

베이커리 구매 데이터 기반

마케팅 전략 제안

구매량 시계열 데이터 분석 및 장바구니 분석을 중점으로

5팀 문상혁

INDEX

01 목적과 목표

02 데이터 탐색

데이터 확인
데이터 전처리

03 데이터 분석

구매량, 시계열 데이터 분석
장바구니 분석

04 결론

마케팅 전략 제안



목적

- (1) 베이커리 매출 증대
- (2) 고객 만족도 상승

목표

1. 상품 구매량과 시계열 데이터를 분석하여 인사이트를 도출한다.
2. 장바구니 분석을 통해 연관성이 높은 상품 조합을 선별한다.
3. 선별한 상품 조합 별로 인사이트를 도출한다.
4. 선별한 상품 조합 및 인사이트를 기반으로 마케팅 전략을 제안한다.



- ✓ 데이터는 총 20,507개로, 거래ID, 상품명, 거래 일시, 하루 중 구매 시기, 주중/주말 구매 여부를 설명한다.
- ✓ 결측치는 존재하지 않았고, 모든 컬럼이 수치형 변수가 아닌 명목형 변수이므로 이상값을 처리하지 않았다.
- ✓ 동일한 거래ID(Transaction)에 동일한 상품(Item)이 있는 경우는 **한번 결제할 때 동일한 상품을 여러개 구매한 경우로 간주**하여, 상품 구매 수량을 나타내는 **구매량(Quantity)** 컬럼을 추가하는 방식으로 중복값을 처리했다.

데이터 설명

데이터	선정 이유	비고
Transaction	거래 ID	-
Item	구매한 제품명	-
date_time	거래 날짜 및 시각	-
period_day	아침/오후/저녁/밤 중 언제 구매했는지 여부	-
weekday_weekend	주중/주말 중 언제 구매했는지 여부	-
Quantity	상품 구매량	Transaction과 Item으로 그룹핑한 후, 데이터를 count하여 추가 생성한 데이터

표 1. 데이터 컬럼별 설명

구매량(Quantity) TOP 5 상품을 파악하고, 해당 상품들의 시간대별 구매량 변화 추이를 분석한다.

구매량 TOP 5 상품

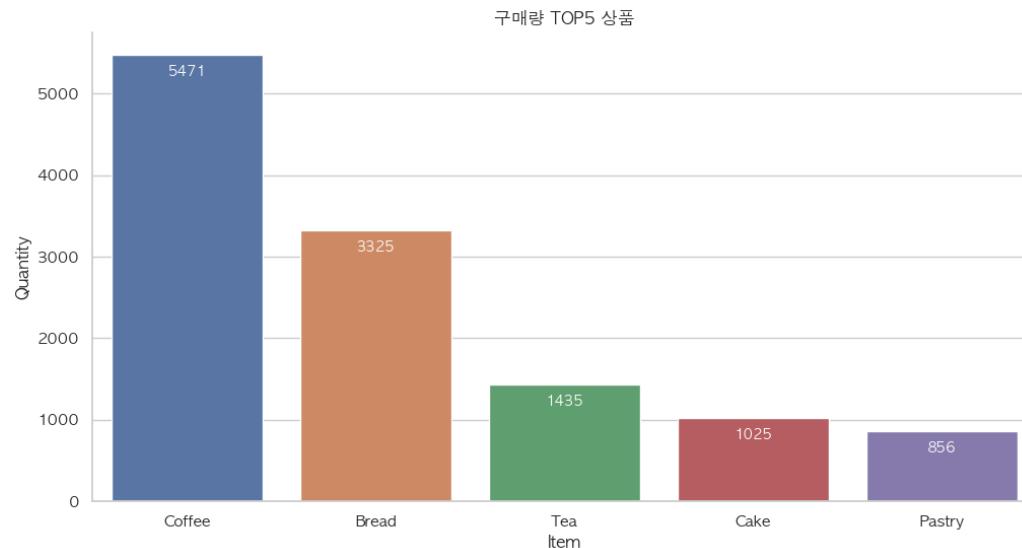


차트 1. 구매량 TOP 5 상품 구매량 막대 그래프

시간대별 상품 구매량 변화 추이

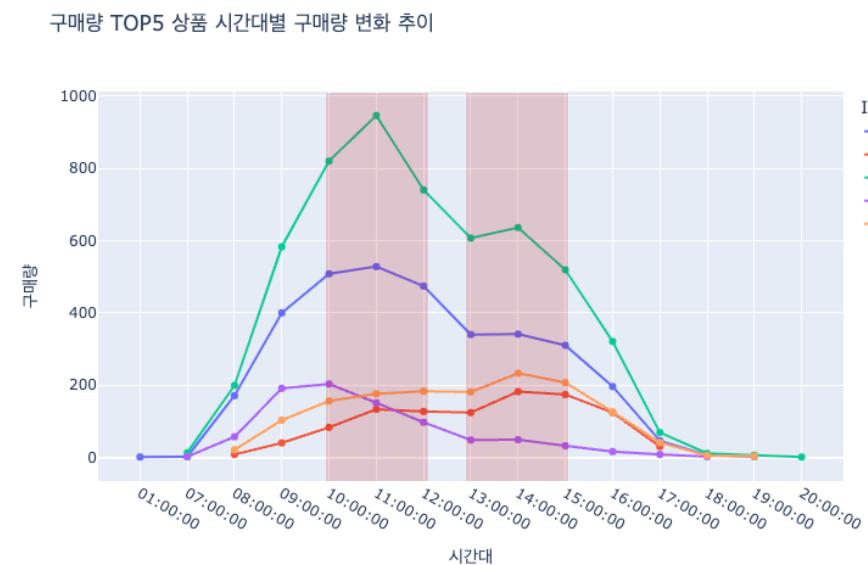


차트 2. 구매량 TOP 5 상품 구매량 막대 그래프

결과 해석

- ✓ 커피의 구매량이 압도적으로 높다.
- ✓ 베이커리에서 식사류도 판매하고 있지만 구매량이 디저트와 음료에 비해 높지 않다.

결과 해석

- ✓ 커피, 빵, 패스트리는 오전에 구매량이 높고 점차 감소한다.
- ✓ 케이크, 차는 오후보다 오후에 구매량이 높다.
- ✓ 커피는 오전과 오후에 판매량이 증가했다 감소하는 구간이 한번씩 존재한다.
- ✓ 모든 상품에서 16시 이후에 구매가 급감한다.

고객 ID가 존재하지 않으므로 prefix 알고리즘 사용 불가하므로 연관 규칙 마이닝 알고리즘 중 성능이 뛰어난 **FP-growth 알고리즘**을 사용한다.

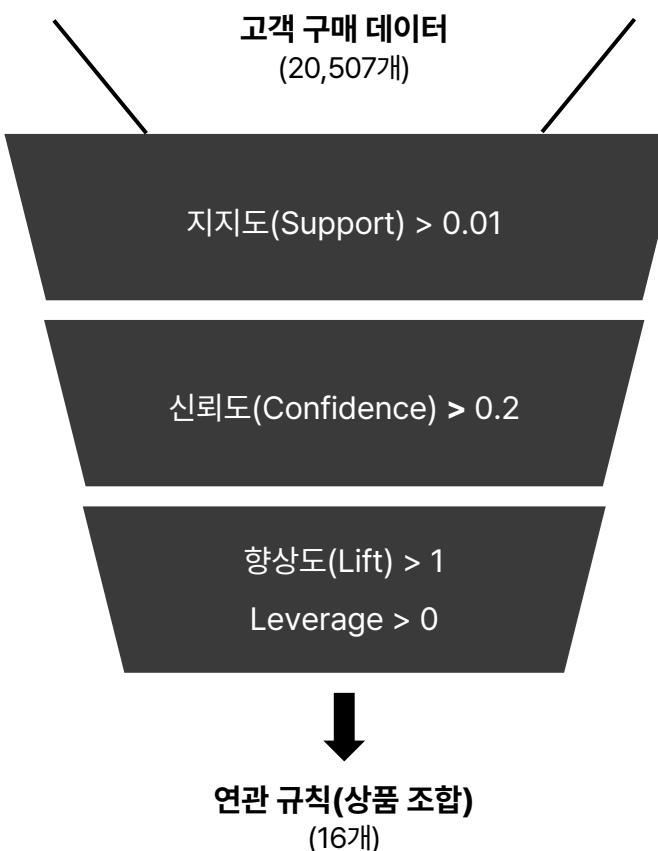
알고리즘 적용을 위한 전처리

- ✓ 결제 건별로 데이터 그룹핑
- ✓ One-Hot 인코딩을 통해 각 상품명 컬럼화

결제 ID	결제 ID
coffee	coffee
bread	pastry
cake	
결제 ID	결제 ID
tea	sandwich
cake	tea
coffee	cake
spanish brunch	

그림 1. 그룹핑 예시

알고리즘 적용하여 규칙 추출



- ✓ 최소 지지도를 0.06으로 설정했을 때
7개의 빈발항목 밖에 남지 않아 **0.01**로 조정
- ✓ 최소 신뢰도를 0.5로 엄격하게 설정했더니
모든 consequents가 Coffee로 나왔다.
- ✓ 더 다양한 제품들의 연관성을 발견하기 위해
최소 신뢰도를 0.2로 낮춰서 진행했다.
- ✓ **신뢰도를 낮추는 것을 보완하기** 위해 지표 추가 활용
 - lift > 1: 두 항목이 서로 독립적이지 않고 실제로 연관이 있음을 의미
 - leverage > 0: 발생 빈도가 우연의 일치보다 높음을 의미

그림 2. 알고리즘 규칙 시작화

조건을 만족하는 규칙(16개)을 향상도와 신뢰도를 기준으로 내림차순 정렬하여 상위 규칙 5개 산출

장바구니 분석 결과 – 조건을 만족하는 상위 규칙

Antecedents	Consequents	support	confidence	lift	leverage
Tea, Coffee	→ Cake	0.01	0.20	1.94	0.004
Cake	→ Tea	0.02	0.23	1.60	0.008
Toast	→ Coffee	0.02	0.70	1.47	0.008
Sandwich	→ Tea	0.01	0.20	1.40	0.004
Spanish Brunch	→ Coffee	0.01	0.60	1.25	0.002
선별된 나머지 규칙 중 하나를 제외하고 모두 Consequents가 Coffee였다.					

표 2. 장바구니 분석 결과표(조건을 만족하는 상위 규칙)

결과 해석

- ✓ 선별된 16개의 규칙에서 대부분의 베이커리 상품과 '커피'가 연관되어 있다는 사실을 알 수 있다.
- ✓ 디저트 중에서 '케이크'가 가장 '커피'와 연관성이 높다.
- ✓ 간편식 중에서 '샌드위치'가 가장 '차'와 연관성이 높다.
- ✓ 간편식 중에서 '토스트'가 가장 '커피'와 연관성이 높다.
- ✓ 식사류 중에서 '스페인식 브런치'와 '커피'와 연관성이 높다.



[식사 - 디저트 - 음료] 세트 상품 출시

- 스페인식 브런치와 케이크 모두 커피와 연관성이 높으므로 [스페인식 브런치 – 케이크 – 커피] 세트 상품 출시
 - 스페인식 브런치, 케이크, 커피 모두 오후(afternoon)에 판매가 많이 되므로 오후 시간에 시간 한정 할인 이벤트 진행
- 기대 효과 :** 구매 연관성이 높은 상품끼리 묶여 세트 상품이 출시되므로 매출 증대와 더불어 고객의 만족도도 상승할 것



베이커리 구매 시 커피 할인 프로모션 진행

- TOP5 외에도 기준에 의해 선별된 규칙을 분석해보면 대부분의 베이커리 상품이 커피와 연관성이 높음
- 커피, 빵, 패스트리는 오전에 구매량이 높으므로 오전에 프로모션 진행

기대 효과 : 베이커리 메뉴와 커피의 매출 증대



케이크 혹은 샌드위치 2개 이상 구매 시 티백 증정 이벤트

- 케이크, 샌드위치 모두 차와 연관성이 높음
- 케이크 혹은 샌드위치를 구매할 때, 판매하는 차 티백을 증정하는 이벤트를 진행

기대 효과 : 케이크, 샌드위치의 매출이 증대되고 고객 만족도가 상승할 것



런치 디저트 세트 출시 및 프로모션(케이크 & 차)

- 케이크와 차는 연관성도 높고 둘 다 점심 직후인 13시부터 구매량이 증가
- 점심 직후 시간대를 타겟하여 13시~15시에 런치 디저트 세트 프로모션 진행

기대 효과 : 케이크와 차 매출 증대 및 차 상품에 대한 추가적인 바이럴 효과(티백은 다른 사람에게 주기 편하기 때문)

END

감사합니다.

5팀 문상혁