

주류 구매자 클러스터 분석을 통한 개인화 마케팅 방안

제2회 유통데이터활용경진대회
비즈니스아이디어 부문

빅아미

목차

문제 정의

주류 유통 시장의 불확실성과
스마트 오더라는 새로운 기회

데이터 분석

클러스터 분석
장바구니 분석

활용 방안

주류 구매 고객 특성별
개인화 마케팅

문제 정의

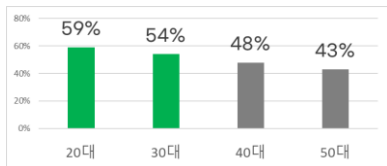


불확실성이 커져가는 주류 시장

시장의 축소, 제품의 다변화, 타겟 확대에 따른 경쟁 심화

주류 소비 유형 다양화

- 홀술 시, 다양한 주종에 도전



- 과음하지 않고 가볍게 술을 즐기는 문화가 필요하다. (94%)

소비 유형별 타겟 마케팅 필요

- MZ 세대
 - 위스키, 와인, 증류 소주 등 고급 주류
 - 미니사이즈, 논알콜 등 니즈 다양화



- 여성층
 - 특이한 맛, 색깔이 예쁜 퓨전 탁주 인기
 - 증류식 소주 선호

경쟁 심화

경제

“회식 줄고 안 팔리는데”...경쟁 치열해지는 주류시장

이성현 기자 lee.sanghyun@mkinternet.com

입력: 2023-04-05 18:29:59

가

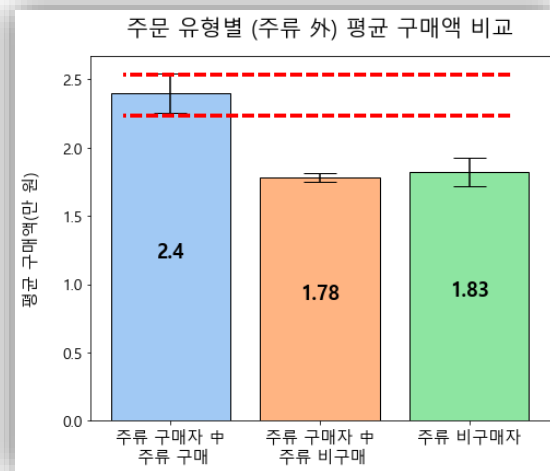
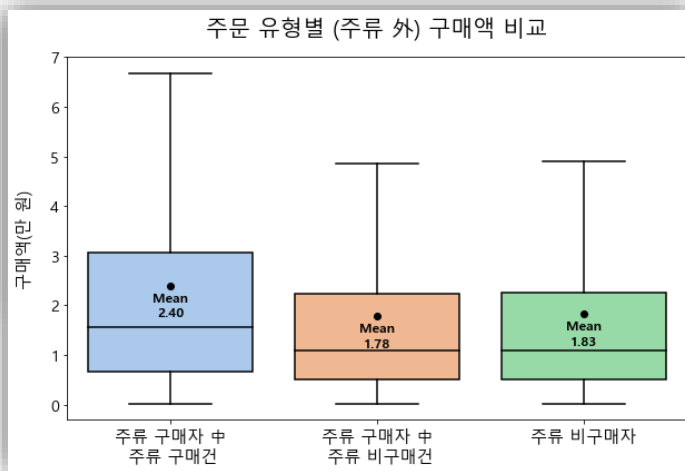
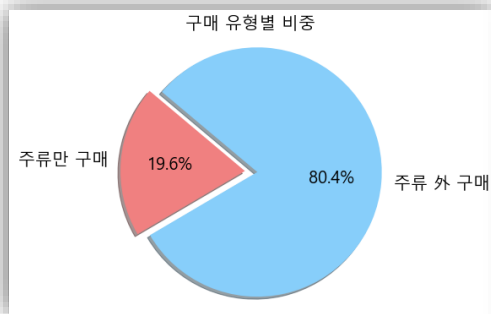


지난 4일 하이트진로 홍천공원에서 맥주 신제품 '엘리'가 출시되는 모습. 하이트진로 임직원들이 엘리의 첫 출고를 축하하고 있다. [사진 제공 = 하이트진로]

팬데믹 여파로 회식이 줄어드는 등 음주 문화가 변하면서 국내 주류 출고량이 7년 연속 감소세를 기록했다. 전반적인 시장 파이가 줄어드는 가운데 주요 기업들이 신제품을 출시하고 새 광고 모델을 채용하는 등 치열하게 경쟁하는 분위기다.

유통업자에게 주류 구매자는 왜 중요한가?

고객은 주류를 구매할 때, 그렇지 않을 때보다 다른 상품들을 더 많이 구매한다



→ 95% 신뢰구간 오차막대를 사용

주류를 구매하는 고객이 그렇지 않은 고객보다 주류 외 상품을 평균 5,700원 더 구매

*차이검정 결과 : p. 25 참고

주류 마케팅의 변화

스마트 오더 시장의 기회와 GS 리테일의 전략

우리동네 GS 내 '와인25+' 스마트오더 시장 내 점유율 약 80%
(22년 기준 매일 경제)

기존 유통 플랫폼 신규 서비스 추가

스마트오더 커머스 스타트업



- 우리동네GS는 O4O* 경쟁력 강화에 집중
- 데이터 활용 스마트 오더 추천으로 O4O 전략 추진



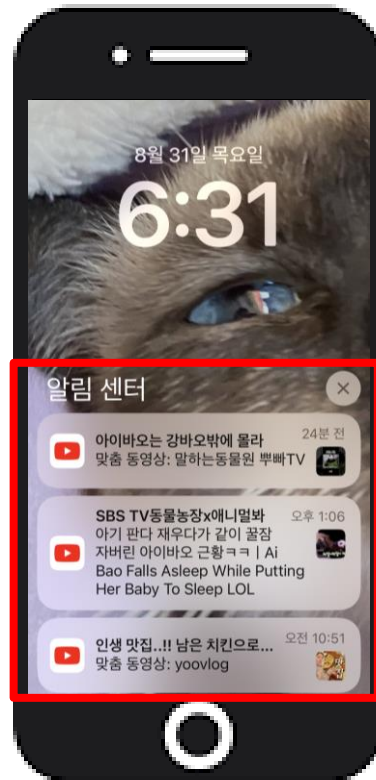
GS리테일의 오프라인 매장을 '허브' 로

트렌디하고 한 가지 컨셉에 특화된 프로모션

문제 정의

**유통(구매) 데이터를 활용하여,
소비자 정보를 기반으로 주류 구매 유형을 정의하고
각 유형별 타겟 마케팅 전략을 수립한다.**

- 스마트 오더의 O2O환경을 최대한 활용
- 마케팅 채널 : 온라인 개인화 마케팅 전략으로 자주
활용되는 APP 푸시 메시지



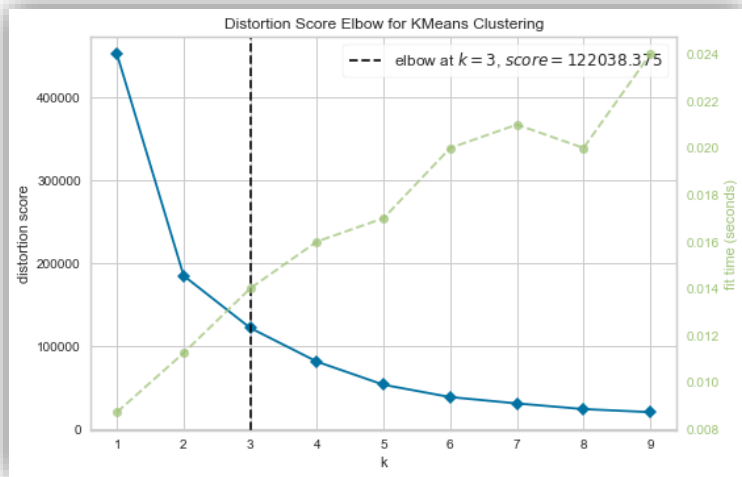
데이터 분석



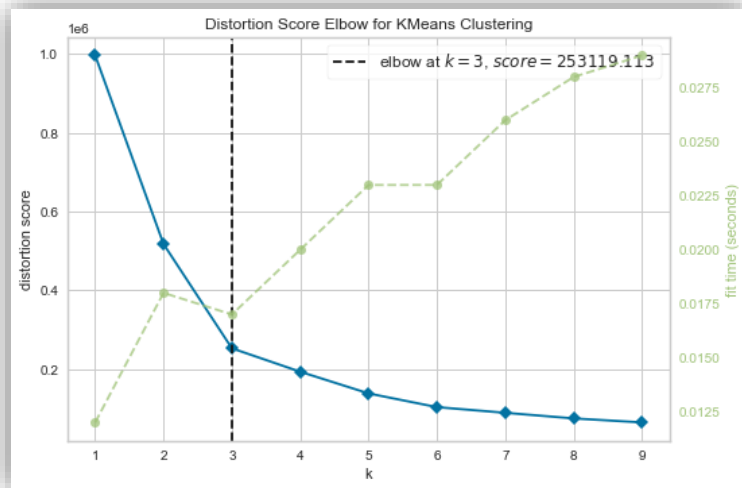
클러스터링(1/2)

1. 주류 구매 건수 2회 이상인 personID의 구매 내역 위주로 분석
2. 연령대를 40대 미만 / 40대 이상으로 구분하여 클러스터링 수행
3. 각 연령대별로 3개의 클러스터가 적합한 것으로 나타남

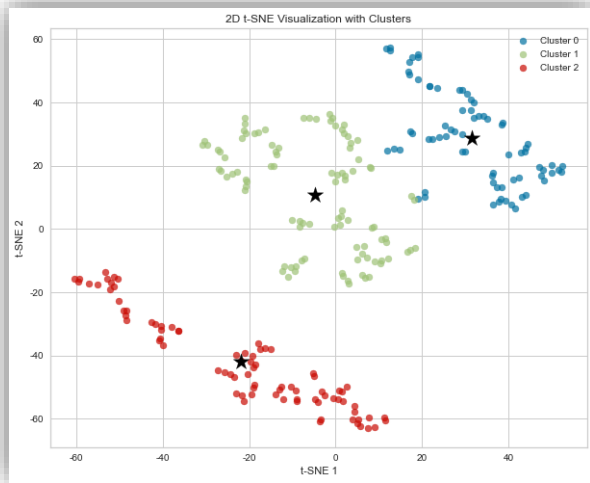
↓ 40대 미만 주류구매자 엘보플랏



↓ 40대 이상 주류구매자 엘보플랏



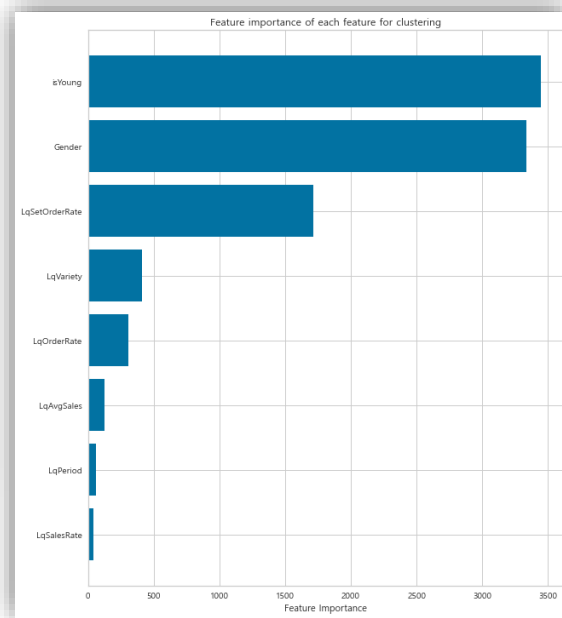
클러스터링(2/2)



40대 미만 주류구매자 클러스터링 결과



40대 이상 주류구매자 클러스터링 결과



1) Min-Max 스케일링 → 2) t-SNE 차원 축소 → 3) K-means 클러스터링

- 30대 이하 / 40대 이상 → 클러스터 0, 1, 2 / 클러스터 3, 4, 5

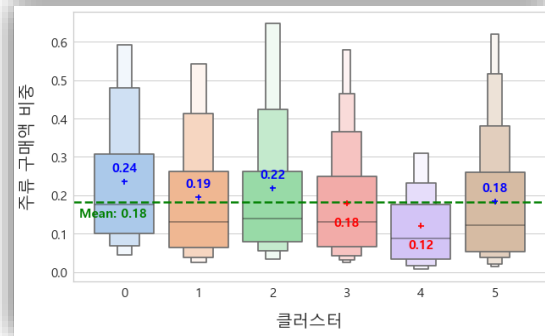
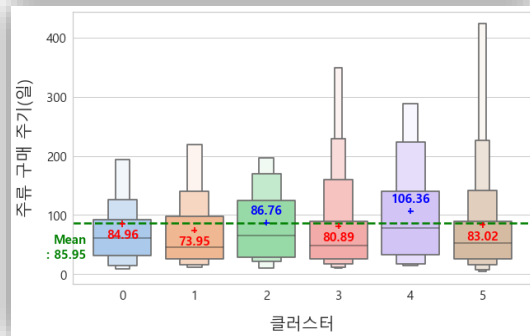
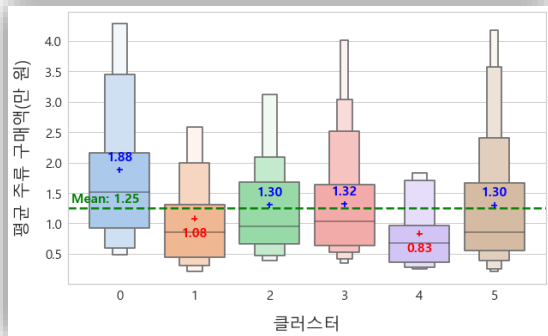
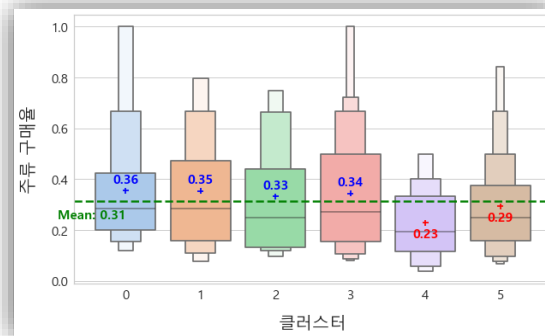
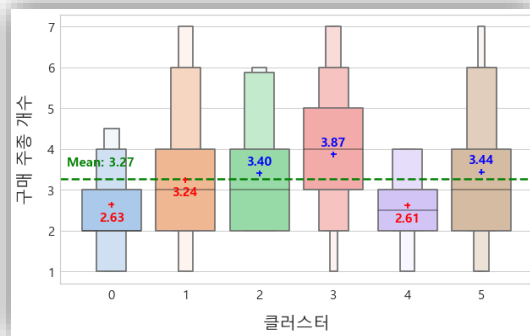
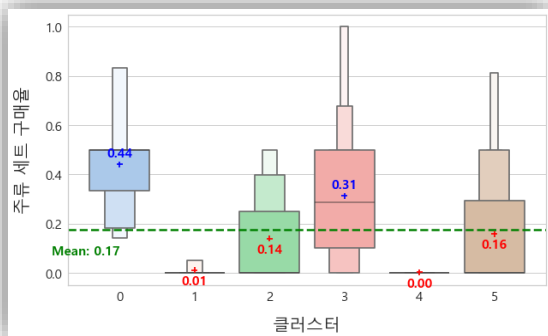
- 성별 → 여 : 클러스터 0, 1, 3, 4 / 남 : 클러스터 2, 5

- lightGBM 활용, feature로 클러스터 분류 학습 결과,

주류 세트 구매율 > 구매주종 개수 > 주류 구매율 > 평균 주류 구매액 > 주류 구매 주기 > 주류 구매액 비중 순으로 feature importance가 높음.

클러스터링 결과

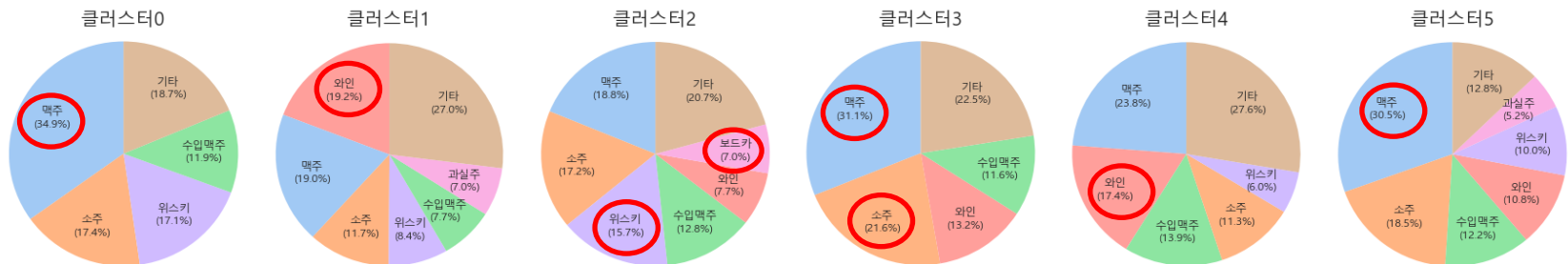
시각화 - 군집별 특징 비교(좌측 상단부터 중요도 순)



클러스터링 결과

클러스터별 선호 주종과 주류 구매 요일별 평균 금액

클러스터별 주류 구매액 중 주종별 비율



요일	C0	C1	C2	C3	C4	C5
월	19401.6	11662.14	11837	13019.9	7291.53	9467.19
화	20023.75	7090.87	11495.33	12447.98	5190.71	10954.44
수	25533.5	11710	20397.35	9684.33	6574.69	13705.86
목	9861.03	8344.9	12385.91	15491.6	6870.24	14770.59
금	25374.33	15609.6	15444.08	11392.67	6371.78	13402.45
토	17469.8	11156.32	12842.86	12648.81	10559.22	12281.41
일	12985.61	11107.2	11636.73	14976.32	8508.87	12840.98

분석 결과

클러스터별 장바구니 분석

클러스터 0	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	국산맥주*	수입맥주	0.01	0.15	4.18
		소주	0.02	0.24	3.09
		연식품(어묵 등)	0.01	0.16	1.67
		즉석/조리식품	0.01	0.19	1.33
		라면류	0.02	0.28	1.33

클러스터 1	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	국산맥주	수입맥주	0.01	0.19	5.71
		비스킷	0.02	0.35	1.20
		라면류	0.01	0.18	1.19
		드레싱/소스	0.01	0.15	1.00

클러스터 2	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	수입맥주	비스킷, 맥주	0.01	0.15	9.36

클러스터 3	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	소주*	맥주	0.01	0.19	2.70
		라면류	0.02	0.26	1.26
		비스킷	0.02	0.34	1.26
		즉석/조리식품	0.01	0.17	1.10

클러스터 4	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	국산맥주	비스킷	0.01	0.31	1.21

클러스터 5	프로모션 상품	관련 상품	지지도	신뢰도	향상도
	국산맥주*	소주	0.01	0.15	2.73
		라면류	0.02	0.25	1.27
		사탕류	0.01	0.16	1.21
		비스킷	0.02	0.34	1.20
		즉석/조리식품	0.01	0.17	1.11

*구매 주종 비율 분석과 상응하는 제품 카테고리
빈발항목집합으로 나온 '맥주', '소주' 상품을 프로모션 상품으로 선정하여 분석진행

분석 결과

클러스터 간 상대적 특성 비교

구분	인구 특성	주류 세트 구매율	구매 주종 개수	주류 구매율	평균 주류 구매액	주류 구매 주기	주류 구매액 비중	최대 주류 구매 요일	구매 주종 특징
C0	Y, 여성	極多	極少	多	極多	짧다	極多	수, 금	국산맥주, 위스키 선호
C1	Y, 여성	極少	少	多	少	많이 짧다	多	금	와인/과실주
C2	Y, 남성	少	多	多	多	길다	多	수	위스키/보드카
C3	O, 여성	多	極多	多	多	짧다	少	목, 일	국산맥주/소주
C4	O, 여성	極少	極少	極少	極少	많이 길다	極少	토	와인/수입맥주
C5	O, 남성	少	多	少	多	짧다	多	목	국산맥주/소주 (무난)

활용 방안



분석 결과 활용

클러스터 특성별 타겟 마케팅 제안

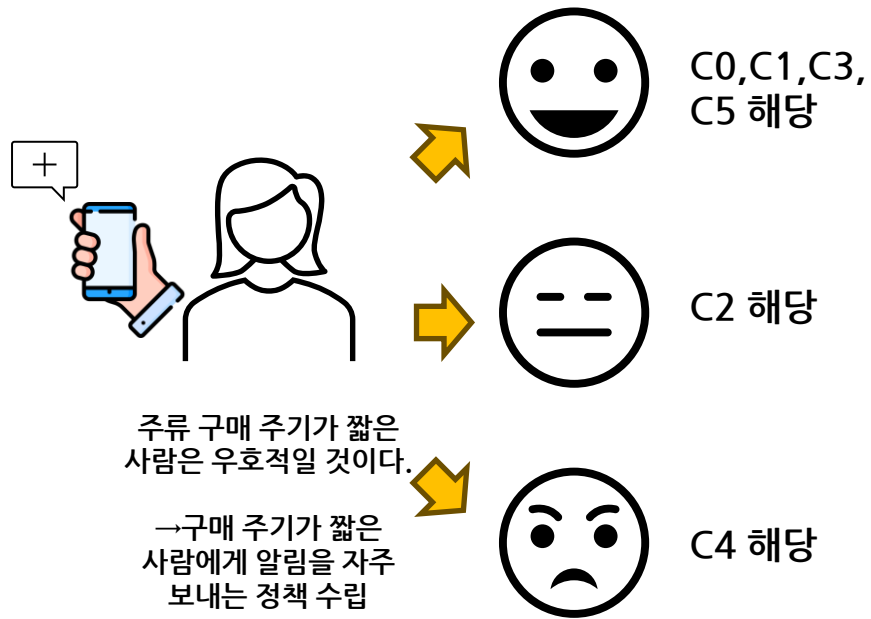
- App 알림 콘텐츠, 알림 시점

구분	알림 선호 예상 정도	중점 프로모션 상품	중점 프로모션 콘텐츠	푸쉬 발송요일
C0 술꾼형 젊은 여성 애주가	우호적	국산맥주/위스키	대량 패키지	수, 금
C1 참새형 젊은 여성 애주가	우호적	와인/과실주/기타	알뜰 할인 상품	금
C2 독주 선호 젊은 남성 애주가	비우호적	위스키/보드카	주류 구매시 라면/비스킷/즉석조리식품 할인 쿠폰	수
C3 대량 여성 약주가	우호적	국산맥주/소주	신제품 출시	목, 일
C4 휴일 음주형 여성 약주가	비우호적	와인/수입맥주	취향별 주류 추천, 기념일 프로모션	토
C5 참새형 남성 애주가	우호적	국산맥주/소주	주류 구매시 라면/비스킷/즉석조리식품 할인 쿠폰	목

분석 결과 활용

클러스터 특성별 타겟 주류 판매 O2O 마케팅 적용 예시

Case1. APP 푸쉬 메시지 발송 횟수



Case2. 클러스터별 맞춤 메시지

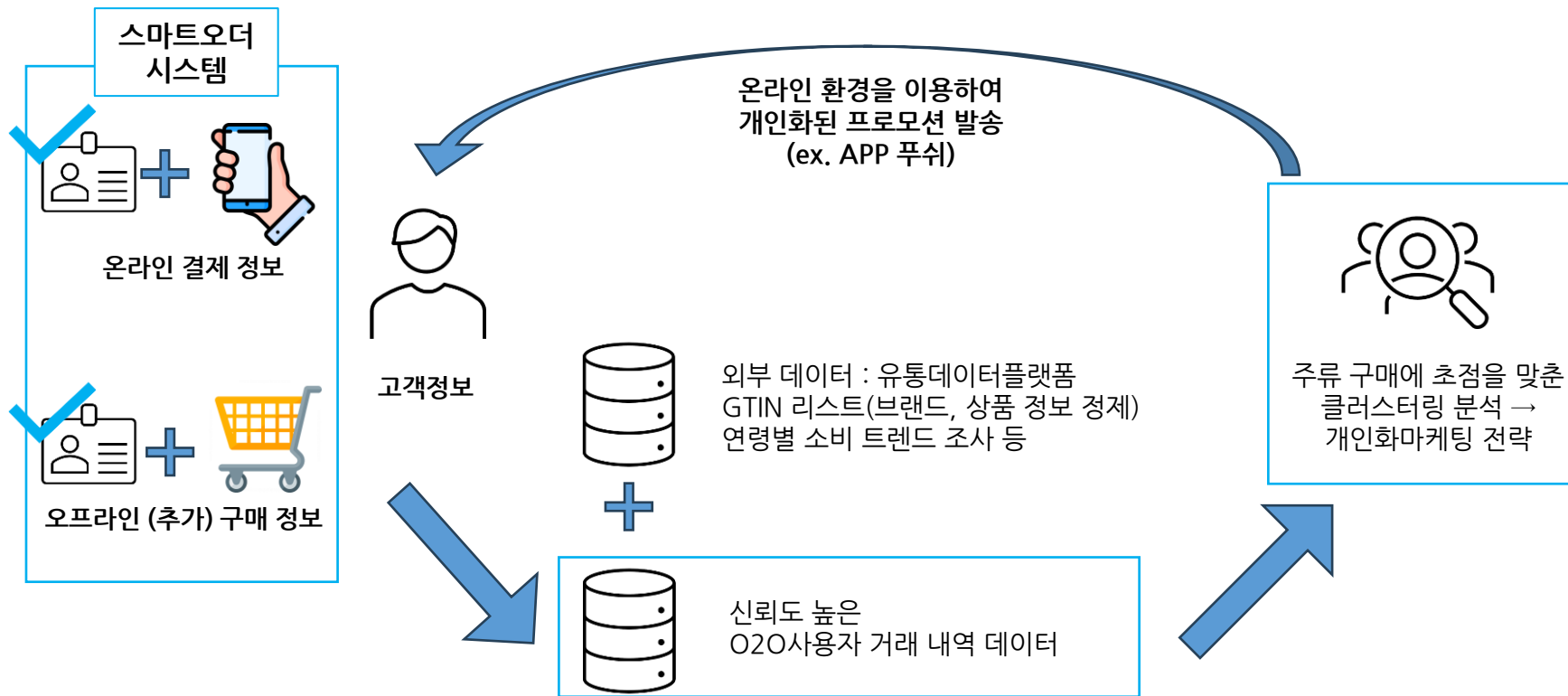


40대 중반 여성
김OO
(C4 해당)



분석 확장 방안

O2O 플랫폼을 활용하여, 데이터를 수집 및 활용하는 선순환 구조



Q&A



[참고] 주류 시장 동향

소주/맥주 소비 감소

수입 주류 및 전통주 등으로 주종의 다양화 및 구매량 증가

`17년 대비 `21년
국내 주류 출고액 변화

- 4.4%

지속적인 하락하여
`21년 약 8조8천억원

(희석소주, 맥주, 탁주,
과실주, 청주, 약주, 기타)

`17년 대비 `21년
주류 수입 비중

5.1 %

`21년 약 1조6천억으로 국내 주류
출고액 대비 15%로 지속적으로 증가

(과실주, 맥주, 위스키, 증류주,
기타)

`22년 대한민국
주류대상 출품작 수

812

`19년 540개 대비 1.5배 증가
이에 전통주 출고액 국내 주류의 1.1%

(와인, 전통주, 맥주, 소주, 기타)

[참고] 활용 데이터

구매 영수증 내역 Dataset

- 출처 : 공모전 제공 유통 판매 데이터 중 구매 영수증 데이터
- 데이터 특징
 - 카테고리 정보 부분 누락 외에는 결측치 없음
 - 118,810 개의 구매 내역
 - 구매일 : '21.01.01 ~ '23.06.30 (구매 시간은 없음)
 - 연령 : 4년 단위로 그룹핑 (23세 ~ 65세)
- 활용 및 분석 전제
 - 물건을 구매한 사람은 자신의 ID 로 구매했다.
 - 해당 기간 별도 필터링 없이 제공된 데이터다.

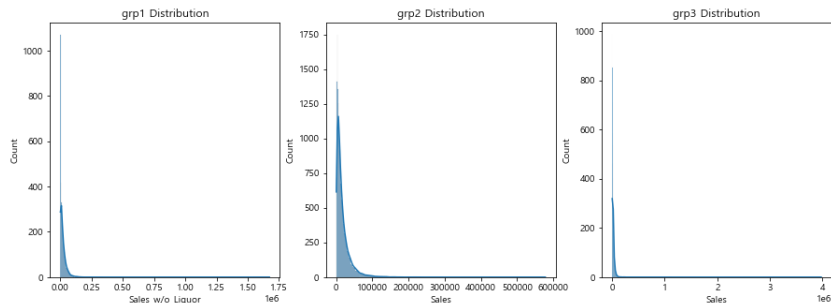
[참고] 정규성 검토

p.5 주류 구매 유형에 따른 구매액 분포

grp1 : 주류 구매자가 주류를 구매한 OrderId의 주류 외 구매액
grp2 : 주류 구매자가 주류를 구매하지 않은 OrderId의 구매액
grp3 : 주류를 한번도 구매하지 않은 OrderId의 구매액

1. One-Way ANOVA 수행

1) 로그 변환 적용으로 정규 분포화



2) Levene's Test 검정 결과, 등분산성 충족($p > .05$)

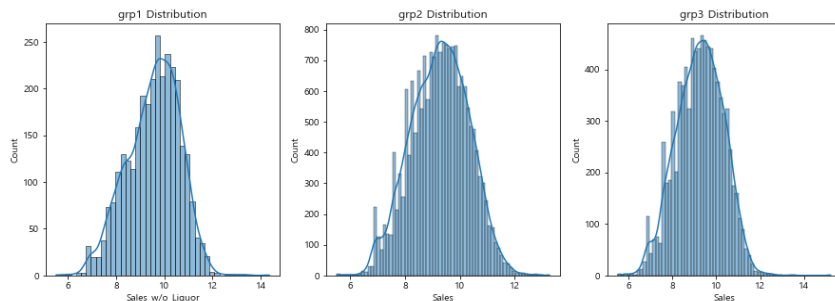
3) One-Way ANOVA 수행 [$F(2,30768)=101.817, p<.001$]

4) Bonferroni Correction 사후검정 수행

grp1 vs grp2: [$t(4270.02)=13.668, p<.001, \text{hedges' } g=0.267$]

grp1 vs grp3: [$t(5649.74)=12.542, p<.001, \text{hedges' } g=0.264$]

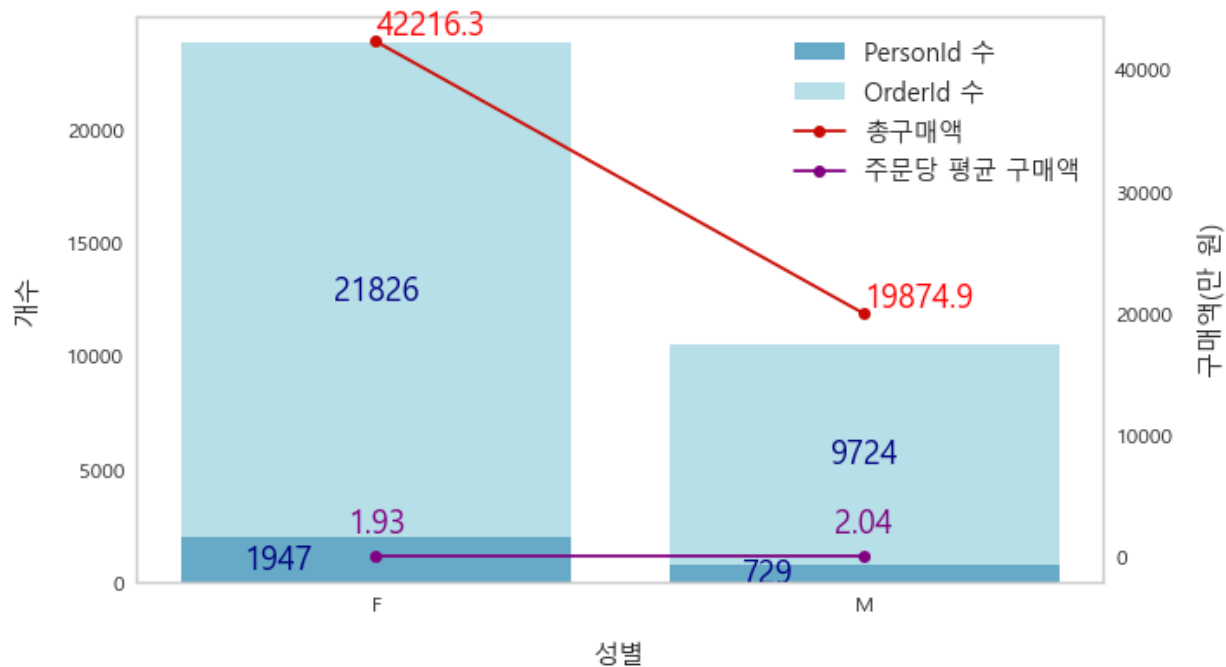
grp2 vs grp3: [$t(16002.77)=-0.100, p=.920$]



그림(좌),(우) : 로그 변환 적용 전후 정규 분포도

[참고] EDA_1

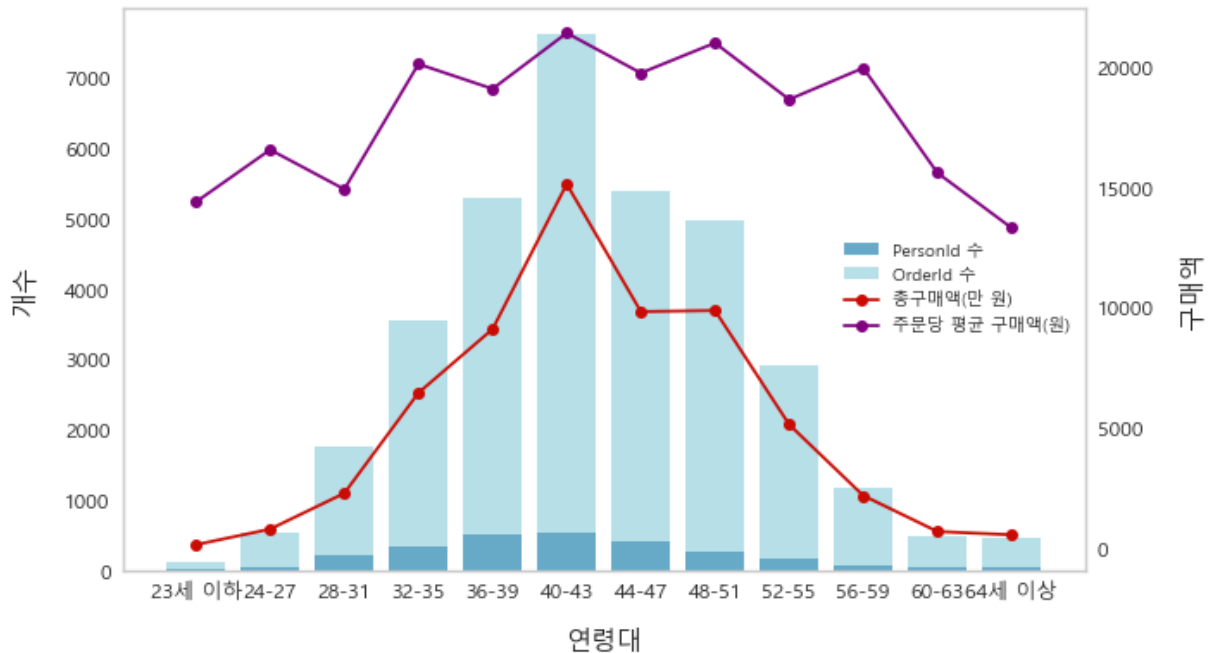
성별에 따른 구매 횟수와 금액 차이



- ✓ 구매 횟수 : 여 > 남
- ✓ 구매자수 : 여 > 남
- ✓ 주문당 평균 구매액 : 남 > 여
- ✓ 남성이 비교적 더 적은 횟수 구매하지만 구매시 많은 액수를 구매함

[참고] EDA_2

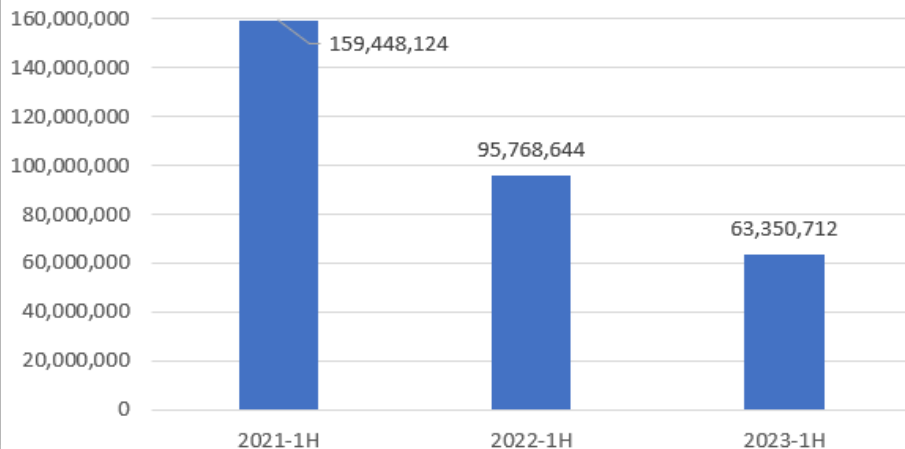
연령에 따른 구매 횟수와 금액 차이



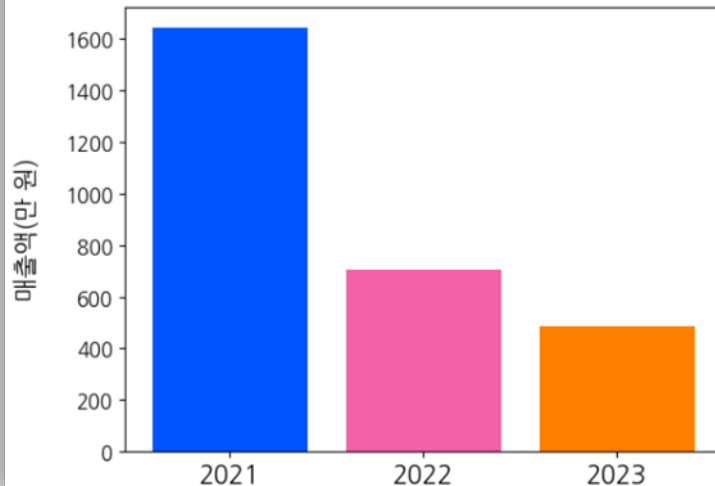
- ✓ 주문횟수 가장 많은 연령대 : 40~43세 그룹
- ✓ 주문당 평균 구매금액이 비교적 많은 연령대 : 32~56(32~59세)그룹

[참고] 기초분석-시기별 구매 경향

2021~2023 상반기 전체 매출액



연도별 1-2분기 주류 매출액(2021 ~ 2023)

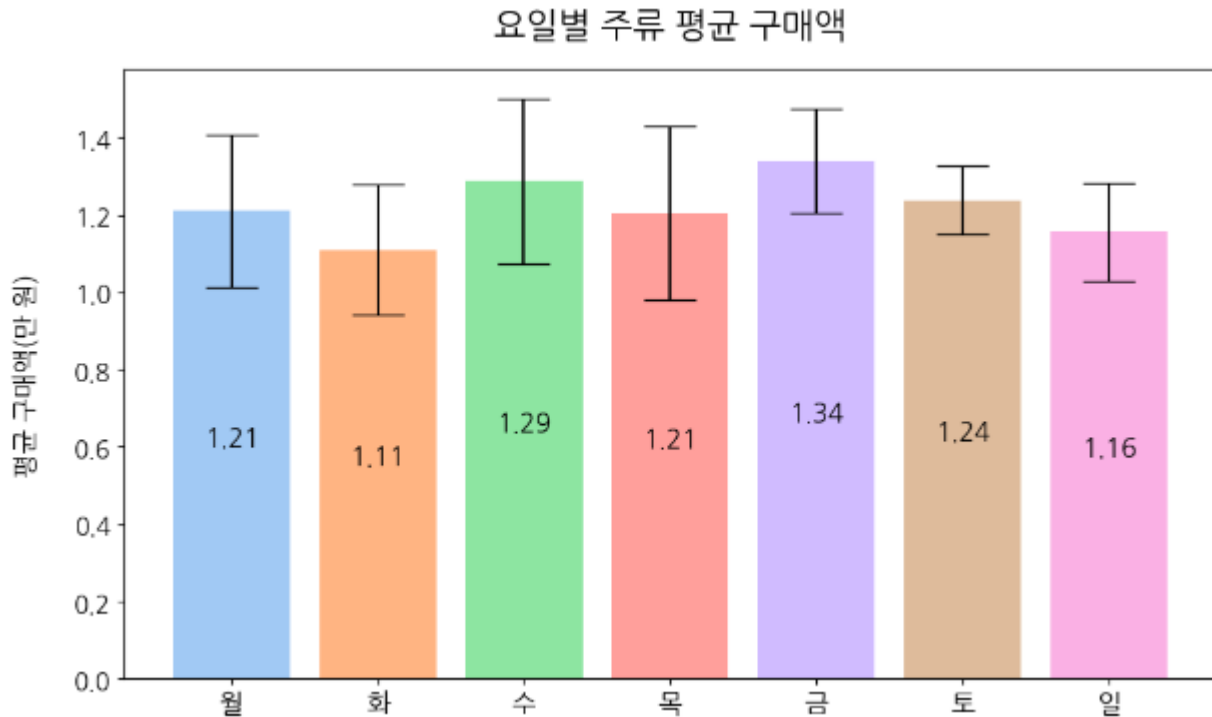


✓ 전체 구매액, 주류 구매액 모두 감소

✓ 2021년 상반기와 2022년 상반기 주류 구매액 감소폭이 전체 구매액의 감소폭에 비해 훨씬 큼

[참고] 기초분석-요일별 구매 경향

푸쉬 메시지 발송 알림 요일 제안 - 요일별 주류 구매액 비교



사람들이 주류에 가장 많은 돈을 쓰는 요일은 금요일이다.
(금 > 수 > 토 > 월 = 목 > 일 > 화)
*오차막대는 95% 신뢰구간을 나타냄

[참고] 전처리

데이터 분류 및 정제

1. 상품 카테고리 정리

- 카테고리명을 4레벨로 구분 및 누락 정보 수정
ex) 주류-주류 → 가공식품-주류-양조주-맥주

2. Retailer 구분 정보 삭제

- 유통업자를 A,B 정보 불 명확화

A ^B _C MasterCategoryFullNa 1	A ^B _C Retailer 2
주류-주류-소주-	A
주류-주류-맥주-	B
주류-주류-맥주-	A
주류-주류--	A
주류-주류-맥주-	A
주류-주류-수입맥주-	B
주류-주류-맥주-	A
주류-주류-맥주-	B
주류---	A
주류---	A
주류-주류-소주-	A
주류-주류-맥주-	B
주류-주류-수입맥주-	A
주류-주류-소주-	B

[클러스터 분석]

- PersonId를 기준으로 그룹화하여 각 사용자의 특성을 나타낼 수 있는 값들로 칼럼 구성
- Min-max 스케일링, T-SNE로 차원 축소
- K-means 클러스터링

Gender	성별
isYoung	40대 이상인지 아닌지 여부
LqAvgSales	주류 구매액
LqSalesRate	주류 세트 구매율
LqOrderRate	주류 구매율
LqPeriod	주류 구매 주기
LqVariety	구매 주종 개수
LqSetOrderRate	6캔 / 12캔 / 24입 / 5L 케그 등 대용량 & 세트 구매 여부

감사합니다.

