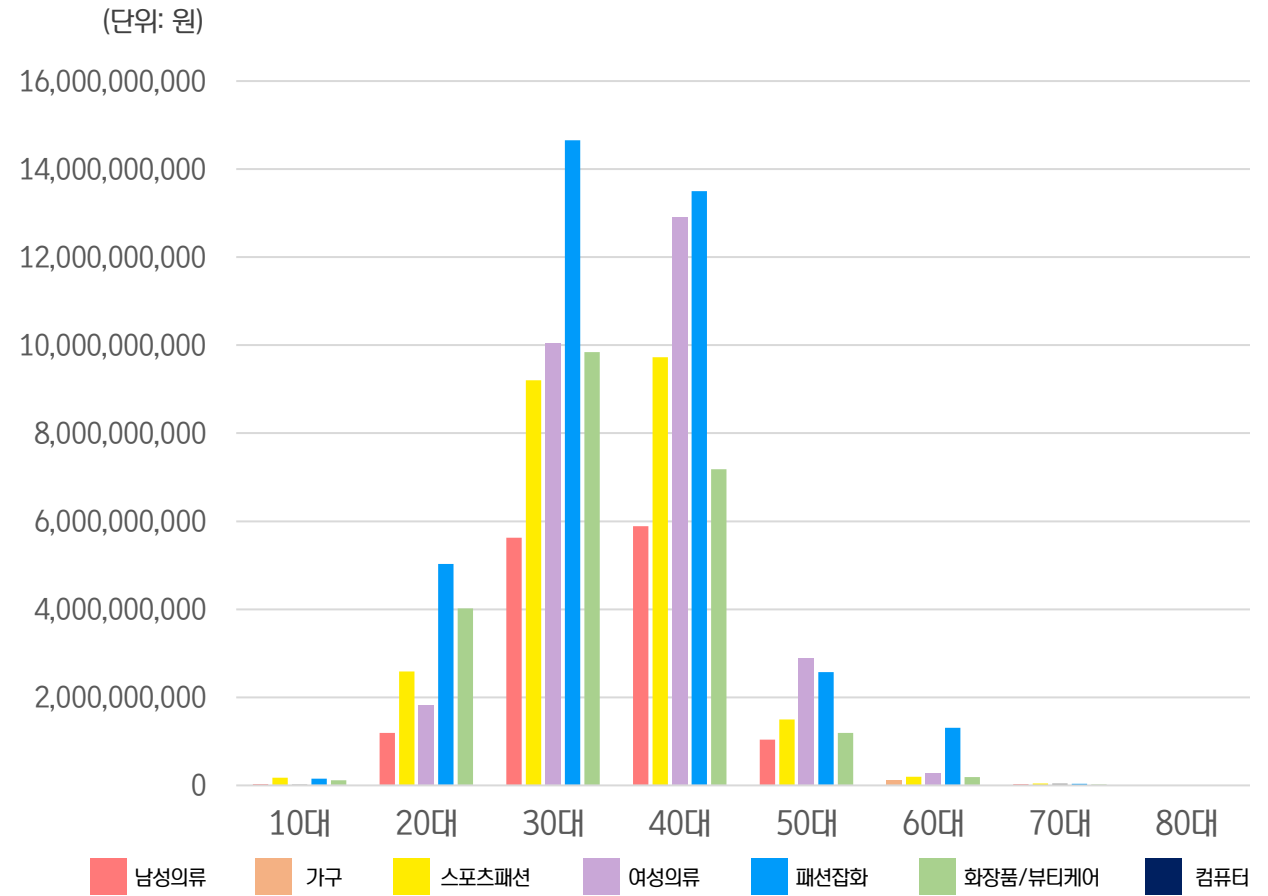
기초 데이터 탐색

② 상품 데이터 탐색

성별 최고 매출 TOP5 제품군



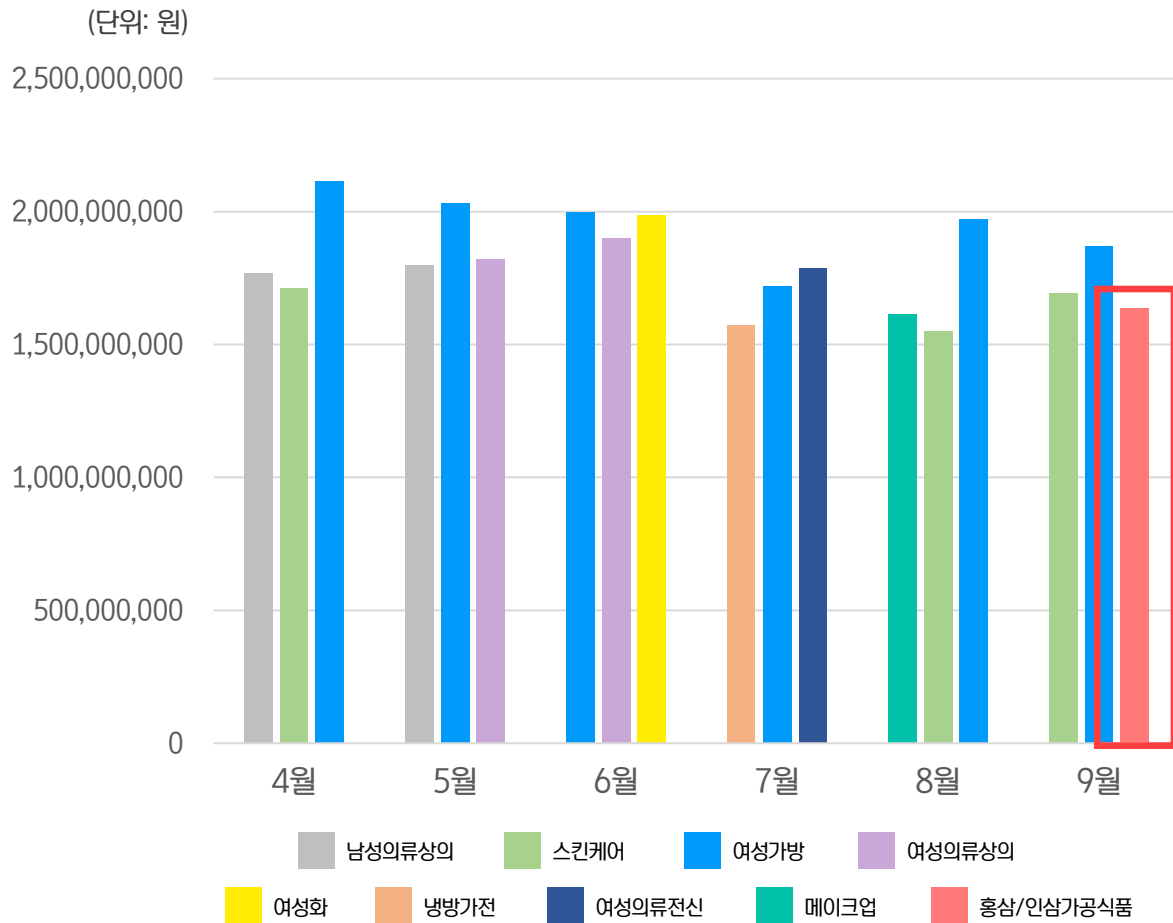
연령별 최고 매출 TOP5 제품군



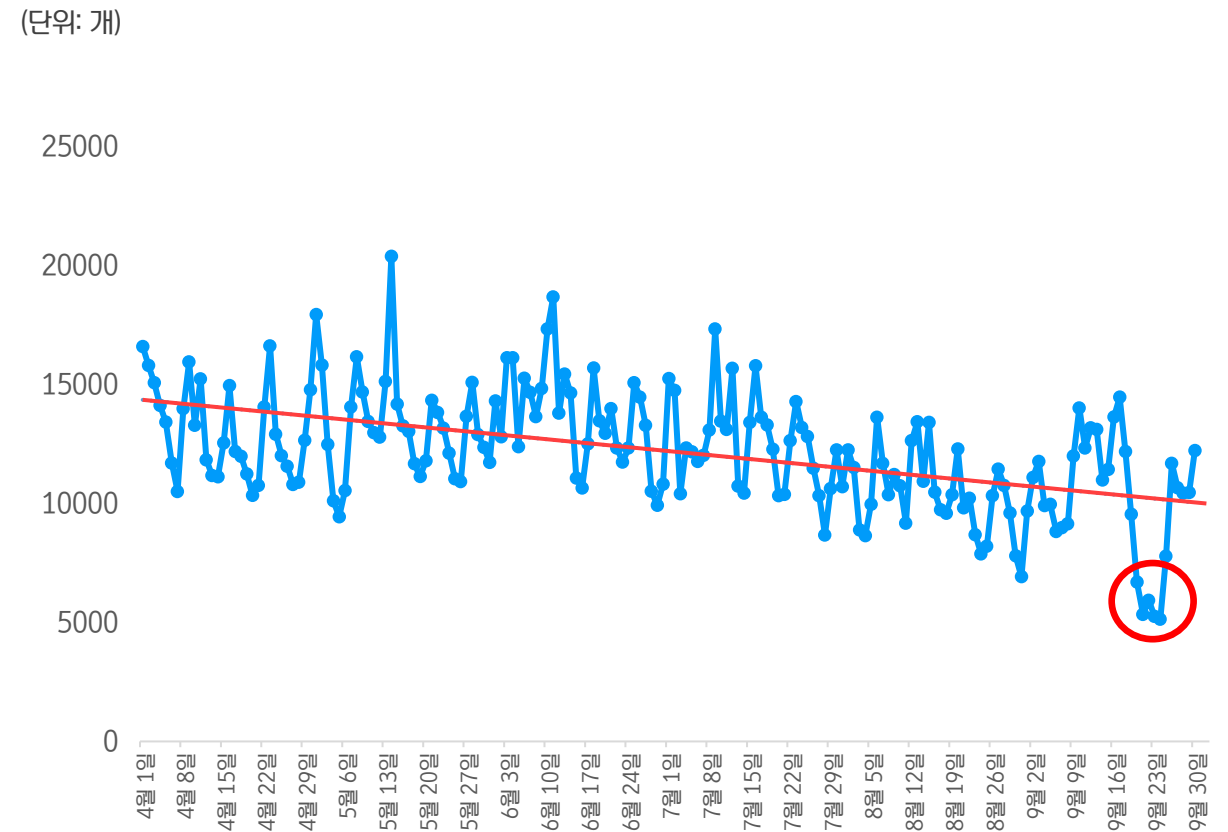
기초 데이터 탐색

③ 세션 데이터 탐색

월별 최고 매출 TOP3 품목



일자별 세션 생성 수





03 문제 원인 파악

미미한 JBP 효과
홈페이지 구성

미미한 JBP(업무 제휴 협약) 효과

① 고려은단

롯데닷컴·고려은단 '업무제휴 협약(JBP)' 체결

입력: 2018-06-29 06:05:58 수정: 2018-06-29 06:14:19 게재: 2018-06-29 15:06:14(98번)



온라인종합쇼핑몰 롯데닷컴은 건강기능식품 기업 고려은단과 고객에게 차별화된 상품과 서비스 제공을 위해 '업무제휴 협약(JBP)'을 체결했다고 29일 밝혔다.

② 브띠엘린

온라인쇼핑몰 롯데닷컴, 프리미엄 유아용품제조 기업 브띠엘린 업무 협약

정의윤 / 기사승인: 2018-07-25 10:37:19



롯데닷컴 김경호 대표(우)가 브띠엘린 표순규 대표(좌)와 업무제휴 협약(JBP)을 체결한 후 악수를 하고 있다.

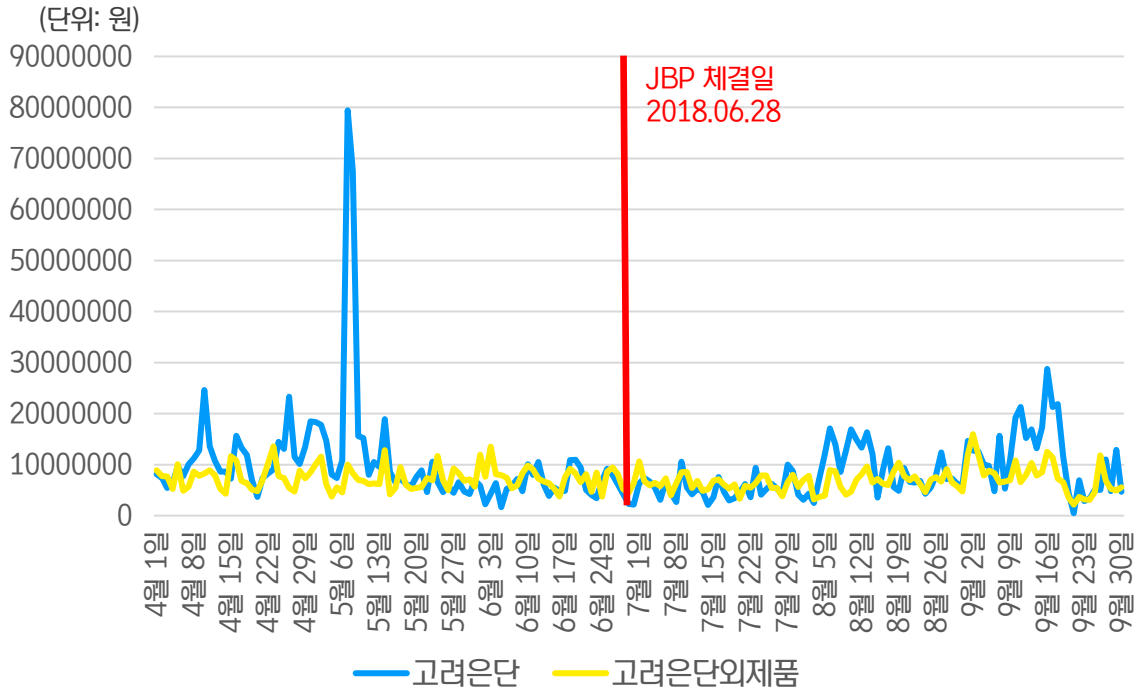
[스페셜경제=정의윤 인턴기자]온라인종합쇼핑몰 롯데닷컴이 프리미엄 유아용품 전문기업 브띠엘린과 고객에게 차별화된 상품과 서비스 제공을 위해 '업무제휴 협약(JBP)'을 체결했다.

✓ 롯데 온라인 쇼핑몰은 현재 매출향상을 위해 다양한 기업들과 업무제휴 협약을 맺고 있음

미미한 JBP(업무 제휴 협약) 효과

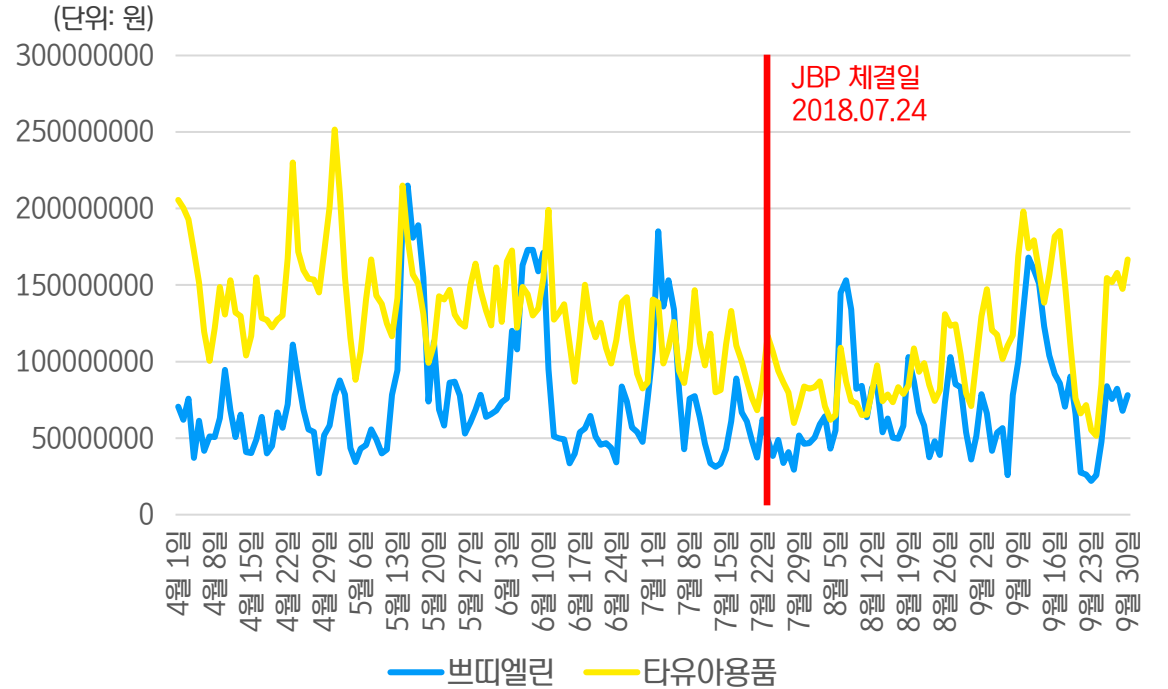
① 고려은단

고려은단 JBP 체결 후 구매액 동향



② 브띠엘린

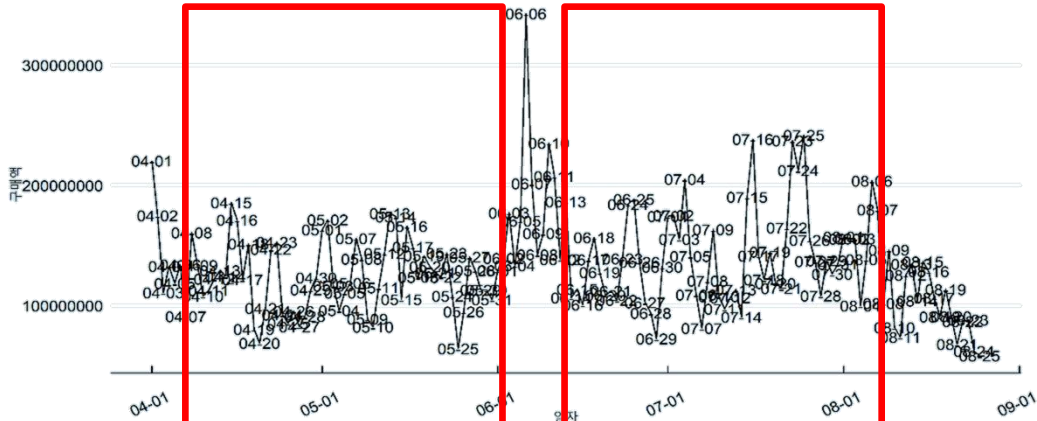
브띠엘린 JBP 체결 후 구매액 동향



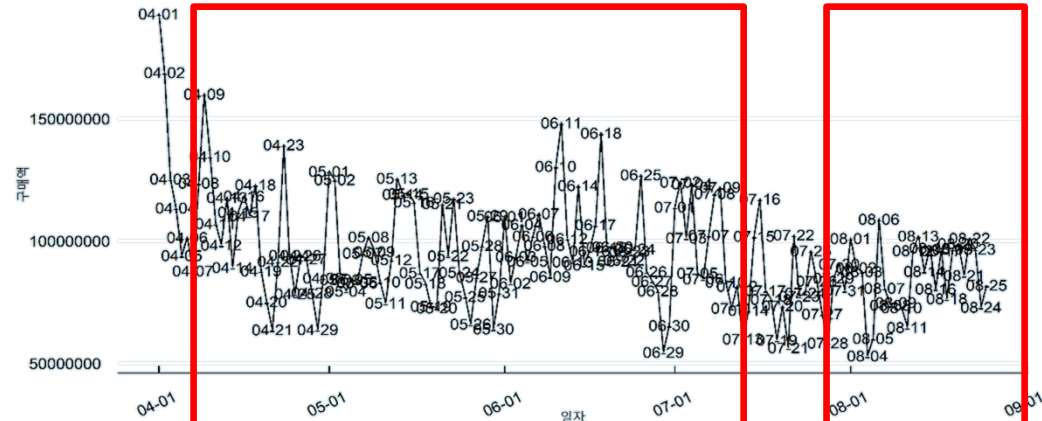
- ✓ 업무 제휴 협약이 지속적인 구매액 상승 곡선을 야기하지 않음
- ✓ 업무 제휴 협약 브랜드의 판매액과 해당 브랜드가 속한 카테고리 내 제품의 구매액을 비교했을 때, 추세의 차이가 미미함
- ✓ 업무 제휴를 통해 매출액 증가 효과를 원한다면, 다른 마케팅 전략이 필요할 것으로 보임

홈페이지 구성

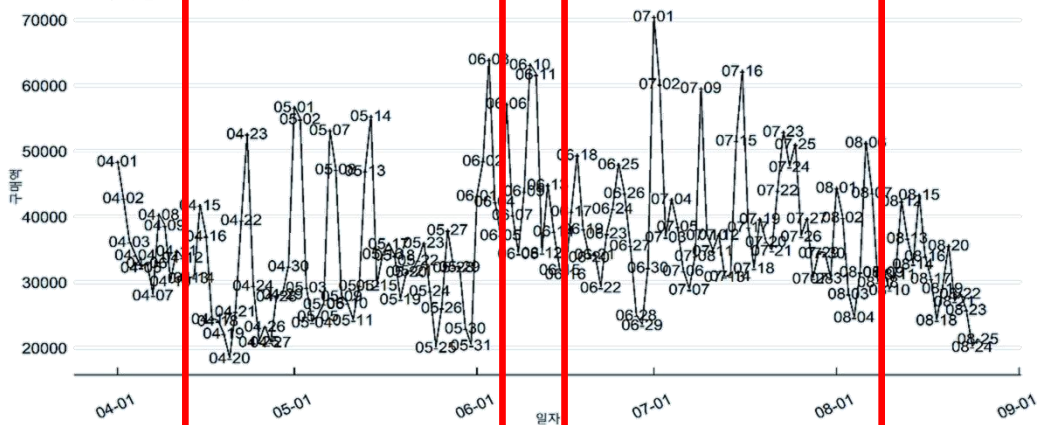
전자제품 구매액



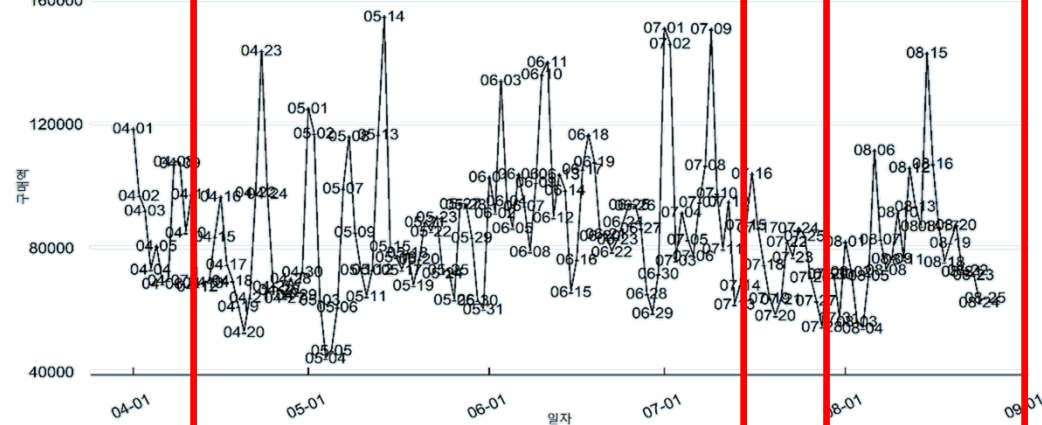
리빙 구매액



전자제품 페이지 뷰 수



리빙 페이지 뷰 수



- ✓ 구매 데이터와 행동 데이터의 추세가 일치하지 않음
- ✓ 원하는 제품을 찾기 위해 많은 시간을 할애해야 함
- ✓ 홈페이지 구성 상에 문제가 있는 것으로 파악되어 이에 대한 개선이 필요해 보임



04

고객 세분화 및 추천전략

고객 세분화
마케팅 전략
홈페이지 구성 전략

고객 세분화

① VIP 고객과 일반 고객 분류

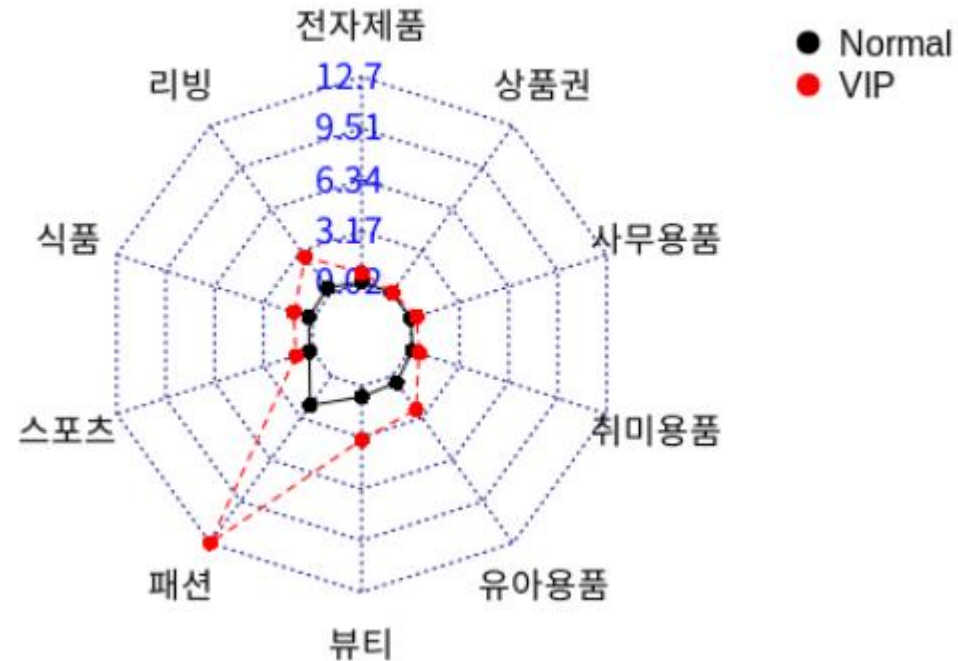
6개월 간의 총 구매 금액



- ✓ 각 고객이 6개월간 구매한 총 금액을 Boxplot으로 시각화
- ✓ 특정 고객들이 아주 고가에 분포되어 있음
- ✓ 이상치 계산 결과 641,450원 이상 구매한 고객그룹
- ✓ 이 그룹을 VIP 고객이라 정의
- ✓ VIP 고객과 일반 고객을 분류하여 분석

고객 세분화

① VIP 고객과 일반 고객 분류

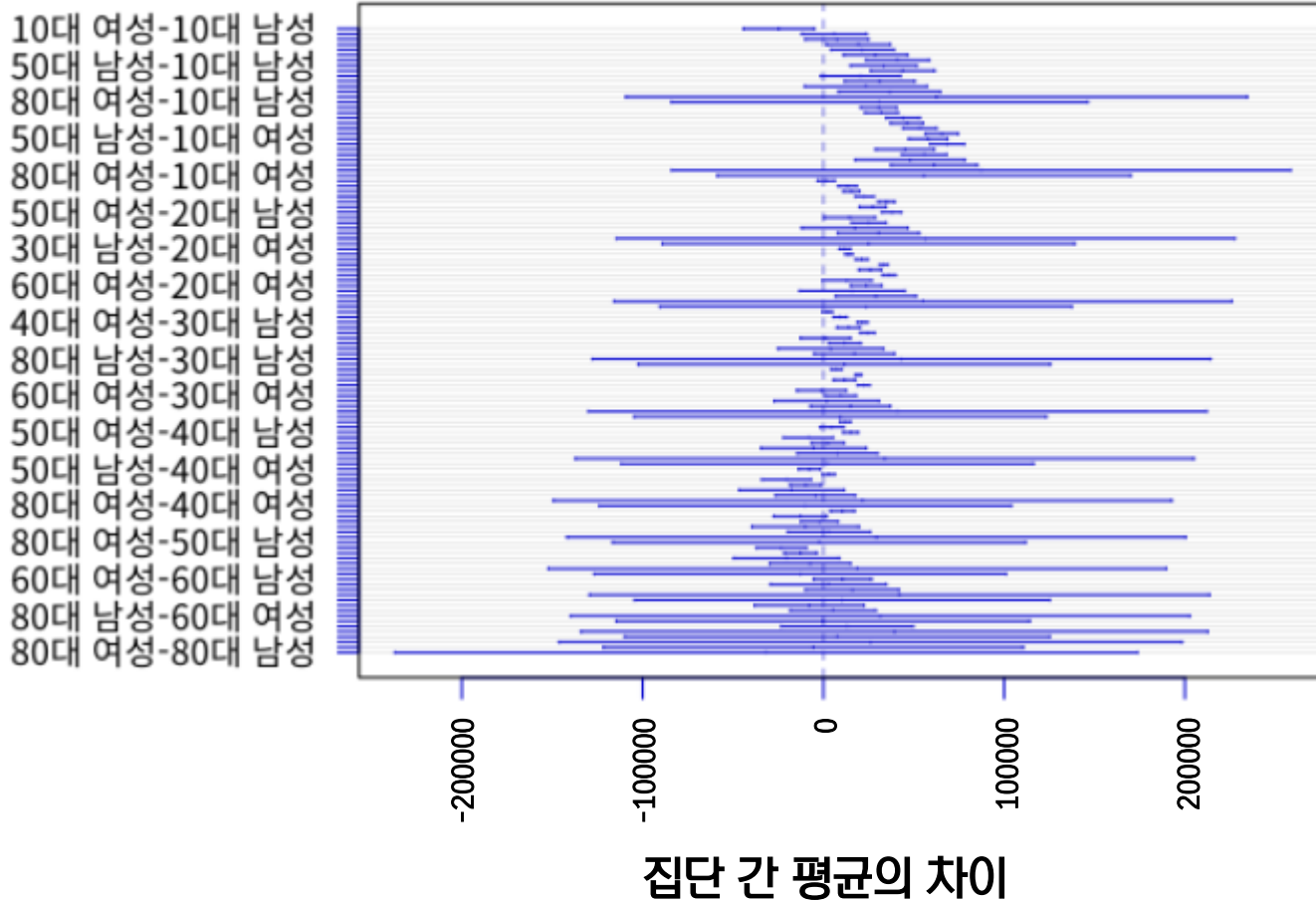


- ✓ VIP 고객과 일반 고객 각각의 상품 대분류별 구매 개수를 변수로 클러스터링 수행
- ✓ VIP 고객이 일반 고객보다 구매 개수가 현저하게 높음
- ✓ 특히 패션 구매가 가장 높음

고객 세분화

② ANOVA 분석 실시 - 일반 고객

95% family-wise confidence level



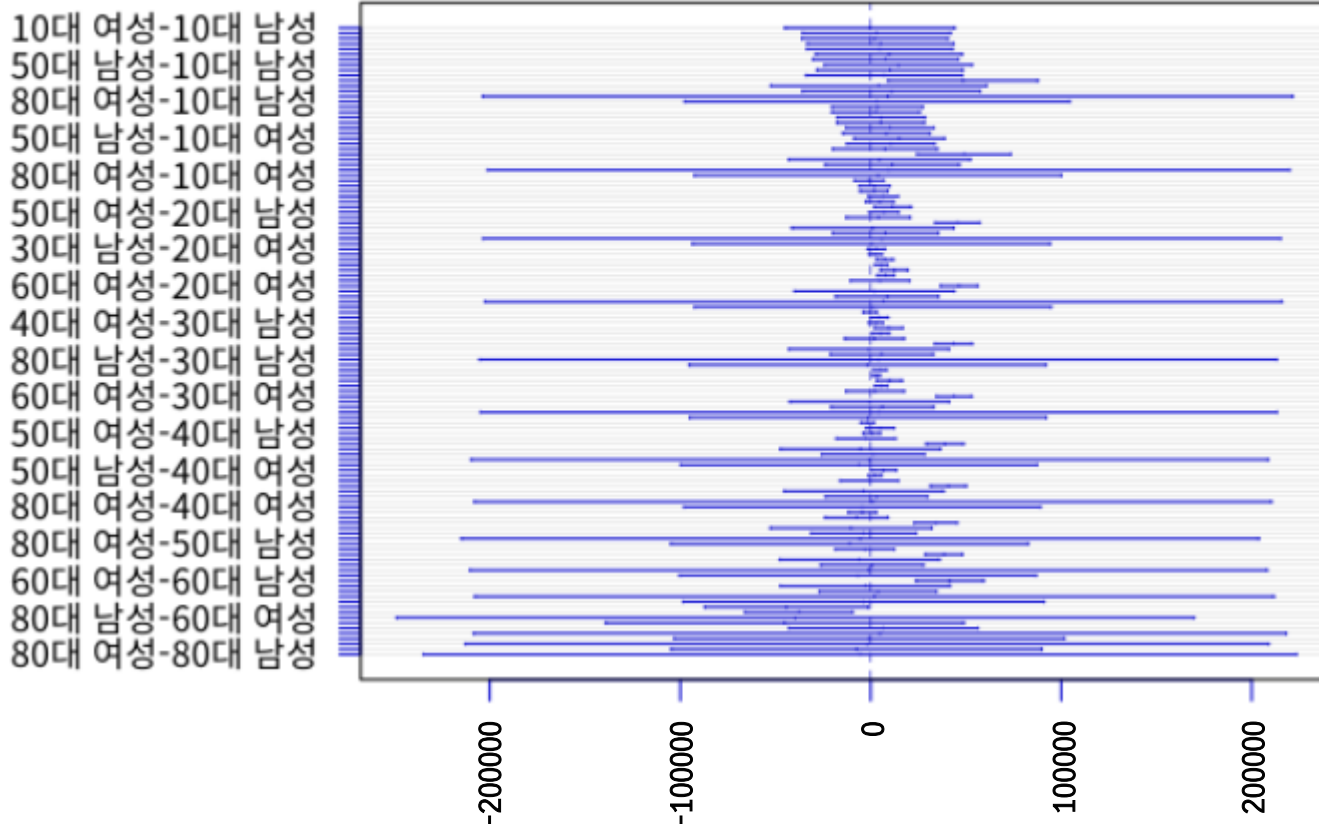
- ✓ 성별 연령별 총 구매 금액으로 ANOVA 분석 실시
- ✓ 구매 금액의 평균의 차이가 없는 집단끼리 묶임
- ✓ 묶인 각 그룹별로 그룹번호를 부여함

그룹번호	1	2	3
해당 집단	10~20대 남성	10대 여성	20대 여성
그룹번호	4	5	6
해당 집단	30대 남성	30대 여성	40~50대 남성
그룹번호	7	8	9
해당 집단	40~50대 여성	60~80대 남성	60~80대 여성

고객 세분화

② ANOVA 분석 실시 - VIP 고객

95% family-wise confidence level

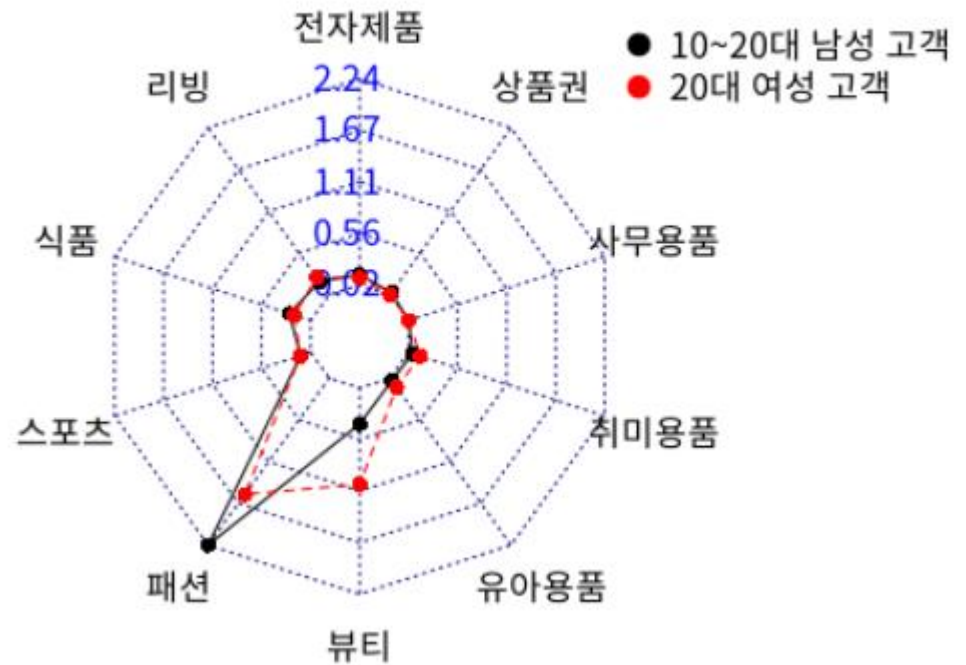
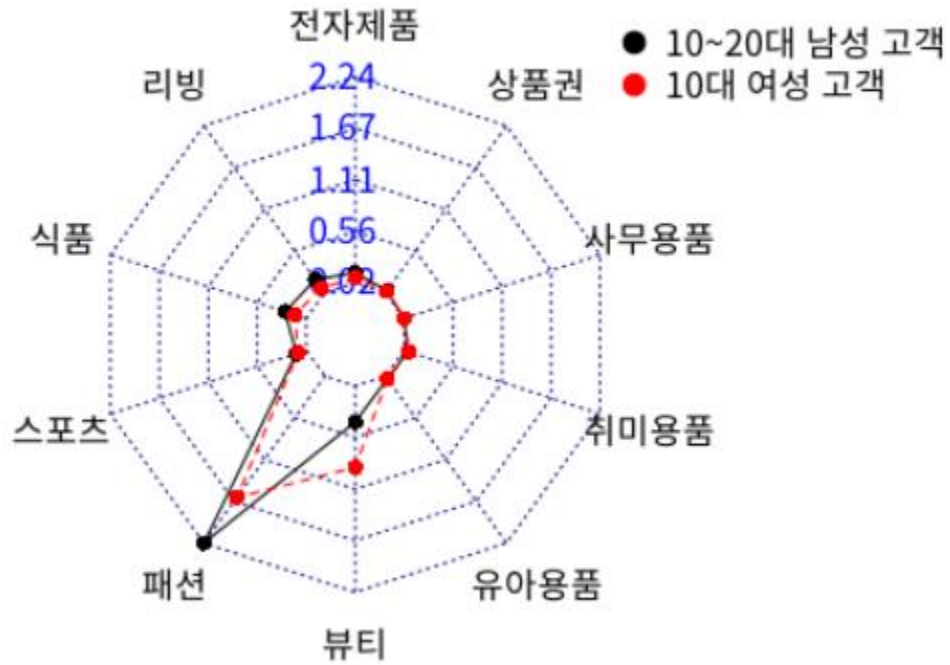


- ✓ 성별 연령별 총 구매 금액으로 ANOVA 분석 실시
- ✓ 평균 구매 금액이 큰 차이를 보이지 않음
- ✓ 따라서, VIP 고객은 따로 세분화하지 않음
- ✓ 일반 고객 분류 번호 9번 다음으로 VIP를 10번으로 지정

집단 간 평균의 차이

고객 세분화

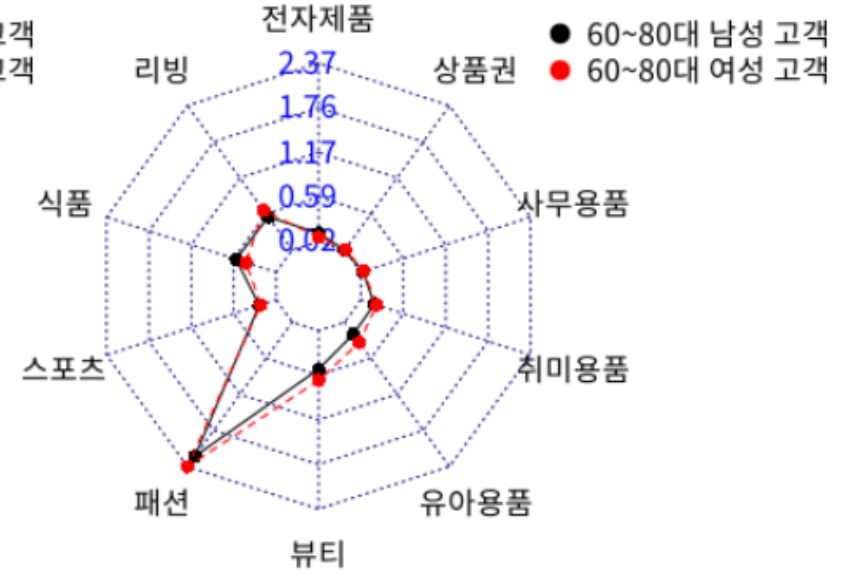
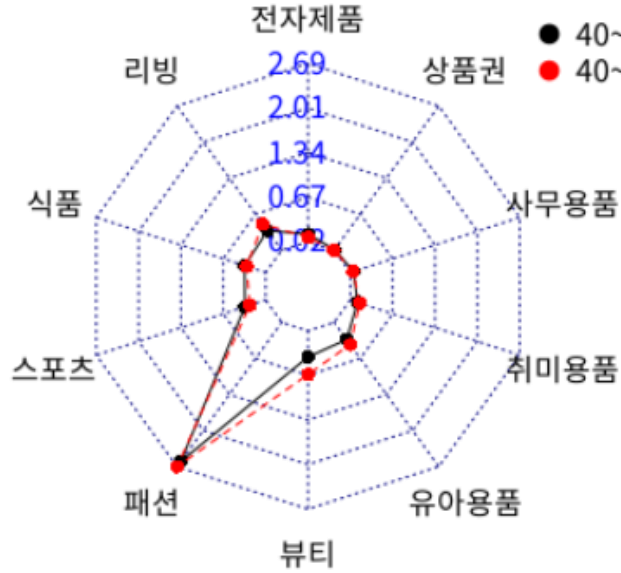
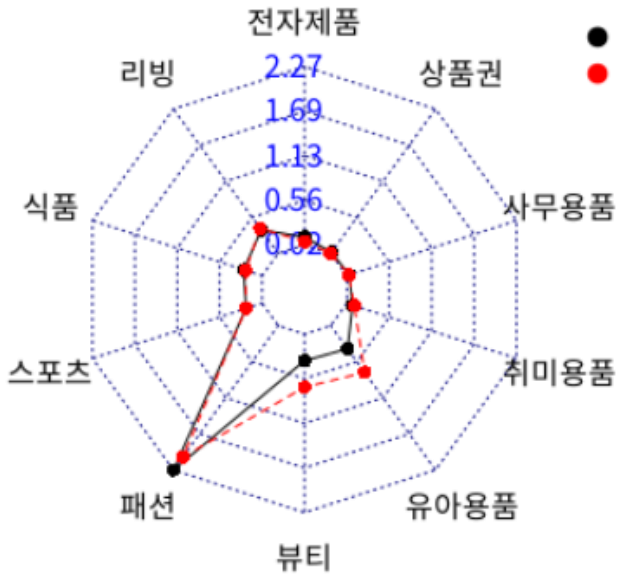
② ANOVA 분석 실시 - 일반 고객의 분류별 비교



- ✓ 모두 패션의 매출이 가장 높음
- ✓ 남성과 달리 여성의 경우 뷰티의 매출도 높음

고객 세분화

② ANOVA 분석 실시 - 일반 고객의 분류별 비교



- ✓ 1-3순위가 패션, 유아용품, 뷰티로 동일함
- ✓ 패션 매출은 비슷하나 유아용품과 뷰티에서 차이가 보임

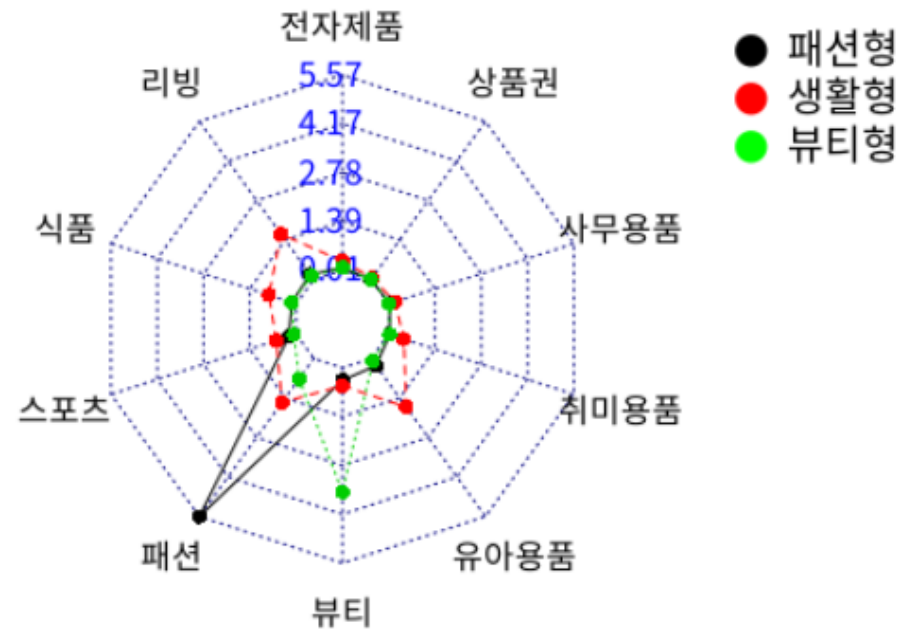
- ✓ 비슷한 구매 행태를 보임

- ✓ 다른 연령대에 비해 식품 구매 수가 유독 많음
- ✓ 영양제와 같은 건강기능식품 구매가 증대하였을 것으로 추측됨

고객 세분화

③ Clustering 수행 - 전체 고객

고객 구매 특성별 레이더 차트

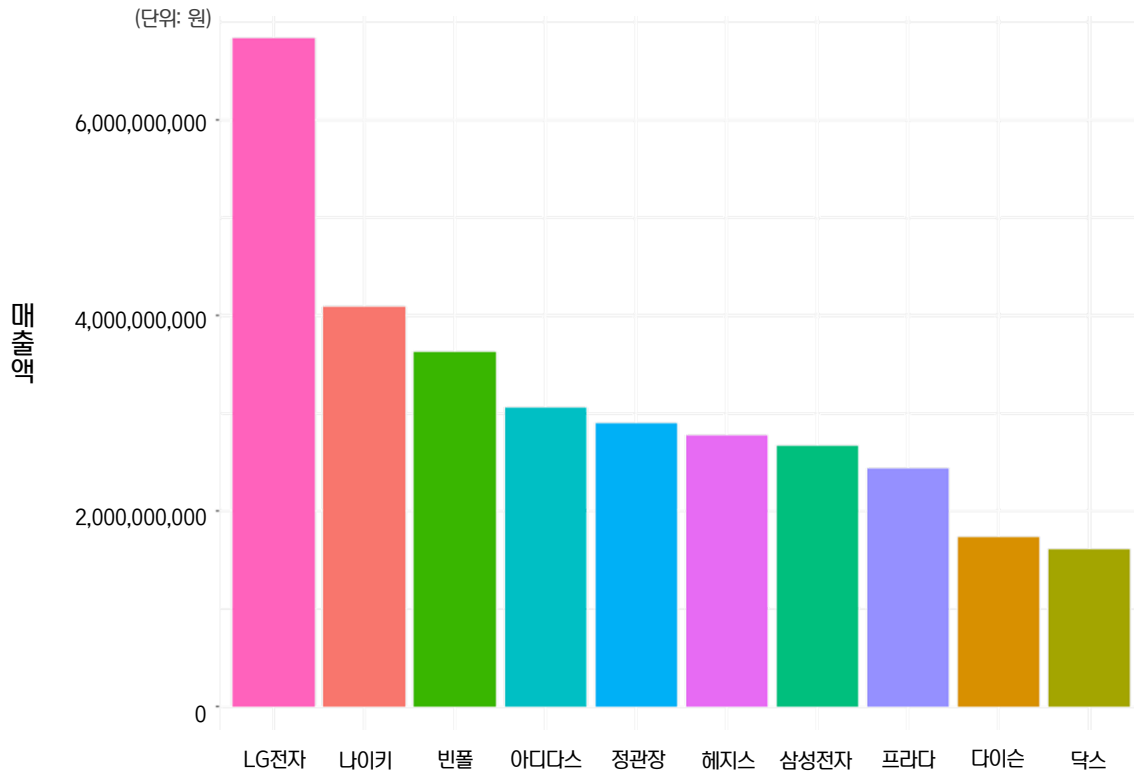


- ✓ 전체 고객의 상품 대분류별 구매 개수를 변수로 클러스터링 수행
- ✓ 적정 클러스터링 개수가 3개로 도출
- ✓ 3개의 집단으로 클러스터링한 결과 크게 패션형, 생활형, 뷰티형으로 나뉨

마케팅 전략

① JBP 전략

최고 매출 브랜드 TOP 10

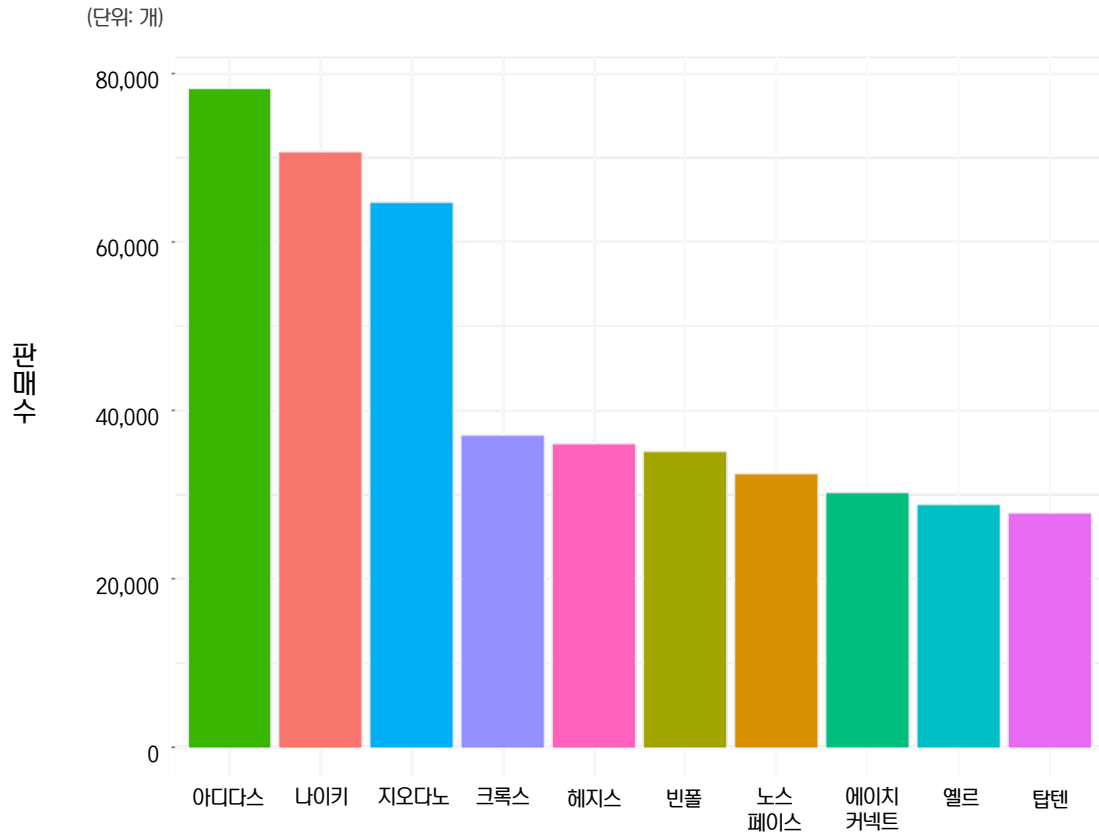


✓ 인기 많은 브랜드와의 프로모션 제안

마케팅 전략

① JBP 전략

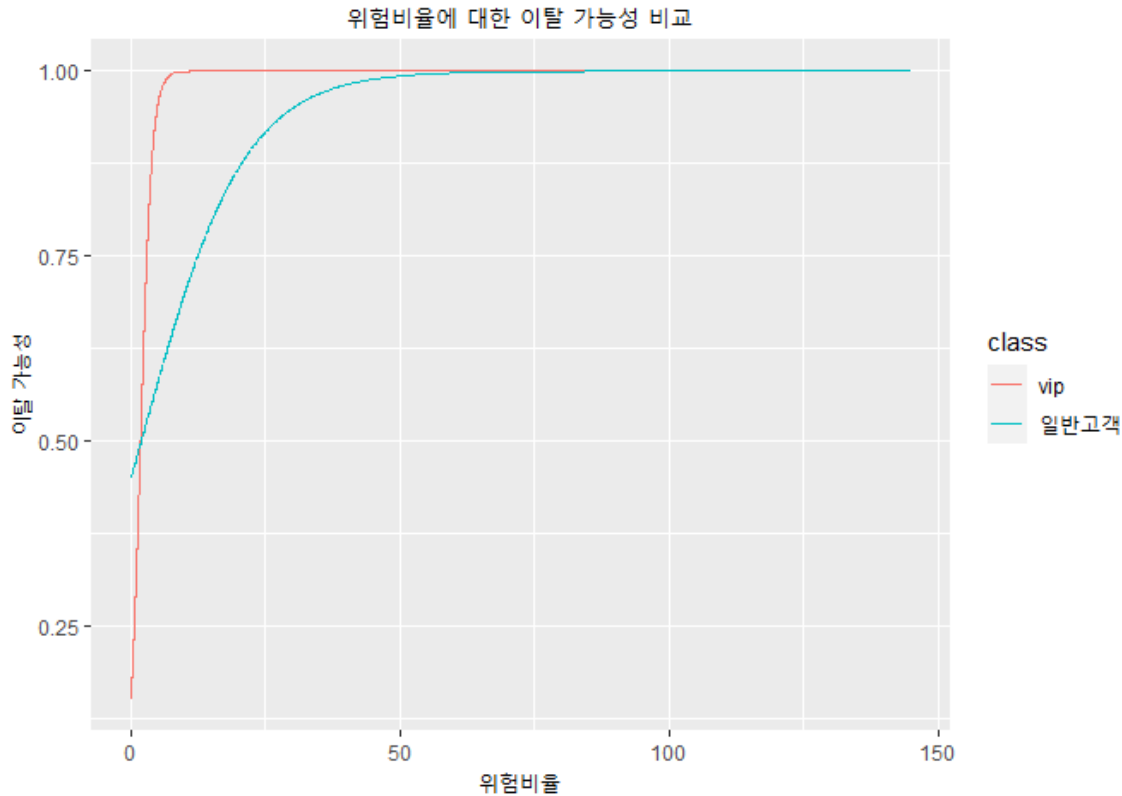
최다 판매 브랜드 TOP 10



✓ 인기 많은 브랜드와의 프로모션 제안

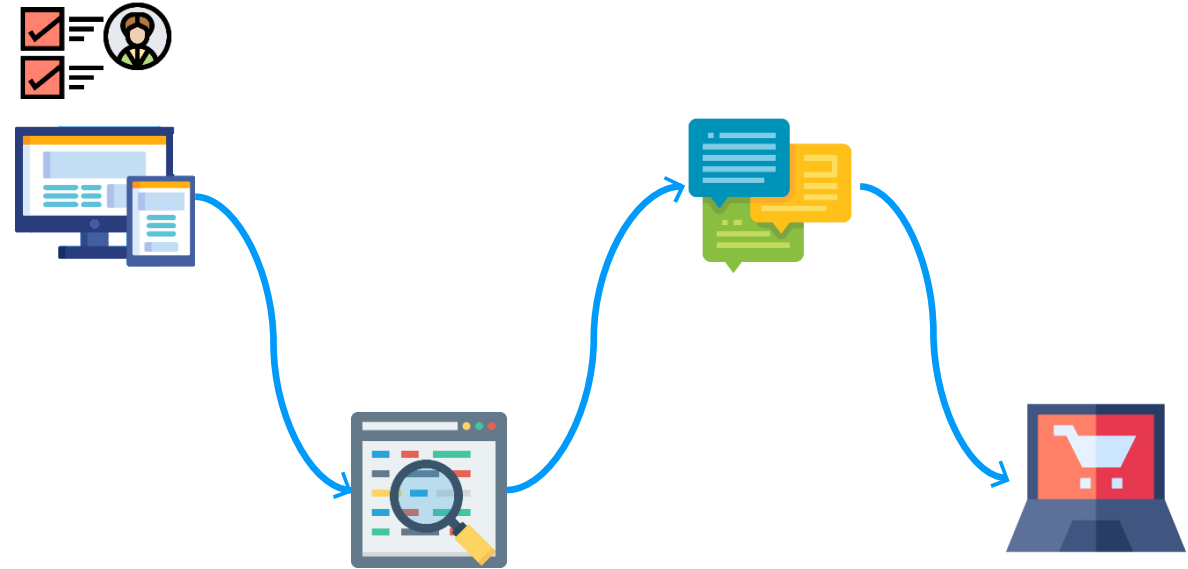
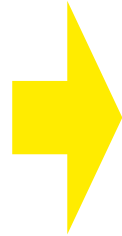
마케팅 전략

② Remarketing



- ✓ VIP고객 : 약간만 오래 접속을 하지 않아도 이탈할 가능성이 높아짐
- ✓ 일반고객 : VIP 고객보다 평균 접속 주기보다 더 오랜 기간 접속하지 않더라도 재접속할 확률이 높음
- ✓ 각 고객 이탈 주기에 맞춰 선호 상품 판매 정보를 e-mail, 쿠폰 발송 등을 통해 push marketing 진행

홈페이지 구성 전략

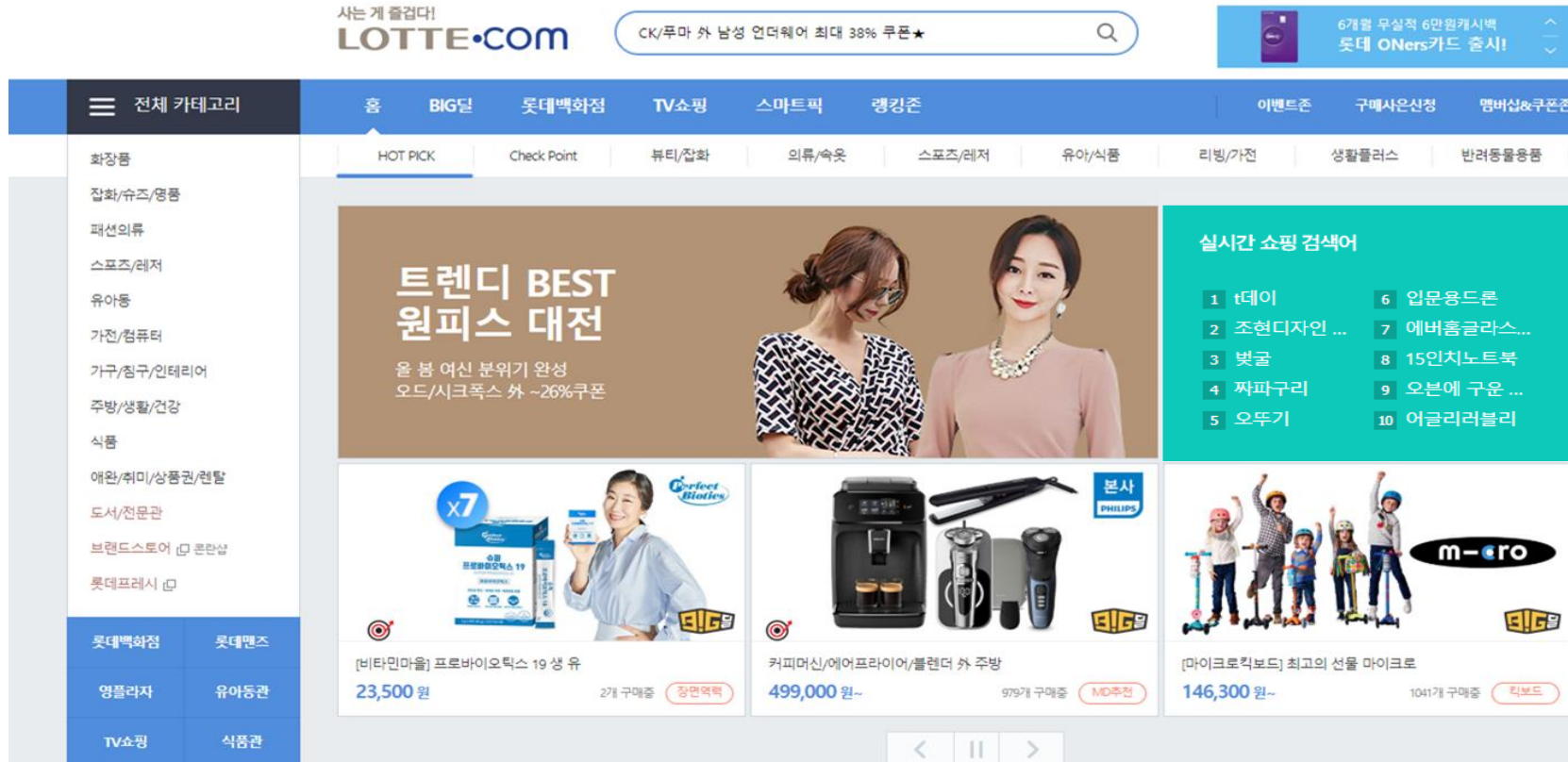


대형 마트의 상품 진열은 고객의 이동 경로를 따라서
고도의 과학에 의해 상품이 배열되어 있다

온라인에서도 구매 경로에 따라서
'자신에게 맞는', '꼭 필요한' 상품을 추천해주자!

홈페이지 구성 전략

① Main Page - 로그인 전

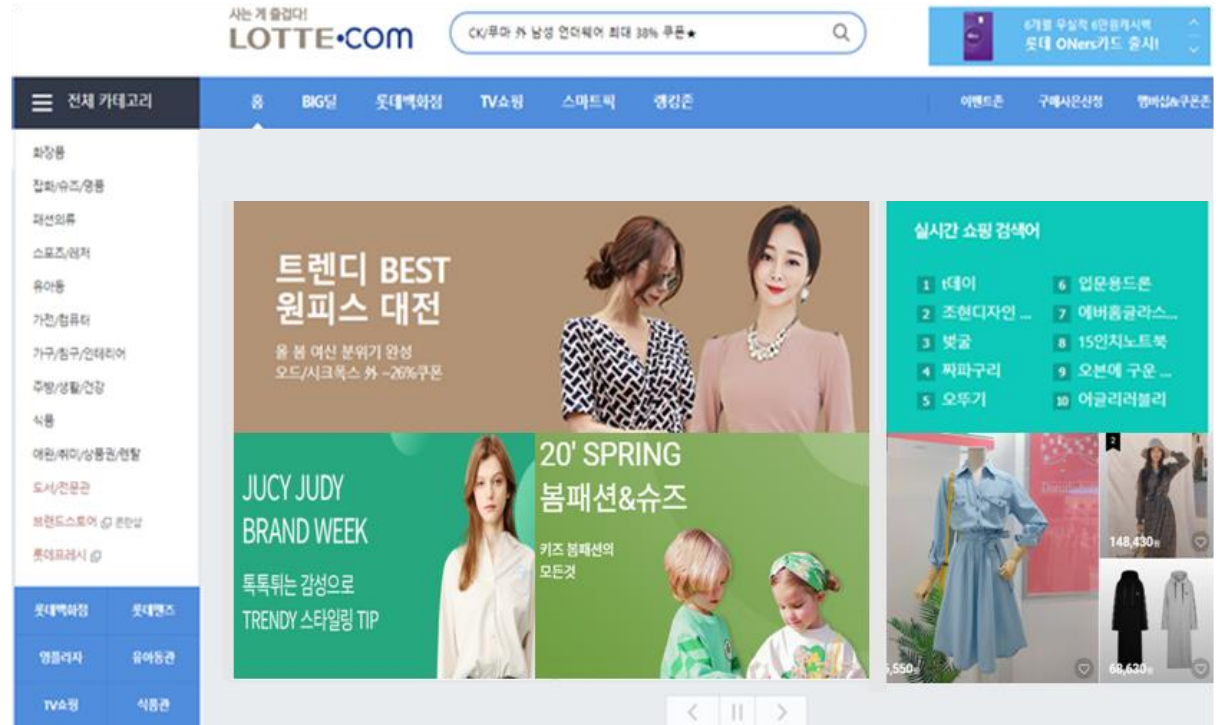
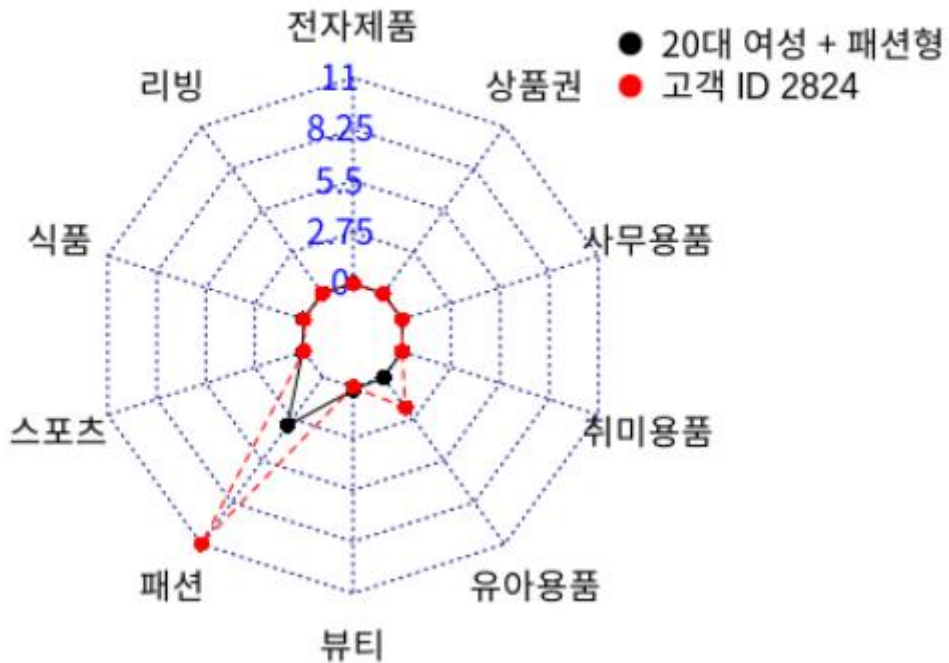


- ✓ 시계열 분석을 통한 트렌디한 제품전 기획
- ✓ 실시간 급상승 검색어 추가

홈페이지 구성 전략

① Main Page - 로그인 후

고객 ID 2824 와 20대 여성 + 패션형 비교



- ✓ ID 2824 고객의 패션과 유아용품 구매 성향이 소속된 그룹보다 강한 것을 확인할 수 있음
- ✓ 이에 따라 메인 페이지를 패션제품과 유아용품 구매 비율로 구성함

홈페이지 구성 전략

② 제품 상세 Page

패션 연관규칙 (여성)

size: support (0.01 - 0.031)
color: lift (1.961 - 8.792)

브래지어 → 여성팬티



- ✓ 연관 분석을 통해 동일 카테고리 내 함께 구매되는 제품 규칙을 발견함
- ✓ 연관 분석 결과 겹옷을 산 고객들은 상의와 하의를 함께 사는 패턴을 보임
- ✓ 따라서 자켓 제품 상세 페이지에서 상의와 하의 제품을 추천해주도록 구성함

홈페이지 구성 전략

③ 장바구니 Page

< ID 2824 고객의 구매내역 >

고객식별번호	성별	나이대	중분류명
2824	F	20	여성의류상의
2824	F	20	남성의류상의
2824	F	20	유아의류상의
2824	F	20	여성가방
2824	F	20	유아동속옷

< ID 2824 고객의 추천내역 >

고객식별번호	성별	나이대	추천상품1	추천상품2	추천상품3
2824	F	20	건강보조식품	닭고기류	여성화

장바구니

일반 (1)

장보기(0)

정기배송(0)

전체선택 선택삭제

롯데백화점

쉬즈미스 린넨 테일러드 자켓
SWWJKK11460
55 / 아이보리(IV)
5/4(월) 발송 예정

1

94,520원
쿠폰변경

111,200원 + 배송비 0원 - 할인 16,680원 = 94,520원(30,000원 이상 무료배송) 묶음배송 상품 보기 >

결제예정금액

상품금액 111,200 원

배송비 0 원

상품할인금액 -16,680 원

총 1건 **94,520 원**

주문하기

선물하기

김롯데님을 위한 추천



- ✓ 아이템 기반 협업 필터링을 통해 다른 카테고리에서 함께 구매되는 제품 규칙을 발견함
- ✓ 분석 결과, ID 2824 고객은 건강보조식품, 닭고기류, 여성화를 구매할 확률이 높음
- ✓ 따라서 장바구니 페이지에서 해당 제품을 추천하도록 구성함

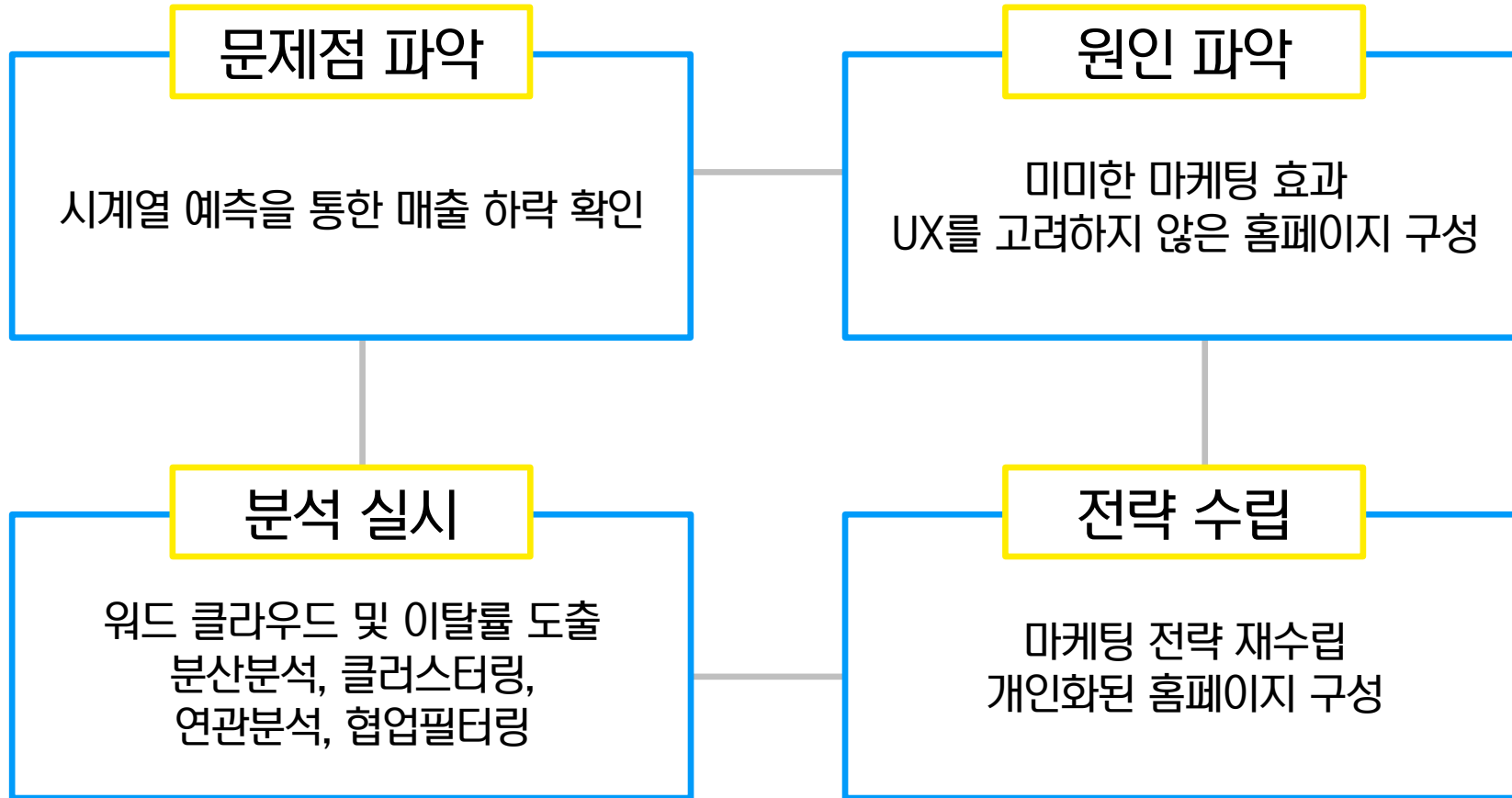


05 결론

요약

한계점 및 향후 연구 과제

역할 분담



한계점 및 향후 연구 과제

데이터의 한계

- 과거 데이터의 총 6개월이라는 짧은 분석기간
- 시계열 분석 중 계절적 요인 파악과 예측의 한계
- 머신러닝을 이용한 클러스터링 및 상품 추천 알고리즘은 새로운 데이터의 유입으로 유의미한 결과 도출가능

고객 세분화 한계

- 아마존의 경우 고객들을 0.1명 단위로 세분화해 상품을 추천하고 홈페이지를 구성
- 상품군 뿐만 아니라 다양한 변수들을 고려하여 고객들의 특성을 파악했다면 더욱 정교하고 세밀한 상품 추천이 가능

알고리즘의 정교성

- 최종 제안 방안은 페이지 구성 방향일 뿐, 구체적인 알고리즘을 제시한 것이 아님
- 실제 개인화된 페이지 구축이 향후 연구 과제로 남음

역할 분담



제우스 (유찬희)
신들의 신



디오니소스 (이용빈)
술의 신



크로노스 (이동일)
시간의 신



헤르메스 (박대한)
전령의 신



아테나 (안영은)
지식의 신



데미테르 (김수빈)
곡식의 신



아폴론 (김아름)
음악의 신

THANK YOU!

