



메뉴 최적화를 위한 홍차 전문 카페 데이터 분석

Tags	1인 프로젝트 Excel MySQL Project Python
Date	@February 24, 2023 → March 30, 2023
Status	Done

이 프로젝트에 있는 모든 분석 내용에 대한 코드는 [GitHub](#)에서 확인할 수 있습니다.

1. 기획 의도 및 분석 목적

기획 의도

2021년부터 부모님께서 운영하시는 홍차 전문 카페에는 메뉴와 관련하여 두 가지 문제가 존재했다.

1. 메뉴에 제품을 추가하거나 제거하는 것부터 새로운 제품을 개발하는 것까지 메뉴와 관련된 의사 결정이 감으로 이루어졌다.
2. 메뉴가 많고 다양해서 직원들이 모든 제품의 레시피를 완전히 파악하고 일관된 품질로 제조하는 데 어려움을 겪고 있었다.

이러한 문제들을 해결하기 위해 POS 시스템에서 수집된 1년 치 매출 데이터를 분석하여 **모든 제품의 인기도를 확인**하고, 이를 토대로 **메뉴를 최적화**하는 프로젝트를 진행했다.

분석 목적

- 상대적으로 판매량이 부진한 **비인기 제품**을 식별하고 **제거**하여 메뉴 구성 조정
- 고객의 선호도를 파악하여 **새로운 제품**을 개발할 때 **참고** 자료로 활용
- 판매가 중단됐지만 상대적으로 인기가 있었던 제품은 **재출시**하여 메뉴 구성 조정

분석 결과 예상 (가설)

1. 홍차 전문 카페이기 때문에 홍차가 들어가지 않는 제품은 인기가 없을 것이다.
2. 여름에는 차가운 음료가 인기 있고, 겨울에는 따뜻한 음료가 인기 있을 것이다.
3. 메뉴판 상단에 위치한 제품들은 접근이 쉬워서 판매량이 많을 것이다.
4. 구매를 유도하는 추천 문구가 달린 제품들은 판매량이 많을 것이다.
5. 판매가 중단되어 메뉴에서 제거된 제품들은 실제로 인기가 없었을 것이다.

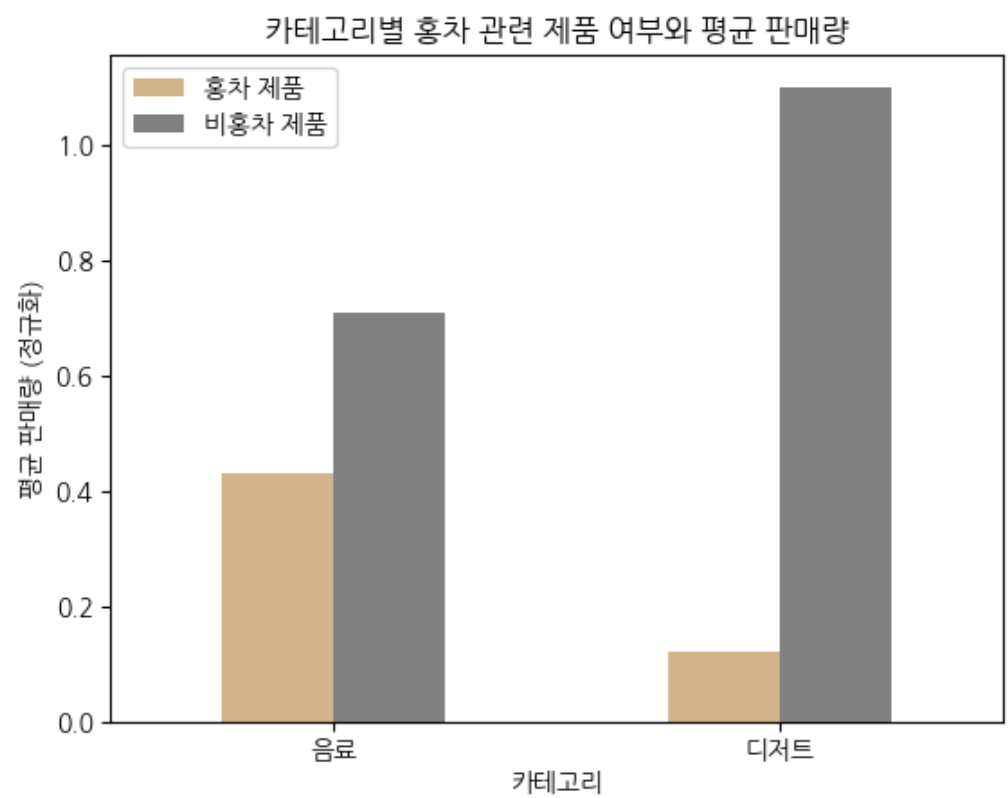
2. 결과

모든 제품이 데이터 수집 기간의 첫날부터 판매되지 않았고 중간에 판매가 중단되거나 새로 출시된 제품도 있었음. 제품의 출시 시기나 판매 기간에 따른 판매량 차이를 보정하기 위해 출시일(`launched_on`)과 단종일(`discontinued_on`) 데이터를 사용하여 **정규화**를 수행함. 각 제품의 판매량을 판매 기간으로 나누어 비교가 가능한 **상대적인 판매량**을 계산함. 분석 결과에서 정규화된 판매량은 별도로 표기함.

모든 제품의 이름은 카페를 특정할 수 없도록 **별명을 사용**했고 제품의 **인기**도는 **판매량을 기준으로 판단**함.

1. 홍차 관련 제품 여부와 판매량 분석

가설: 홍차 전문 카페이기 때문에 홍차가 들어가지 않는 제품은 인기가 없을 것이다.



카테고리	홍차 제품 판매량 (정규화)	비홍차 제품 판매량 (정규화)
음료	0.43	0.71
디저트	0.12	1.10

- 카테고리를 새롭게 분류한 기준: 음료(밀크티, 음료, 차, 커피, 홍차) & 디저트(디저트)

홍차 관련 음료, 디저트의 평균 판매량이 그렇지 않은 음료, 디저트의 평균 판매량보다 약 **1.651배**, **9.167배** 적음.

홍차 전문 카페라고 해서 고객이 홍차 관련 제품을 더 선호하는 것은 아니라고 할 수 있음.

전체 제품 판매량 상위 TOP 10

제품명	제품 온도	총판매량*
밀크티 A	COLD	5265
커피 A	COLD	4364
커피 A	HOT	2876
밀크티 A	HOT	1534
디저트 S	N/A	1365
디저트 CB	N/A	1317
커피 Y	COLD	324
밀크티 B	COLD	964
디저트 CA**	N/A	820
커피 C	HOT	773

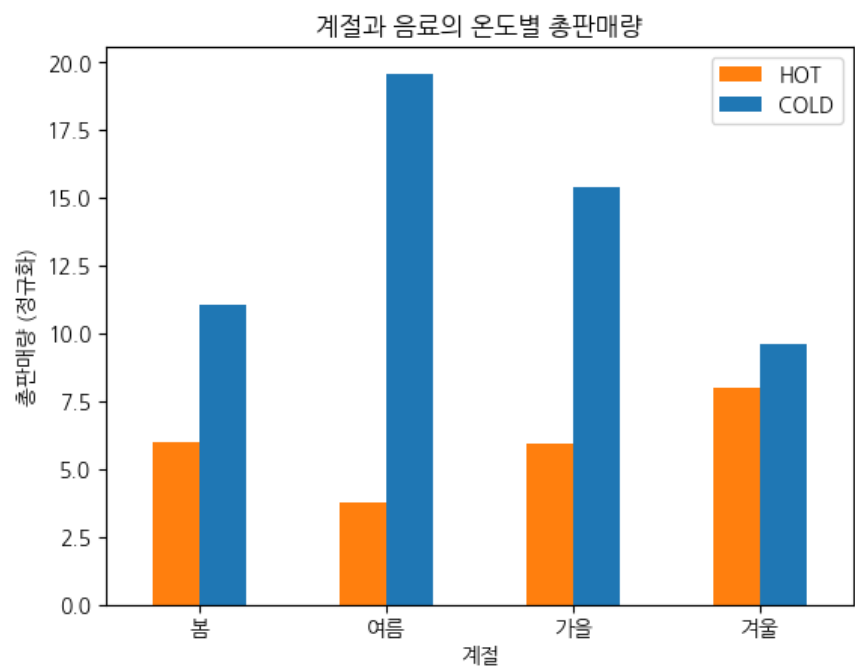
* 정규화된 판매량을 기준으로 정렬했으나 실제 판매량을 조회함.

** 디저트 CA는 홍차맛 디저트

카테고리 구분 없는 전체 제품 판매량 상위 10개 중 6개가 홍차가 들어가지 않은 제품임.

2. 계절과 음료의 온도별 판매량 분석

가설: 여름에는 차가운 음료가 인기 있고, 겨울에는 따뜻한 음료가 인기 있을 것이다.



계절	HOT 총판매량 (정규화)	COLD 총판매량 (정규화)
봄	6.01	11.05
여름	3.75	19.59
가을	5.93	15.39
겨울	7.98	9.64

- 4개의 계절로 분류한 기준: 봄(3, 4, 5월), 여름(6, 7, 8월), 가을(9, 10, 11월), 겨울(1, 2, 12월)

모든 계절에 걸쳐 차가운 음료의 총판매량이 따뜻한 음료의 총판매량보다 많음. 고객은 **계절과 상관없이 주로 차가운 음료를 선호**하지만, 겨울에는 비슷한 비율로 음료를 선택하는 경향이 있음.

겨울 음료 판매량 상위 TOP 10

제품명	제품 온도	총판매량*
밀크티 A	COLD	948
커피 A	HOT	889
커피 A	COLD	683
밀크티 A	HOT	592
커피 Y	COLD	103
커피 C	HOT	246
밀크티 B	COLD	166
홍차 A	HOT	151
밀크티 B	HOT	123
커피 C	COLD	103

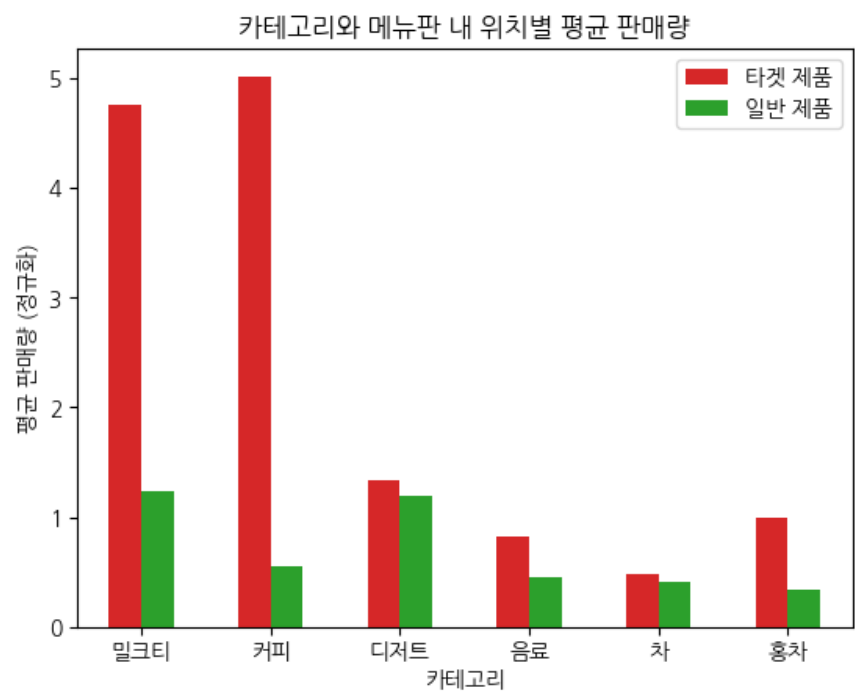
* 정규화된 판매량을 기준으로 정렬했으나 실제 판매량을 조회함.

겨울 음료 판매량 상위 10개중 5개의 제품이 차가운 음료(COLD)임.

3. 메뉴판 내 위치와 판매량 분석

가설: 메뉴판 상단에 위치한 제품들은 접근이 쉬워서 판매량이 많을 것이다.

메뉴 디자인 심리학에 따르면, 고객의 시선은 **Golden Triangle**의 형태, 메뉴판의 **정중앙 ▶ 우측 상단 ▶ 좌측 상단** 순으로 이동한다고 함.



카테고리	타겟 제품* 판매량 (정규화)	일반 제품 판매량 (정규화)
밀크티	4.76	1.24
커피	5.02	0.55
디저트	1.33	1.20
음료	0.83	0.45
차	0.48	0.41
홍차	1.00	0.34

* 타겟 제품은 메뉴판 상단에 위치한 제품을 의미함.

- 고객의 키오스크, 카운터 결제 여부를 확인할 수 없어서 전체 데이터를 추출 및 가공하고 시각화를 진행함.
- 키오스크와 카운터 메뉴판 둘 다 정중앙의 개념이 없어서 좌측과 우측 상단의 위치한 제품들만 분석함.

키오스크와 카운터 메뉴판의 좌측과 우측 상단에 위치한 제품들이 그렇지 않은 제품들보다 **평균 판매량이 많음**.

특히 디저트, 음료, 차, 홍차 카테고리처럼 뚜렷한 인기 제품이 없는 카테고리도 덜 분명하지만, **동일한 패턴이 관찰됨**.

4. 추천 문구(키워드)가 달린 제품 판매량 분석

가설: 구매를 유도하는 추천 문구가 달린 제품들은 판매량이 많을 것이다.

	추천 제품	일반 제품
평균 판매량 (정규화)	4.28	0.99

키오스크와 카운터 메뉴판에 추천 문구가 달린 제품들(밀크티 A, 디저트 PA, 커피 Y)이 그렇지 않은 제품들보다 평균 판매량이 약 **4.323배 많음**.

BEST, 시그니처 등의 추천 문구가 고객의 구매를 유도하고 있을 수 있음.

5. 단종된 제품 판매량 분석

가설: 판매가 중단되어 메뉴에서 제거된 제품들은 실제로 인기가 없었을 것이다.

	단종된 제품	판매 제품
평균 판매량 (정규화)	0.82	0.24

데이터 수집 기간에 판매가 중단된 제품 26개가 현재 판매되고 있는 비인기 제품 26개(판매량 하위 제품)보다 평균 판매량이 약 **3.417배 많음**.

메뉴에서 제거된 제품들이 실제로 인기가 없었던 것이 아니라고 할 수 있음.

3. 제안

1. 카테고리 당 7개의 메뉴 아이템만 남기고 비인기 제품을 제거하기

메뉴의 카테고리 하나에 들어가는 최적의 제품 개수는 통상적으로 7개라고 한다. 카테고리에 제품이 너무 많으면 메뉴가 어수선하고 혼란스러워 보일 수 있으며, 고객이 메뉴를 결정하고 주문하는 데 어려움을 느낄 수 있기 때문이다.

6개의 카테고리 중 3개의 카테고리(디저트, 음료, 커피)는 7개가 넘게 제품을 포함하고 있다.

디저트 카테고리는 24개, 음료 카테고리는 10개, 커피 카테고리는 12개의 제품이 속해있다.

정규화된 판매량이 적은 순서대로 제품을 메뉴에서 제거하여 카테고리당 7개의 제품을 유지하는 것을 제안한다.

디저트 카테고리 제품 판매량 하위 TOP 14

제품명	총판매량*
디저트 CP	156
디저트 CC	198
디저트 N	267
디저트 G	348
디저트 PB	349
디저트 KC	257
디저트 E	67
디저트 SB	84
디저트 CE	436
디저트 RC	87
디저트 KR	307
디저트 CC	468
디저트 D	103
디저트 CF	527
디저트 RV	112
디저트 CD	582
디저트 KB	366

* 정규화된 판매량을 기준으로 정렬했으나 실제 판매량을 조회함.

- 판매가 중단되지 않은 제품들의 전체 판매량의 합계를 계산함 (음료인 경우 HOT, COLD 합산).

음료 카테고리 제품 판매량 하위 TOP 3

제품명	총판매량*
음료 T	66
음료 P	83
음료 S	155

커피 카테고리 제품 판매량 하위 TOP 5

제품명	총판매량*
커피 CL	156
커피 T	198
커피 M	267
커피 R	348
커피 W	349

2. 차 카테고리 제품을 집중적으로 개발하고 단종된 제품을 개선해서 재출시하기

모든 계절에 걸쳐 6개의 카테고리 중 차 카테고리가 가장 판매량이 저조하다. 이는 차가운 음료가 계절과 무관하게 잘 팔리는 반면, 차 카테고리에 속한 6개 제품 중 오직 2개만이 COLD 옵션을 제공하기 때문으로 추정된다.

따라서 차 카테고리의 모든 제품에 대해 COLD 옵션을 추가하거나 혹은 HOT, COLD 옵션을 모두 제공하는 새로운 차 제품을 개발해서 카페인 부담이 적고 건강한 제품을 원하는 고객의 니즈를 충족시키는 것을 제안한다.

또한, 판매가 중단된 제품들이 현재 판매되고 있는 비인기 제품들보다 인기가 더 좋았다는 점을 고려할 때, 전체 판매량 상위에 위치한 단종 제품들을 제품 판매량 하위에 위치한 제품들과 교체하는 것을 제안한다.

계절별 비인기 카테고리

계절	카테고리	총판매량*
봄	차	249
여름	차	272
가을	차	255
겨울	차	196

* 정규화된 판매량을 기준으로 정렬했으나 실제 판매량을 조회함.

전체 단종 제품 판매량 상위 TOP 3

제품명	총판매량*
디저트 M	563
디저트 F	437
디저트 W	186

3. 추천 문구(키워드)를 추가하고 메뉴판 내 위치 변경하기

분석 결과에서 키워드가 달린 제품들이 그렇지 않은 제품들보다 평균 판매량이 많고 키오스크와 카운터 메뉴판의 좌측과 우측 상단에 위치한 제품들이 그렇지 않은 제품들보다 평균 판매량이 많다는 걸 알 수 있었다.

현재 단 세 가지의 제품에만 BEST, 시그니처 등의 키워드가 사용되고 있는데 그중 하나인 커피 Y는 2023년 9월 20일에 출시되었음에도 불구하고 판매량이 기존 제품들보다 월등하게 좋다.

그러나 나머지 두 제품은 이미 인기가 많았기 때문에 키워드가 있어서 혹은 메뉴판 내 위치가 우수하여 판매량이 많은 것인지, 아니면 그냥 원래부터 인기가 많아서 그런 것인지 단정 지을 수 없다.

따라서 카테고리마다 **판매량을 늘리고자 하는 하나의 제품을 선택하여 해당 제품에 키워드를 추가하고 메뉴판 내 위치를 좌측 혹은 우측 상단으로 변경한 후 구매를 유도**하되 변경된 사항이 키워드와 메뉴판 내 위치뿐인지, 그리고 실제로 판매량에 변화를 불러오는지 확인해야 한다.

판매량 추이를 계속 관찰하고 키워드 사용 횟수를 조정하여 최적의 키워드 개수를 찾는 것도 중요하다.

추천하는 키워드 리스트

- 가장 잘 팔리는 제품(BEST, BEST SELLER)
- 새로 출시된 제품(NEW)
- 식물성 재료만으로 만든 디저트(VEGAN)
- 스무디류 제품(SUMMER ONLY) 등

전체 제품 판매량 상위 TOP 10

제품명	총판매량*
커피 A	7240
밀크티 A	6799
커피 C	1457
디저트 S	1365
디저트 CB	1317
밀크티 B	1287
커피 Y	324
홍차 A	919
디저트 CA	820
디저트 PA	733

* 정규화된 판매량을 기준으로 정렬했으나 실제 판매량을 조회함.

- 판매가 중단되지 않은 제품들의 전체 판매량의 합계를 계산함 (음료인 경우 HOT, COLD 합산).

추천 문구가 달린 제품들은 전체 제품 판매량 상위 10개 중 2위 (밀크티 A), 7위 (커피 Y), 10위 (디저트 PA)다.

4. 데이터 설명

데이터 수집 기간

2023년 1월 1일 - 12월 31일

데이터 수집 방법

모든 매출 관련 데이터는 우노스페이 데이터를 xls로 추출 후 사용

`product_sales` 데이터 셋은 직접 엑셀 스프레드시트에 입력

데이터 수집의 한계

주문/결제별 구매 항목 알 수 없음

- 고객들이 구매하는 제품들의 조합 파악 불가능

제품 판매의 시간대 알 수 없음

- 각 제품의 시간대별 판매량 분석 불가능

일일 매출 데이터: `daily_sales`

속성	설명	예시
date_id	날짜 고유 식별자	1
date	날짜	1/1/23
day_id	요일 고유 식별자 - 일, 월, 화, 수, 목, 금, 토 순 서	1
total_sales	매출 합계(₩)	1,000,000
order_count	결제 건수	15
average_order_value	영수 단가 (₩)	10,000

총 데이터 수: 365개 데이터

- `date`의 포맷을 1/1/23에서 2023-01-01로 변경 후 자료형도 변경

제품 데이터: `product`

속성	설명	예시
product_id	제품 고유 식별자	1
name	제품명	아메리카노
price	해당 제품의 가격 (₩)	6,000
category	해당 제품이 속한 카테고리 - 디저트, 밀크티, 음료, 차, 커피, 홍차	커피
temperature	해당 제품의 온도 - HOT, ICE	HOT
launched_on	출시일 - 제품의 첫 판매일	1/2/23
discontinued_on	단종일 - 제품의 마지막 판매일	12/31/23

총 데이터 수: 116개 데이터

- (메뉴판 상) 초코 라떼는 (키오스크 상) 아이스 초코와 핫 초코로 구분
- 쇼케이스에서 병이나 캔 형태로 판매되는 밀크티, 홍차 제외
- ICE는 COLD로 수정하고 디저트 카테고리 제품의 NULL 값은 N/A 입력
- 수집 기간 전부터 판매된 제품들은 첫 영업일인 1월 2일을 출시일로 처리
- 수집 기간 이후에 판매된 제품들은 12월 31일을 단종일로 처리

제품 판매 데이터: **product_sales**

속성	설명	예시
date_id	날짜 고유 식별자	1
product_id	제품 고유 식별자	1
quantity_sold	해당 제품의 일일 판매량	10

총 데이터 수: 12,546개 데이터

- 음료 추가 옵션 (에스프레소 샷, 시럽, 원두 종류, 휘핑크림) 제외
- 진열장에서 판매되는 기타 제품 제외
- 카라멜 박스 (6개입)은 카라멜 종류당 하나가 판매된 것으로 간주

요일별 매출 데이터: **sales_by_day_of_week**

속성	설명	예시
day_id	요일 고유 식별자 - 일, 월, 화, 수, 목, 금, 토 순서	1
day_of_week	요일 - 일요일, 월요일, 화요일, 수요일, 목요일, 금요일, 토요일	일요일
total_sales	매출 합계 (₩)	10,000,000
order_count	결제 건수	1000
average_order_value	영수 단가 (₩)	10,000

총 데이터 수: 7개 데이터

6. 회고

- 네이버 지도, 카카오 지도, 소셜 미디어 (e.g. 인스타그램, 네이버 블로그) 등에 올라온 **고객의 리뷰와 피드백의 텍스트 데이터**를 수집하고 분석하여 제품에 대한 고객의 선호도나 의견을 이해하는 데 사용해 보면 좋을 것 같다.

텍스트 마이닝을 통해 특정 제품에 대한 긍정적인, 또는 부정적인 감정을 추출하고 분류하여 제품의 인기도와 품질에 대한 인사이트를 얻을 수 있을 것이다.

- 오프라인 판매만 이루어지는 상황에서 고객을 추적할 방법이 없어 **재방문 고객 여부**를 확인하거나 고객이 **관광객인지 현지인인지 구별**하는 것이 불가능했다.

홍차 전문 카페인 만큼 관광객이나 첫 방문 고객은 홍차 관련 음료를 선택할 가능성이 높다. 현지인 (동네 고객)이나 재방문 고객은 어떤 제품을 선호하고 재구매하는지 알 수 없어서 아쉬웠다.