

연관 규칙 분석을 활용한 오프라인 마케팅 판매 촉진 전략

박재영

개요

	날짜	메뉴	포스번호	영수번호	시간	수량	매출	메뉴번호	요일	날씨
0	20221125	[40863]카라멜팝콘(M)	2	73	1956	1	5000	40863	금	비
1	20221030	[40860]컵와인 까베르네소비뇽	1	293	1912	2	18182	40860	일	맑음
2	20221109	[00413]아이스 아메리카노(Ice Coffee)	1	43	1931	1	4545	413	수	흐림
3	20221019	[40863]카라멜팝콘(M)	1	20	1925	1	5000	40863	수	맑음
4	20221112	[15867]츄러스	1	109	1328	1	4545	15867	토	비

원본 데이터

목표

오프라인 매점의 판매기록을 바탕으로

어떻게 하면 매출 및 판매량을 향상시킬 수 있을지 인사이트를 도출

Contents

01 데이터 전처리

- 1) 데이터 전처리
- 2) 고객별 구매 데이터

03 EDA 및 분석 2

- 1) 연관 규칙 분석

02 EDA 및 분석 1

- 1) 인기 메뉴
- 2) 메뉴와 외부요인
- 3) 매출 분석

04 인사이트

데이터 전처리

- 1) 데이터 전처리
- 2) 고객별 구매 데이터

데이터 전처리

데이터 전처리

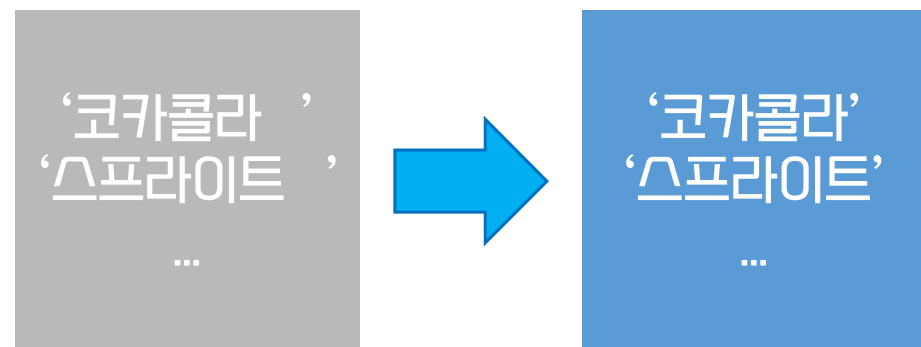
데이터 원형 (결측값 없음)

	날짜	메뉴	포스번호	영수번호	시간	수량	매출	메뉴번호	요일	날씨
0	20221125	[40863]카라멜팝콘(M)	2	73	1956	1	5000	40863	금	비
1	20221030	[40860]컵와인 까베르네소비뇽	1	293	1912	2	18182	40860	일	맑음
2	20221109	[00413]아이스 아메리카노(Ice Coffee)	1	43	1931	1	4545	413	수	흐림
3	20221019	[40863]카라멜팝콘(M)	1	20	1925	1	5000	40863	수	맑음
4	20221112	[15867]츄러스	1	109	1328	1	4545	15867	토	비

1. 메뉴번호와 메뉴명 분리

메뉴
카라멜팝콘(M)
컵와인 까베르네소비뇽
아이스 아메리카노(Ice Coffee)
카라멜팝콘(M)
츄러스

2. 메뉴 앞뒤 공백 제거



데이터 전처리

데이터 전처리

3. 날짜-시간 병합

일시
2022-11-25 19:56:00
2022-10-30 19:12:00
2022-11-09 19:31:00
2022-10-19 19:25:00
2022-11-12 13:28:00

4. 메뉴 정리 (총 37종 -> 35종)

- ‘(배민)양념 소떡소떡(1p)’ 제거
실제 메뉴판 이미지에서 발견되지 않음.
적은 양의 데이터 (251개)
11/10~27 에만 한정판매
- ‘초코라떼 (HOT)’와 ‘핫초코’ 통합
판매일자가 겹치지 않음.
매출액이 4545로 같음.
‘초코라떼 (HOT)’ 19개 < ‘핫초코’ 606개

변경전

	날짜	메뉴	포스번호	영수번호	시간	수량	매출	메뉴번호	요일	날씨
1611	20221023	초코라떼 (HOT)	1	179	1714	1	4545	40855	일	맑음

	날짜	메뉴	포스번호	영수번호	시간	수량	매출	메뉴번호	요일	날씨
1611	20221023	핫초코	1	179	1714	1	4545	448	일	맑음

변경후

데이터 전처리

데이터 전처리

5. 매출액 정리

- 수량에 따른 매출액이 조금씩 차이나는 데이터 수정
(수량 1 매출 기준)

변경전

메뉴번호	메뉴	수량	매출
40860	컵와인 까베르네소비뇽	2	18182
34433	순살닭강정	2	14545
9175	뉴욕핫도그	2	12727

변경후

메뉴번호	메뉴	수량	매출
40860	컵와인 까베르네소비뇽	2	18182
34433	순살닭강정	2	14546
9175	뉴욕핫도그	2	12728

데이터 전처리

고객별 구매 데이터

동일한 **날짜** 내에서
포스번호와 **영수번호**가 동일한 데이터들은
한 고객의 구매 데이터로 정의



연관 규칙 분석에 활용

	일시	요일	날씨	포스번호	영수번호	메뉴번호	메뉴	수량	매출
12	2022-10-19 19:07:00	수	맑음	2	2	15867	츄러스	1	4545
13	2022-10-19 19:07:00	수	맑음	2	2	32291	코카콜라	1	3182
14	2022-10-19 19:07:00	수	맑음	2	2	34433	순살닭강정	1	14546



pivot_table

	Beer set	Couple set A	Couple set B	츄러스	코카콜라	순살닭강정
구매번호						
1019-2-2	0	0	0	1	1	1

EDA 및 분석 1

- 1) 인기 메뉴
- 2) 메뉴와 외부요인
- 3) 매출 분석

EDA 및 분석 1

인기 메뉴

1. 메뉴 매출 순위

츄러스의 매출액이 가장 높고,
샤퀴테리의 매출액이 가장 낮다.

	메뉴	매출	rank_price
0	츄러스	38214360	1
1	카라멜팝콘(M)	37210000	2
2	카라멜 팝콘(L)	21245667	3
3	Couple set A	20159370	4
4	뉴욕핫도그	10806072	5
30	하겐다즈 바닐라	776408	31
31	Couple set B	465456	32
32	사케	450000	33
33	뱅쇼(G)	280000	34
34	샤퀴테리 Box (4구)	95452	35

2. 메뉴 판매량 순위

판매량 순위도 매출 순위와
마찬가지이다.

	메뉴	수량	rank_count
0	츄러스	8408	1
1	생수	7736	2
2	카라멜팝콘(M)	7442	3
3	코카콜라	2714	4
4	카라멜 팝콘(L)	2337	5
30	하겐다즈 바닐라	122	31
31	사케	45	32
32	뱅쇼(G)	28	33
33	Couple set B	16	34
34	샤퀴테리 Box (4구)	7	35

EDA 및 분석 1

인기 메뉴

3. 단위당 수익 비교

일반적으로 판매량 대비 수익이 클수록 효율적인 판매

Family set 가
가장 효율적인 판매 기록,
생수 비효율적

메뉴	매출	수량	단위당수익	rank_unit
Family set	4976349	161	30909.0	1
Couple set B	465456	16	29091.0	2
Beer set	6218320	380	16364.0	3
Couple set A	20159370	1386	14545.0	4
샤퀴테리 Box (4구)	95452	7	13636.0	5
미닛메이드 포도	1021716	281	3636.0	30
스프라이트	4149328	1304	3182.0	32
코카콜라	8635948	2714	3182.0	32
토레타	2294222	721	3182.0	32
생수	7032024	7736	909.0	35

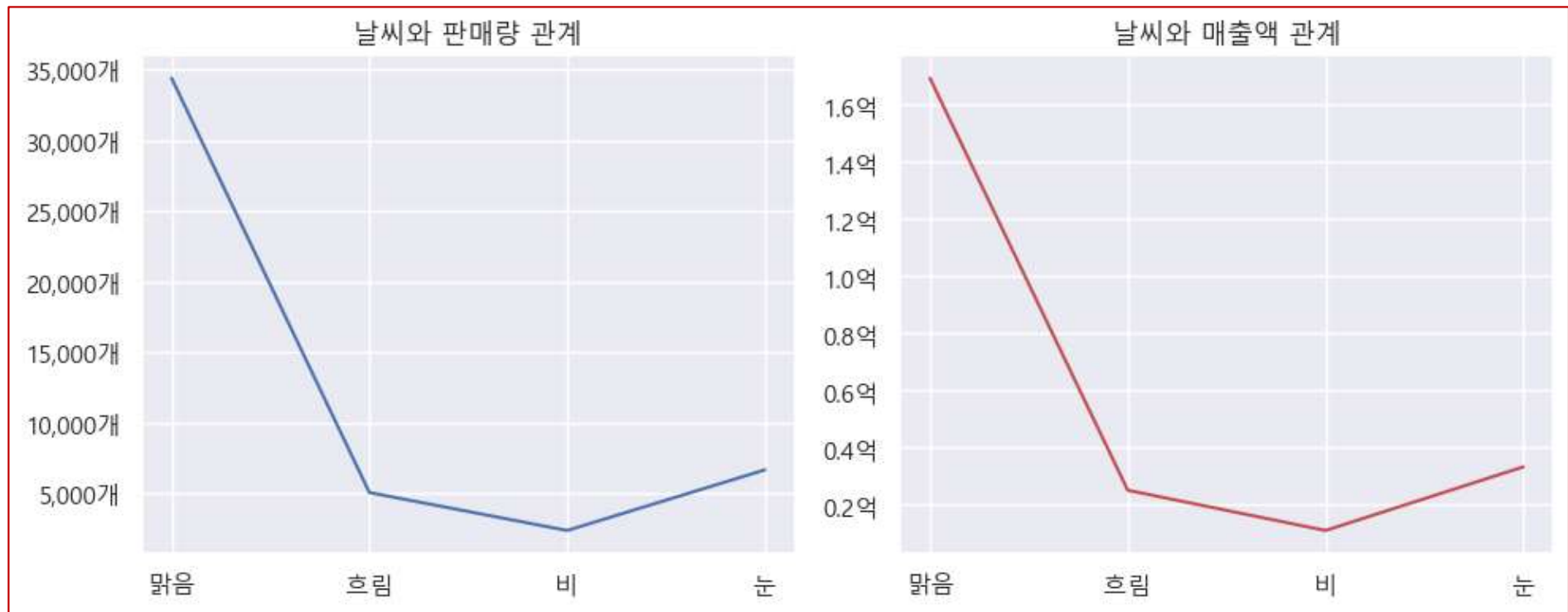
EDA 및 분석 1

메뉴와 외부요인

1. 날씨와 판매량/매출액 비교

날씨에 따른 판매량과 매출액이 비슷한 추세를 보인다.

맑음일때 판매량과 매출액이 특출나고 **눈, 흐림, 비** 순서로 이어진다.

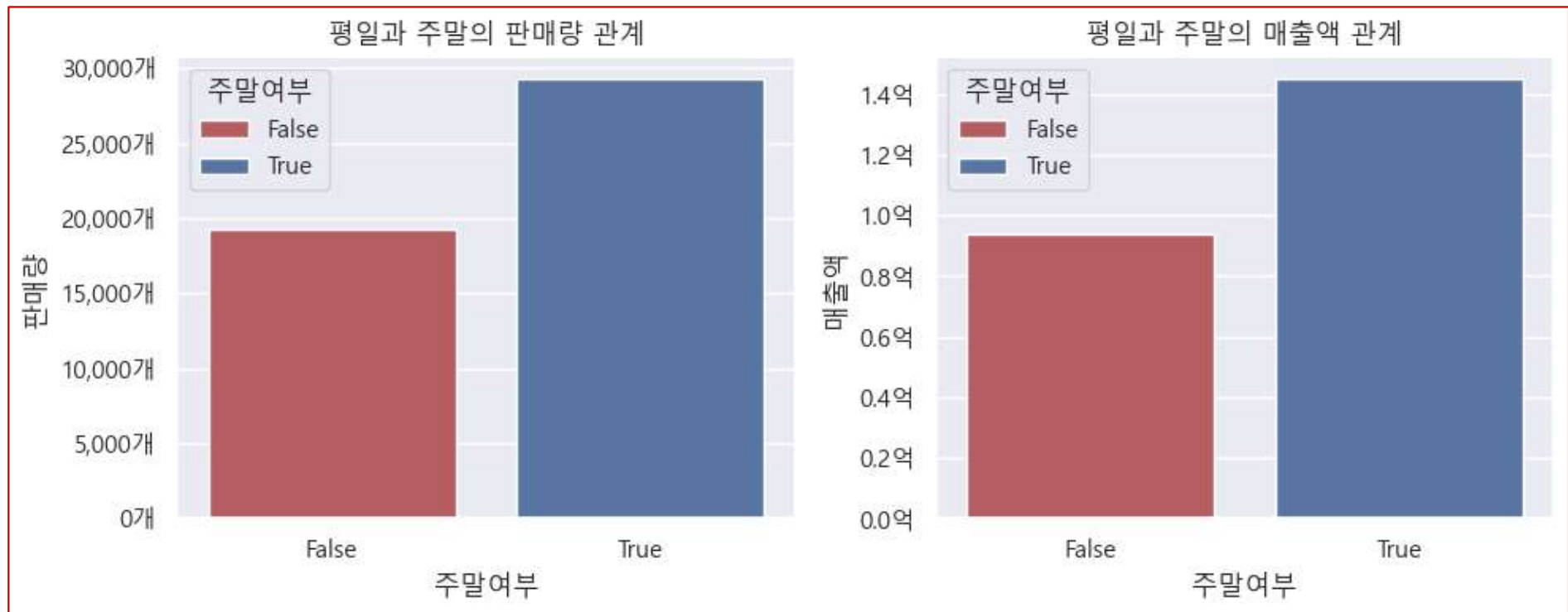


EDA 및 분석 1

메뉴와 외부요인

2. 평일과 주말 판매량/매출액 비교

평일보다는 **주말**일때 판매량과 매출액이 우수하다.

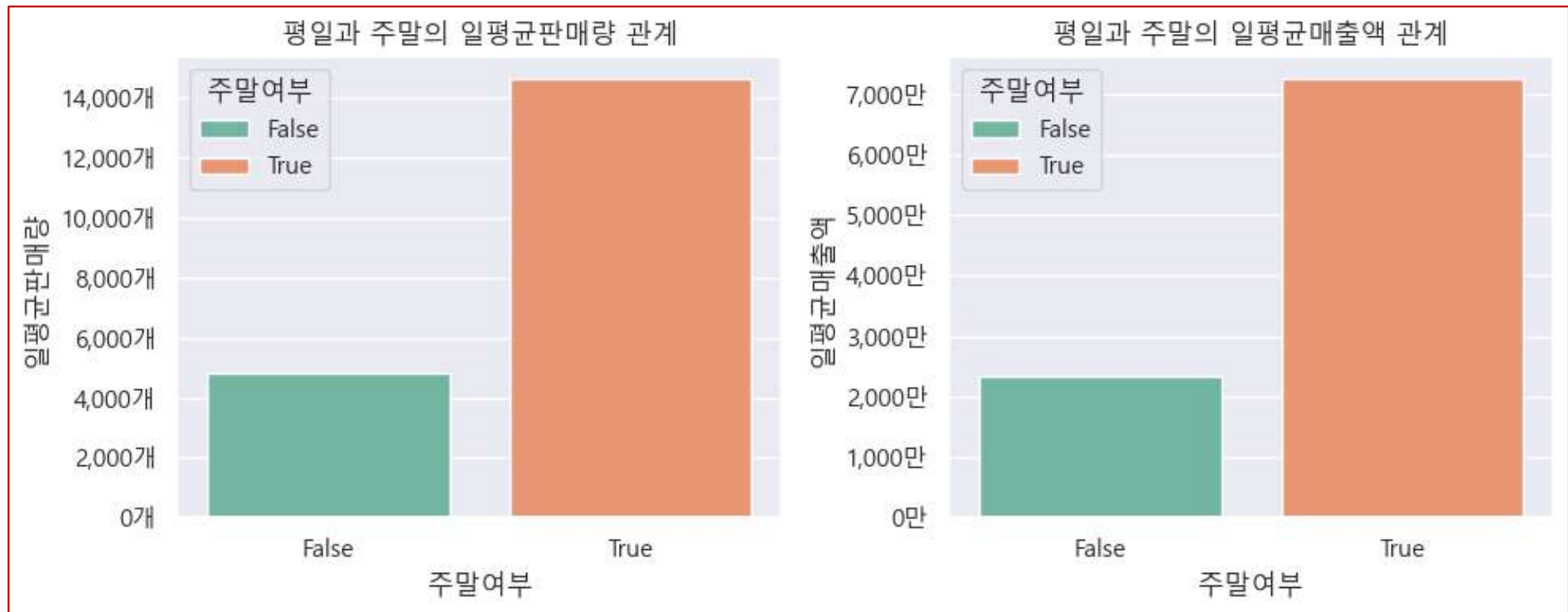


EDA 및 분석 1

메뉴와 외부요인

2. 평일과 주말 판매량/매출액 비교

평일은 4일, 주말은 2일임에도 **일평균**으로 비교하면 더 차이가 두드러진다.



EDA 및 분석 1

메뉴와 외부요인

3. 요일에 따른 매출 비교

앞선 결과에 이어 각 **요일별**로 세분화했을때 차이를 다시 확인할 수 있다.
또한 화요일 공연은 3일 뿐이었음에도 일평균매출액이 낮지 않은 편이다.

	요일	요일횟수	총매출액	일평균매출액
0	화	3	7831192	2610397
1	수	11	24697657	2245242
2	목	11	23837246	2167022
3	금	11	37416436	3401494
4	토	11	78652352	7150214
5	일	11	66287200	6026109



EDA 및 분석 1

매출 분석

1. 매출 이상치 확인 (특별한 행사나 판매상 문제가 있었는지)

- threshold 1.5 인 **사분위수 방법**에서는 매출 이상치 X
- threshold 2 (95%에 해당)인 **Z-score 방법**으로도 2개의 이상치만 발견

	날짜	매출	z매출
34	2022-12-03	8814397	2.015623
56	2022-12-31	10440204	2.713083

- 따라서, 전체 58일치 데이터 중 10%인 **상위 5개, 하위 5개 데이터**를 이상치로 설정

	날짜	매출	요일
56	2022-12-31	10440204	Saturday
34	2022-12-03	8814397	Saturday
39	2022-12-10	8146609	Saturday
13	2022-11-05	7926200	Saturday
29	2022-11-26	7711198	Saturday

<총매출액 상위 5일>

	날짜	매출	요일
10	2022-11-02	1804970	Wednesday
32	2022-12-01	1722251	Thursday
27	2022-11-24	1703131	Thursday
26	2022-11-23	1609506	Wednesday
42	2022-12-15	1274079	Thursday

<총매출액 하위 5일>

EDA 및 분석 1

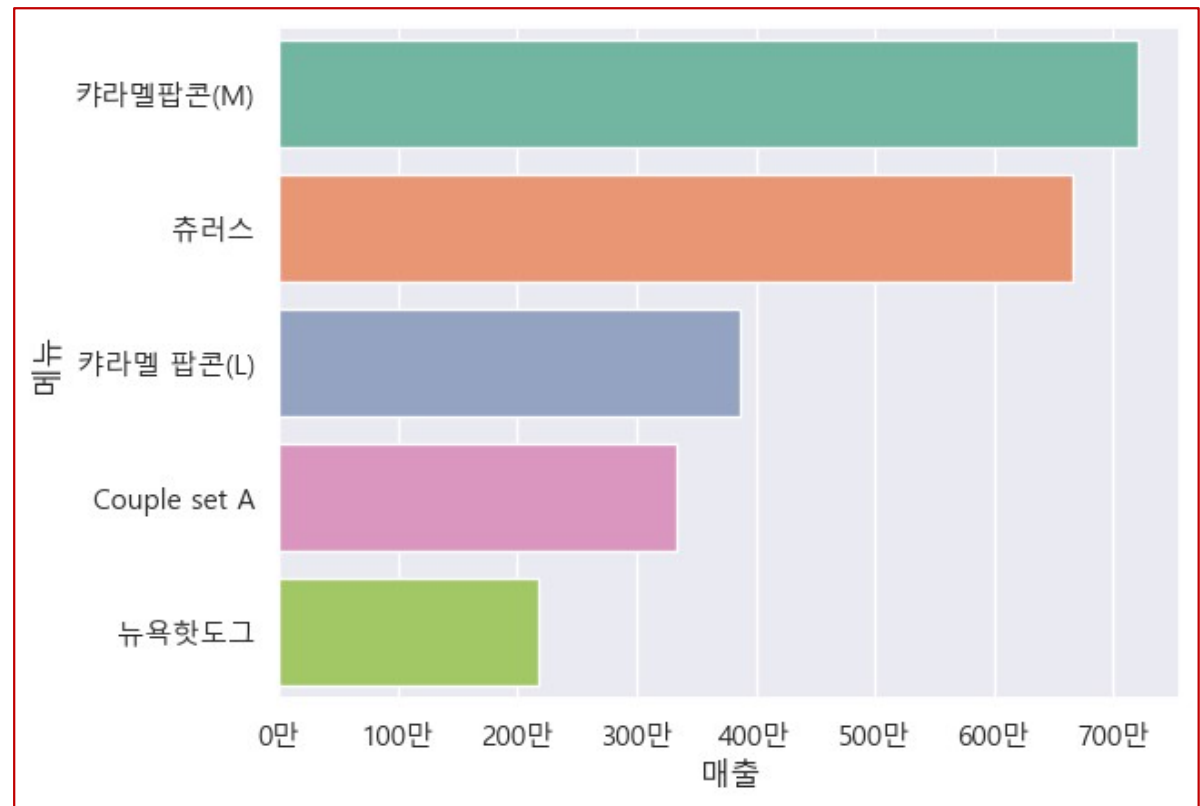
매출 분석

2. 매출이 높을 때 인기 메뉴

- 1위와 2위만 다를 뿐, 기존 인기 메뉴 매출 상위 5개의 메뉴가 그대로 유지

	메뉴	매출	rank_price
0	쥬러스	38214360	1
1	카라멜팝콘(M)	37210000	2
2	카라멜 팝콘(L)	21245667	3
3	Couple set A	20159370	4
4	뉴욕핫도그	10806072	5
30	하겐다즈 바닐라	776408	31
31	Couple set B	465456	32
32	사케	450000	33
33	뱅쇼(G)	280000	34
34	샤퀴테리 Box (4구)	95452	35

<인기메뉴 - 1. 매출 순위>



EDA 및 분석 1

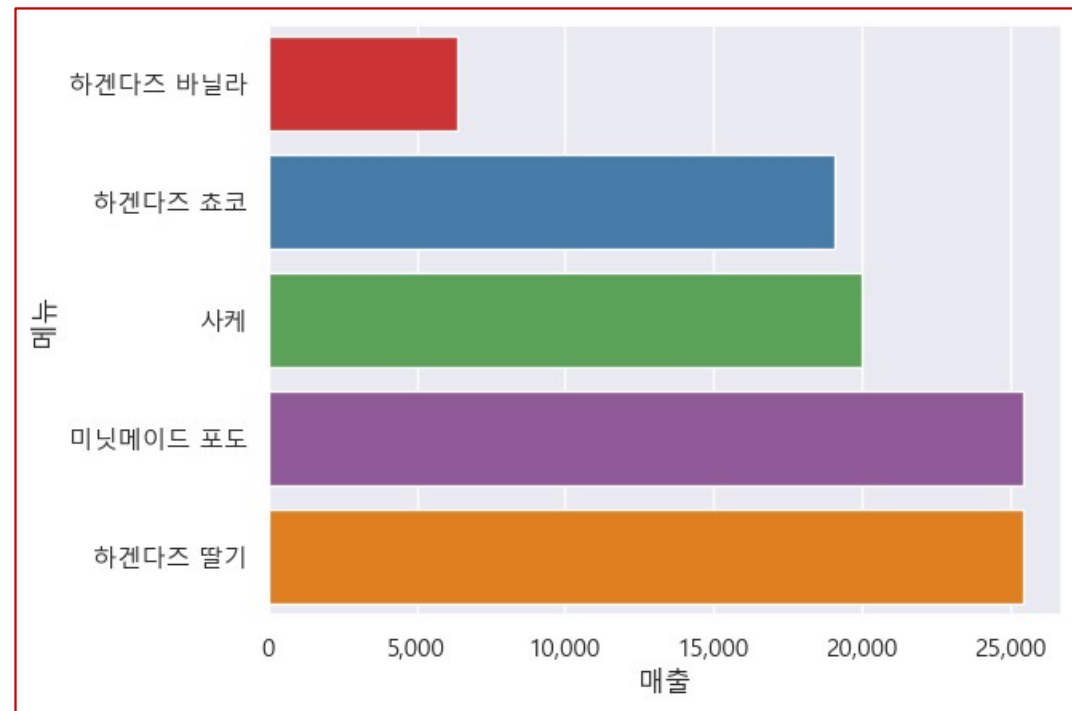
매출 분석

3. 매출이 낮을 때 비인기 메뉴

- 기존 인기 메뉴 매출 하위권의 메뉴가 그대로 유지

	메뉴	매출	rank_price
27	미닛메이드 포도	1021716	28
28	하겐다즈 초코	954600	29
29	하겐다즈 딸기	782772	30
30	하겐다즈 바닐라	776408	31
31	Couple set B	465456	32
32	사케	450000	33
33	뱅쇼(G)	280000	34
34	샤퀴테리 Box (4구)	95452	35

<인기메뉴 - 1. 매출 순위>



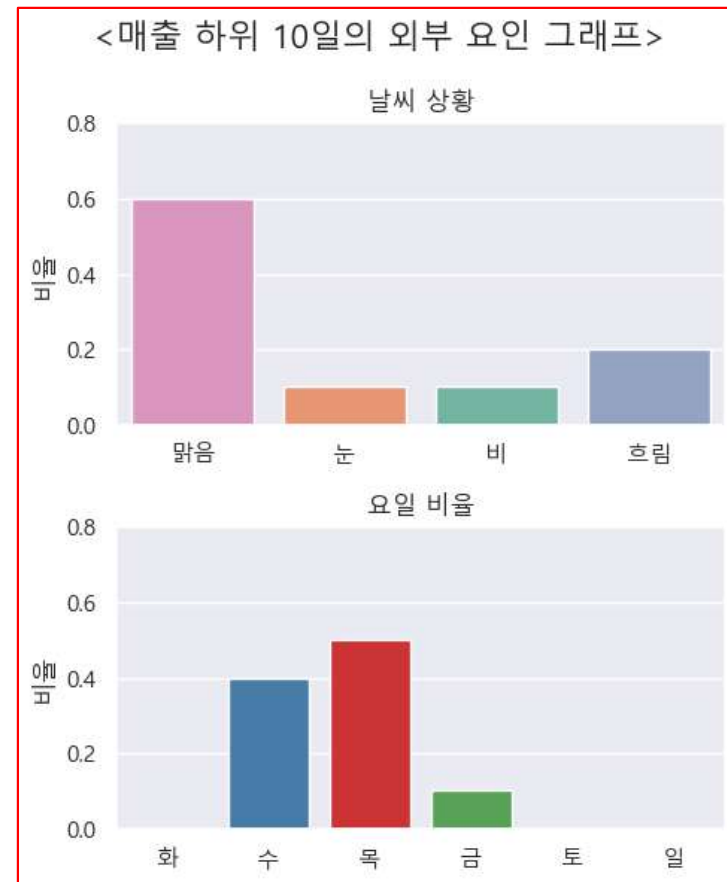
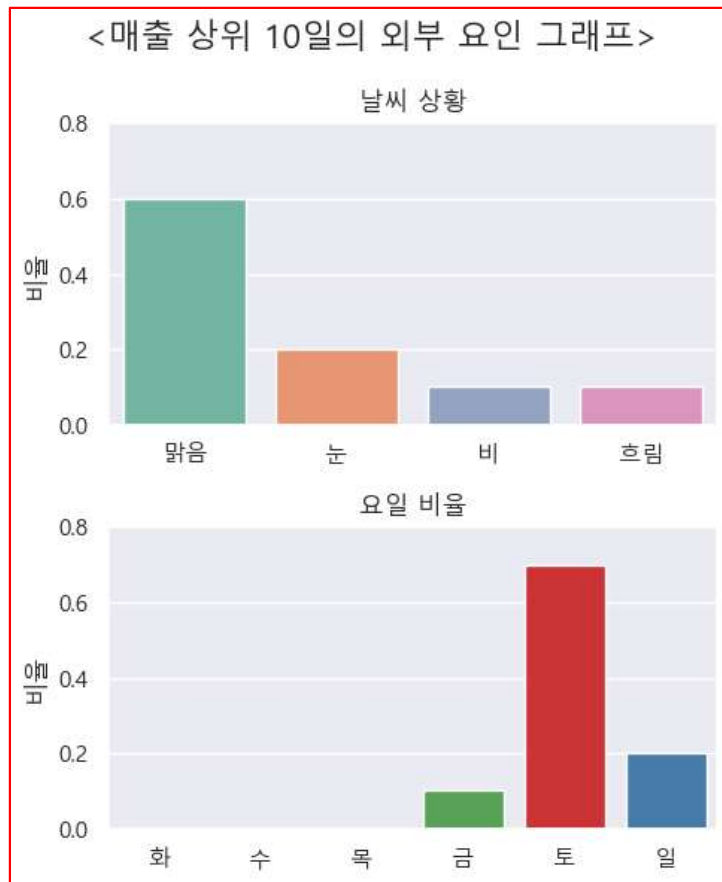
특별한 행사나 판매상 문제는 없었던 것으로 보임.

EDA 및 분석 1

매출 분석

4. 매출에 더 큰 영향을 끼치는 요인

- 날씨보다는 주말인지 아닌지가 매출에 더 큰 영향



EDA 및 분석 2

1) 연관 규칙 분석

EDA 및 분석 2

연관 규칙 분석

연관 규칙 분석이란?

항목들의 관계를 조건형식으로 찾아가는 분석 방법

마케팅에서는 **고객들의 상품 구매 데이터**를 이용하여 **품목 간의 연관성**을 확인

신뢰도(Confidence)

X를 포함하는 거래 내역 중 Y가 포함된 비율

$$P(Y|X) = P(X \cap Y) / P(X)$$

지지도(Support)

X와 Y를 동시에 포함하는 거래 비율

$$P(X \cap Y)$$

향상도(Lift)

두 관계에 대한 설명

- Lift > 1 : X와 Y가 양의 상관관계
- Lift = 1 : 무관계. 독립
- Lift < 1 : X와 Y가 음의 상관관계

$$P(Y|X) / P(Y)$$

EDA 및 분석 2

연관 규칙 분석

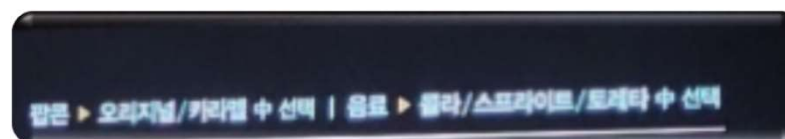
Apriori 알고리즘

- I. 설정한 **최소지지도** 이상인 지지도를 갖는 **빈발 품목 집합(Frequent Item Set)** 생성
- II. 빈발 품목 집합의 **모든 부분집합**에 대해 설정한 **최소신뢰도**를 넘는 연관 규칙 탐색
(최소 지지도와 최소신뢰도 0.005로 설정)

EDA 및 분석 2

연관 규칙 분석

antecedents	consequents	support	confidence	lift	leverage
(카라멜팝콘(M))	(생수)	0.019	0.095	0.504	-0.019
(생수)	(카라멜팝콘(M))	0.019	0.103	0.504	-0.019
(카라멜팝콘(M))	(코카콜라)	0.011	0.055	0.737	-0.004
(코카콜라)	(카라멜팝콘(M))	0.011	0.150	0.737	-0.004
(카라멜팝콘(M))	(스프라이트)	0.005	0.026	0.675	-0.003
(스프라이트)	(카라멜팝콘(M))	0.005	0.137	0.675	-0.003
(카라멜팝콘(M))	(아이스 아메리카노(Ice Coffee))	0.005	0.025	0.553	-0.004
(아이스 아메리카노(Ice Coffee))	(카라멜팝콘(M))	0.005	0.113	0.553	-0.004



팝콘과 음료를 함께 판매하는 **커플세트(매출 4위)**가 있지만, **생수(판매량 2위)**와 **아이스 아메리카노(판매량 6위)**는 음료 선택지에 없음.

또한, 음료도 세트가 아닌 **단품**으로 구매하는 고객이 많음.

EDA 및 분석 2

연관 규칙 분석

antecedents	consequents	support	confidence	lift	leverage
(코카콜라)	(츄러스)	0.008	0.107	0.523	-0.007
(츄러스)	(코카콜라)	0.008	0.039	0.523	-0.007



츄러스는 매출과 판매량 모두 1위인 메뉴
반면에, **뉴욕핫도그**는 매출 4위, 판매량 7위

EDA 및 분석 2

연관 규칙 분석

antecedents	consequents	support	confidence	lift	leverage
(클라우드(캔))	(카라멜팝콘(M))	0.005	0.118	0.582	-0.004
(카라멜팝콘(M))	(클라우드(캔))	0.005	0.025	0.582	-0.004



맥주(매출, 판매량 5위)와 함께 구매하는 메뉴는 카라멜팝콘(매출, 판매량 2~3위)

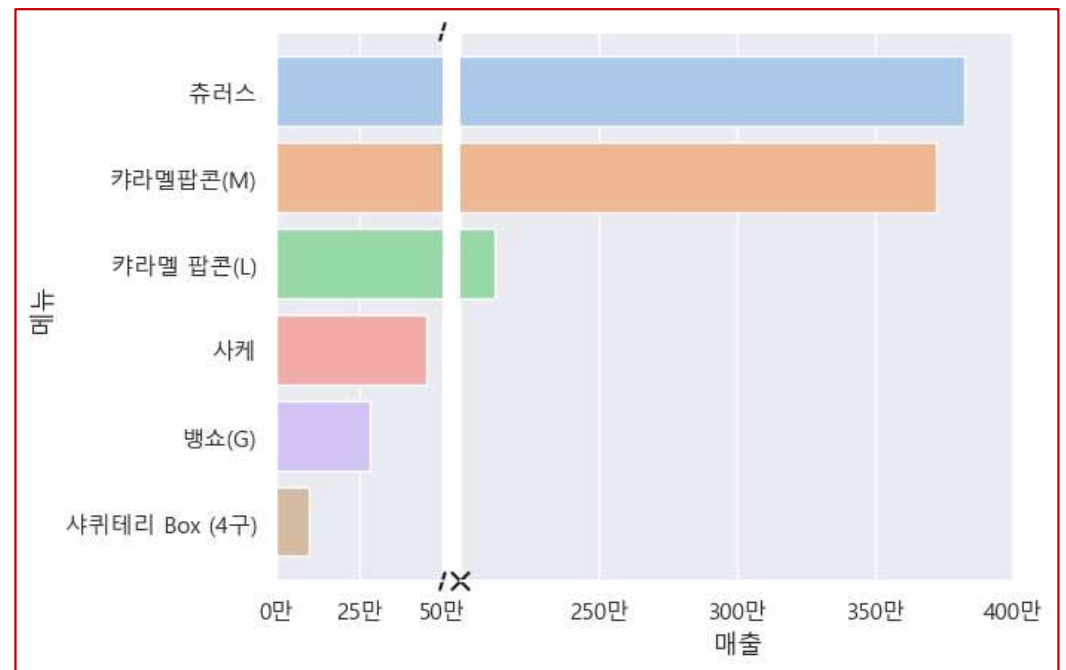
순살닭강정은 매출 6위

인사이트

인사이트 1

날씨가 맑을수록, 평일보다 주말일 때 매출이 높다.

인기 메뉴와 비인기 메뉴의 매출 격차가 매우 크다.



인사이트 2

커플세트 메뉴의 음료 선택지 다양성을 늘려야 할 필요가 있다.

높은 판매량을 보인 생수와 아이스 아메리카노를 세트 메뉴에 포함한다면 **confidence**에 따라 커플세트의 판매량을 **최대 26% 더 높일 수 있다.**

(커플세트의 판매량이 증가한 만큼 생수/아이스 아메리카노의 판매량은 감소할 것이다.)

(다만, 세트메뉴로의 접근성을 높임으로써 단품 구매 고객도 세트메뉴 구매로 유도할 수 있을 것이다.)



커플세트 (매출 4위)

생수 (판매량 2위)

아이스 아메리카노(판매량 6위)

인사이트 3

싱글세트 구성품을 변경하여 해당 메뉴의 매출을 높일 수 있다.

핫도그 대신 가장 높은 매출과 판매량을 나타낸 츠러스를
싱글 세트에 적용한다면
단품으로 츠러스와 음료를 사는 고객을 세트 메뉴로 전환하여
매출을 16% 더 높일 수 있을 것이다.



인사이트 4

맥주 세트 메뉴도 구성의 변화가 필요하다.

닭강정은 식사 대용이 될 수 있다.

그보다는 간단한 간식으로 즐길 수 있는 팝콘을 활용하는 것이 좋아 보인다.



맥주 (매출, 판매량 5위)
카라멜팝콘(M/L) (매출, 판매량 2~3위)

맥주 세트 (매출 12위)
순살닭강정 (매출 6위)

감사합니다