

• • •

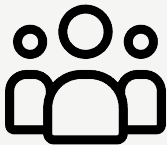
변형 RFM 재구매 비율을 이용한 제품 세분화

- 소속: EPOCH
- 팀명: 겸민이들
- 팀원: 김부겸, 박정민

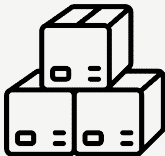
Table of contents



주제 선정이유



고객 중심



제품 중심

RFM



Recency

얼마나 최근에 구매했는가?



Frequency

얼마나 자주 구매했는가?



Monetary

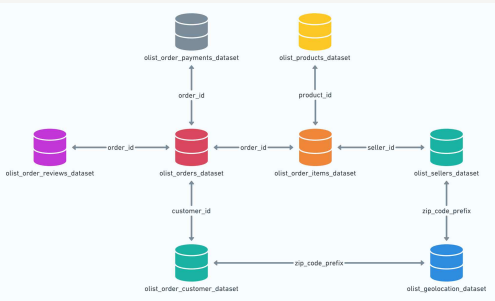
얼마나 돈을 썼는가?



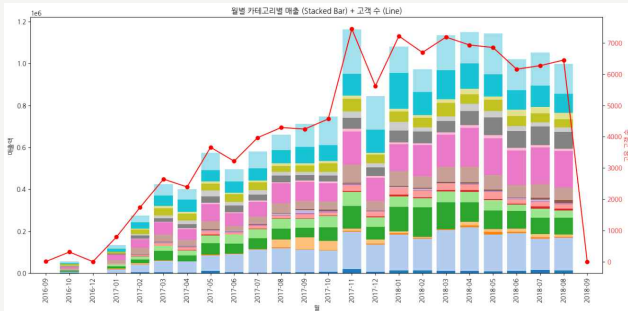
RFM 예시



Olist DataSet



Recency의 무의미



새로운 RFM

Repurchase_Ratio

재구매가 얼마나 이루어지는가

Frequency

얼마나 빈번하게 판매되었는가

Monetary

평균 매출액은 얼마인가

사용한 데이터

olist_orders_dataset	order_purchase_timestamp, order_delivered_customer_date, order_status, order_id, customer_id
olist_products_dataset	products_id, products_category_name
olist_customers_dataset	customer_id, customer_unique_id
olist_order_items_dataset	order_id, price, freight_value
olist_category_name_translation	product_category_name_english

전처리



배송 완료 건만 사용



가격에 운송비를 포함



결측치 제거

Table

ID

주문, 상품,
판매자, 고객

비용

상품 가격,
운송비, 합계

이름

카테고리명,
카테고리명_영어

날짜

배송일

유형

고객 유형
(신규/재구매)



RFM 지표 생성

Repurchase_Ratio

- 고객별 첫 구매일 계산
- 신규 고객과 재구매 고객 분리
- 재구매 고객수/전체 고객수

재구매 비율이 얼마나 되는가?

Frequency

- 카테고리별 전체 주문수

얼마나 빈번하게 판매되었는가?

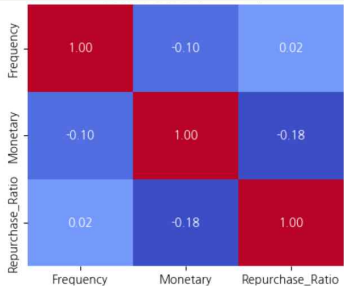
Monetary

- 카테고리별 평균 매출액

평균 매출액이 얼마인가?



PCA를 하지 않은 이유



주성분	설명 분산 비율
PC1	0.4055
PC2	0.3273
PC3	0.2671

K-Means

연속형 데이터



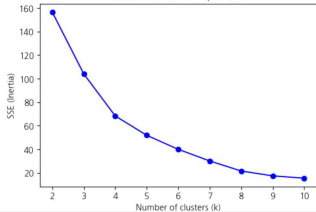
단순성



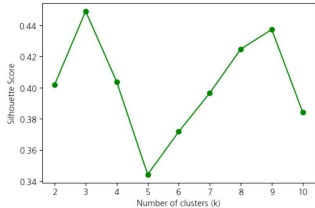
중심점 분석

K=3

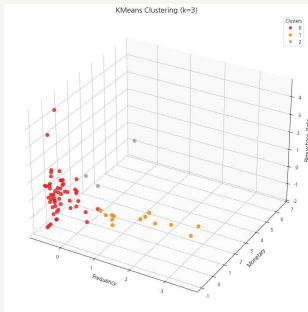
Elbow Method for Optimal k



Silhouette Scores for Different k



군집화 결과



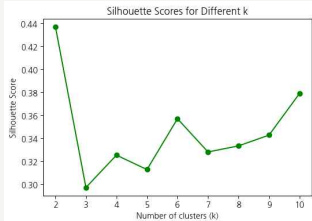
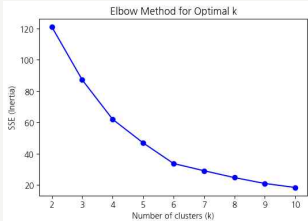
Cluster 1(13개)

- 재구매율(R) 중간
- 빈도(F) 높고
- 가격(M) 낮음

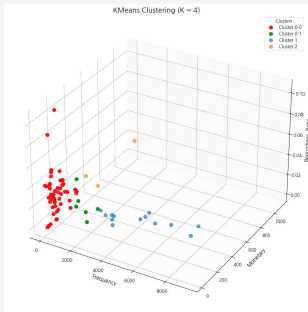
Cluster 2 (3개)

- 재구매율(R) 낮음
- 빈도(F) 낮음
- 가격(M) 높음

Cluster 0 재군집



군집화 결과



Cluster 0_0(49개)

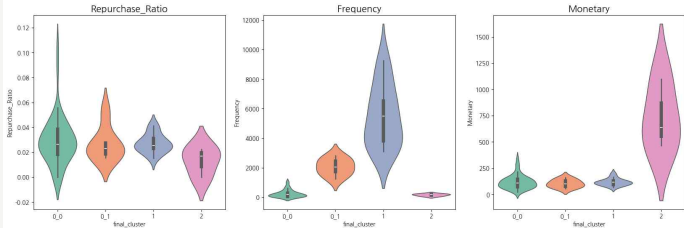
- 재구매율(R) 다양
- 빈도(F) 낮음
- 가격(M) 낮음

Cluster 0_1 (6개)

- 재구매율(R) 중간
- 빈도(F) 중간
- 가격(M) 낮음

Cluster 특징

최종 4개 클러스터별 지표 분포 (Violin Plot)



대표 Category

Cluster 0_0

- Book
- Fashion
- Flowers
- Food

Cluster 1

- Computers_accessories
- Garden_tools
- Perfumery
- Sports_leisure

Cluster 0_1

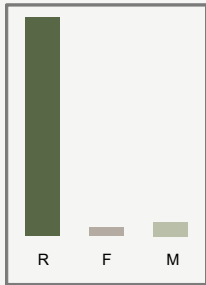
- Baby
- Office_Furniture
- Stationery
- Pet_Shop

Cluster 2

- Computer
- Home_appliances
- Small_appliances_home_oven_and_coffee



Cluster 0_0



롱테일 구조

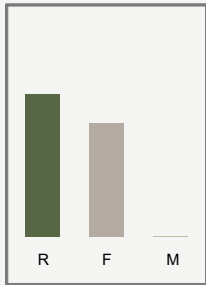
- 재구매율 다양
- 빈도 낮음
- 금액 낮음

'잘 안팔리는 상품'이 아닌 '다양성과 경쟁성을 완성시키는 전략적인 자산'

맞춤 전략

- 오래된 재고 할인
- 연관 검색어 노출

Cluster 0_1



안정적인 매출

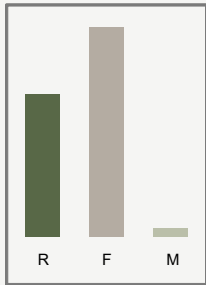
- 재구매율 중간
- 빈도 중간
- 금액 매우 낮음

상품을 꾸준히 찾지만, 여러 플랫폼 이용 가능성이 높음

맞춤 전략

- 재구매 알림
- 주기적인 쿠폰 제공

Cluster 1



트래필 유발 상품

- 재구매율 중간
- 빈도 높음
- 금액 낮음

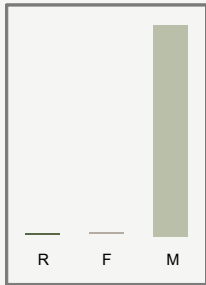
가장 빈도가 높은 대중적인 인기 상품군

하지만, 가격이 낮아 수익성 떨어짐

맞춤 전략

- 객단가 기반 프로모션
- 지능적인 교차 판매

Cluster 2



고가 상품

- 재구매율 매우 낮음
- 빈도 매우 낮음
- 금액 높음

가격이 높은 상품으로 이루어져, 일회성 구매

맞춤 전략

- 묶음 구매 할인
- 할부 옵션 강조

결론

기대 효과

기존 RFM 대비
제품 분석에 용이

카테고리별 맞춤 전략으로
매출/유지율 개선

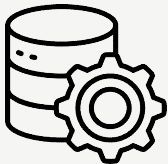
한계

절대적인 성능 지표 부재

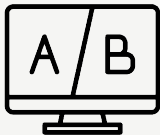
전략적 제안 차원에서 그침



추후 개선점



데이터 보완



A/B 테스트 진행



인사이트 도출

...

Q&A

...



감사합니다