

BDA 데이터 분석/활용 공모전

Track 1 : 시각화 인사이트

이쿠조 MEMBERS
신재우 김세원 박희주 남화승

track1 : 이쿠조

1. 데이터 전처리 .

불필요한 컬럼 제거 | 수치형 변수로 변환 | 파생변수 생성

1. 데이터 전처리

(1) 불필요한 컬럼 제거

결측치

- 덤상품수량
- GS홈쇼핑주문번호
- 상품타입 ('선물세트' 제외)

단일값

- 단독배송여부
- 배송유형 ('3PL' 제외)

의미하지 않은 값

- 등록일자
- 플랜트 정보

(2) 문자형 변수 → 수치형 변수

0 ~ n의 수치로 변환

거래처	주문/배송상태	클레임 상태
사은품 여부	상품유형	마감구분

* 마감구분 > 풀필먼트

풀필먼트라는 마감구분이 따로 있는 것이 아니라 플랜트 정보에 영향을 받은 것이기에 자재명을 통해 상온 / 냉장 / 냉동으로 재분류

(3) 파생변수 생성

월	from 주문일
출고날짜오차	운송장등록날짜 - 출고예정일
예상출고소요기간	출고예정일 - 주문일
실제출고소요기간	운송장등록날짜 - 주문일
자재명_분류	from 자재명
묶음수량	초기자재수량 / 상품수량

출고 및 배송은 소비자의 구매확정에 있어 중요한 부분으로, 해당 컬럼들을 통해 소비자들이 만족 및 불만족한 포인트를 찾아낼 수 있다.

자재명 컬럼으로부터 브랜드 및 상품분류를 추출하여 소비자들이 많이 찾는 CJ 브랜드를 알아낼 수 있다.

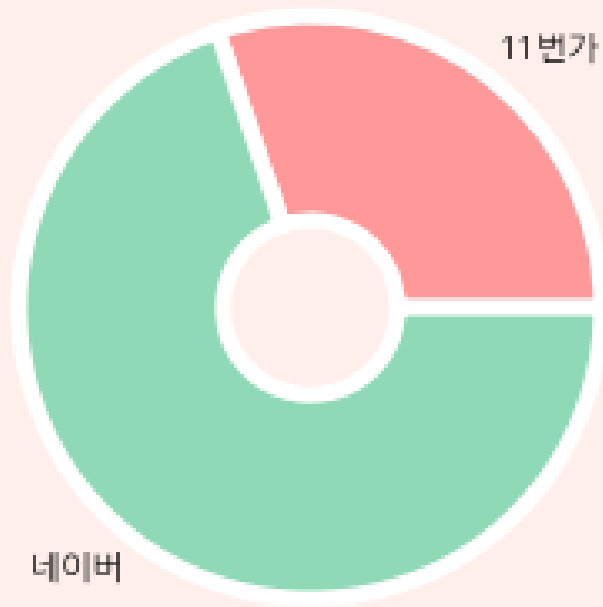
track1 : 이쿠조

2. Main Data 데이터 분석 .

Main Data : 매출이 발생한 데이터 ; '배송완료' & '클레임 X (-)'

1) 거래처

(1) 매출액 비교



p-value : 0.0

네이버 >> 11번가

(2) 거래처별 판매량 순위

비비고	74475
다담	37685
햇반	21653
고메	19912
백설	4151
기타	1663
스팸	1436
쁘띠첼	1282
냉면	1025
꼬마	763
닭가슴살	636
고추장	552
다시다	532
하선태	355
밀당의고수	289

11번가

햇반	77908
비비고	44038
백설	12213
고메	10225
스팸	5430
냉면	3128
기타	1820
다담	1505
쁘띠첼	1301
닭가슴살	1277
고추장	1259
명가	827
꼬마	706
다시다	639

네이버

2) 주문일

(1) 월별 매출

< 4월 : 11번가 데이터만 존재 >

- 주문일별 매출액 평균이 같지 않다.
- '2022-04-11' 에 매출이 급증한 이유는 11번가의 고유 행사인 십일절로 인한 것이라고 결론지을 수 있다.

11번가

< 5월 : 11번가 & 네이버 데이터 혼재 >

- 주문일별 매출액 평균이 같지 않다.
- '2022-05-11' 에 매출이 급증한 이유는 11번가의 고유 행사인 십일절로 인한 것이라고 결론지을 수 있다.

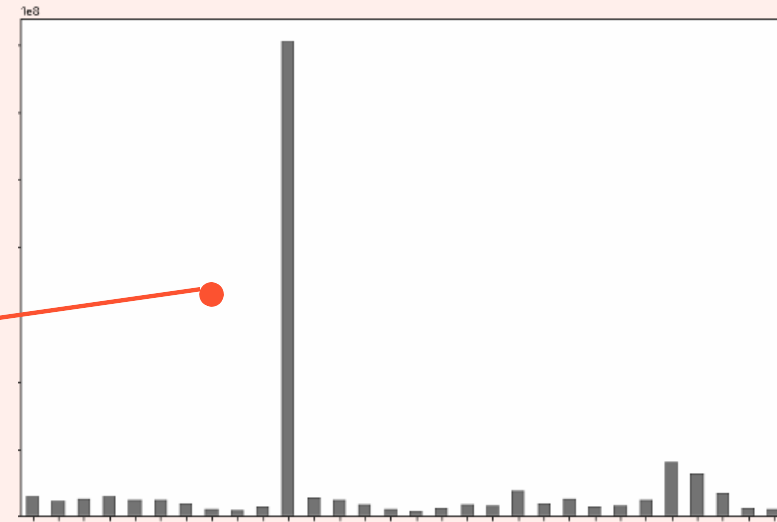
< 6월 : 네이버 데이터만 존재 >

- 주문일별 매출액 평균이 같지 않다.
- '2022-06-10' 에 매출이 급증한 이유를 알아보기 위해 (2)번 분석(5월 거래처별 매출 분석)을 진행하였다.

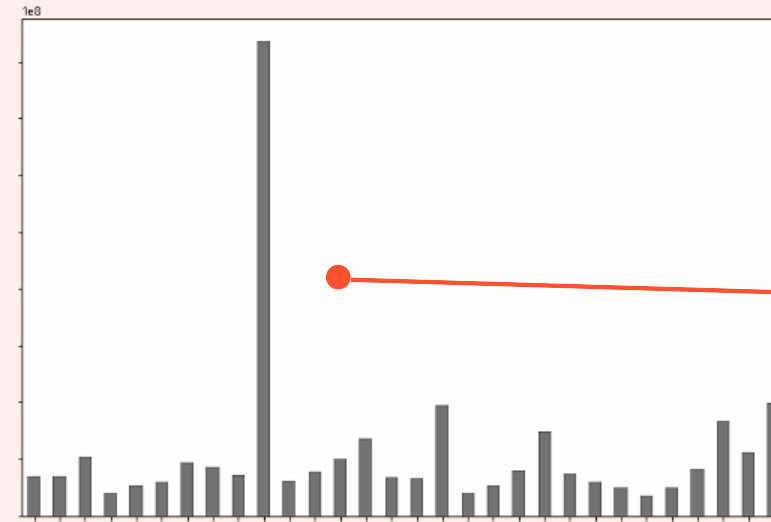
네이버

< 5월 : 11번가 & 네이버 데이터 혼재 >

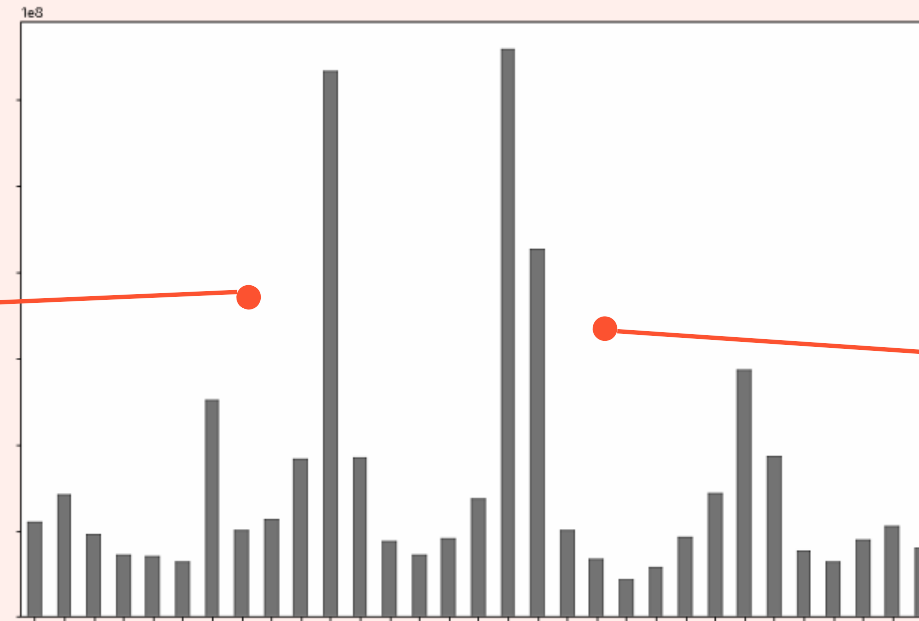
- 주문일별 매출액 평균이 같지 않다.
- '2022-05-17', '2022-05-18' 에 매출이 급증한 이유를 알아보기 위해 (2)번 분석(5월 거래처별 매출 분석)을 진행하였다.



4월



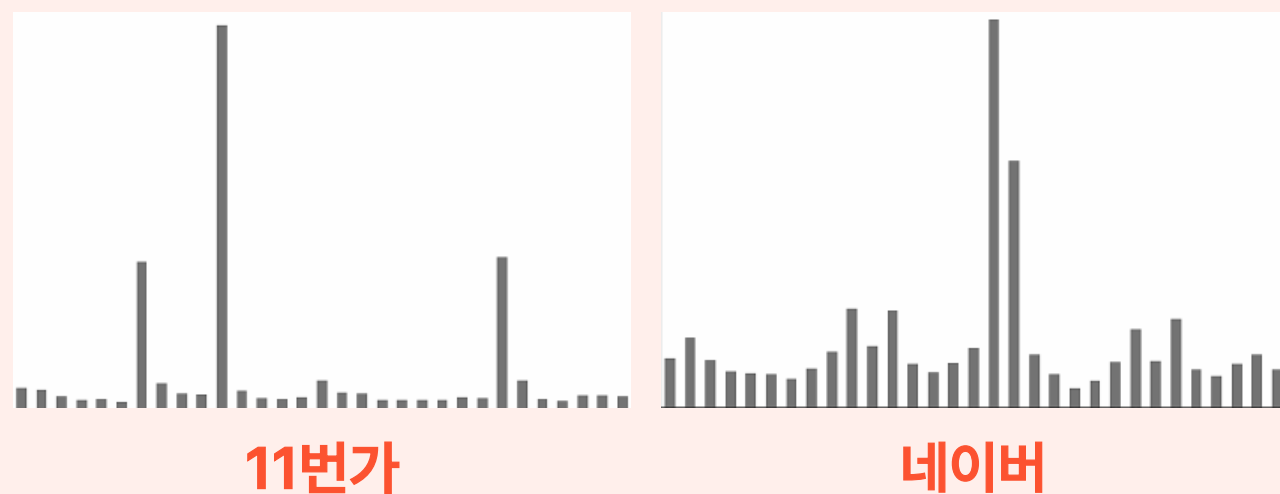
6월



5월

p-value : 0.0

(2) 5월 거래처별 매출



- 두 거래처 모두 5월 주문일별 매출액의 평균이 같지 않다. 즉, 각 주문일마다 특성이 있다.

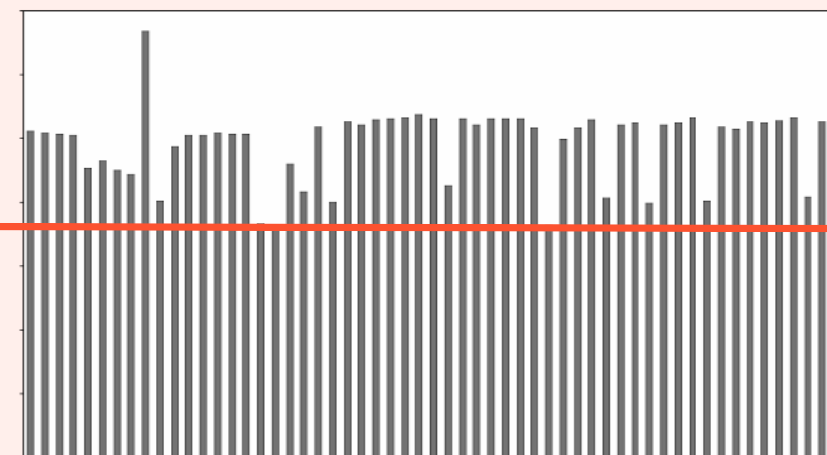
• '2022-05-11' 은 11번가의 십 일절로 인한 현상이다.

• '2022-05-17', '2022-05-18' 에 발생한 매출 급증은 네이버로 인한 것이고, 네이버의 어떤 행사가 원인인지는 확실치 않다

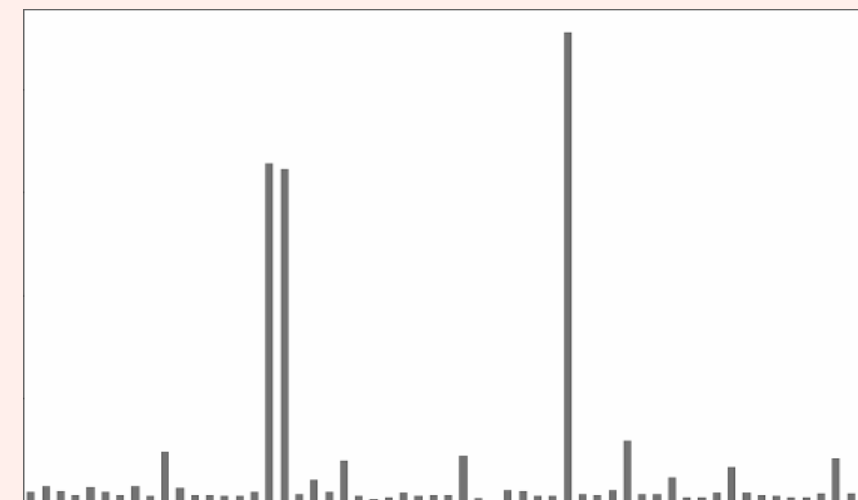
*분석한 핫반 = 핫반210GX24(1입)/상온

(3) 네이버 주문일별 핫반 통계량

최저가격선



- 최저가격선을 통해 5/17, 5/18의 평균 가격이 평소에 비해 훨씬 저렴하다는 것을 알 수 있다. (18,000원대)
- 이를 통해 할인행사의 존재를 유추할 수 있다.
- 패턴을 분석하면, 평소에는 4~7일 간격으로 작은 할인을, 한 달에 한 번씩 큰 할인을 진행 한다는 것을 알 수 있다.



- 5/17, 5/18, 6/10 를 제외한 다른 할인 날짜 에는 매출이 급증하지 않았다. 눈에 띄는 매출 증대는 있었지만, 위 3일에 비하면 저조하다.
- 차이는 '할인된 가격'이다.
18,000원대 초반 vs. 20,000원대

3) 자재명_분류

(1) 상품군 세분화

판매량이 많은 상위 브랜드 5개 세분화

비비고

만두, 국, 반찬, 밥

백설

기름, 조미료, 가루, 면, 반찬

햇반

밥, 조리된 밥

다담

양념

고메

피자, 튀김, 중화, 베이커리, 양식

비비고

만두, 국, 반찬, 밥

(2) - 1 | 비비고 묶음수량 변곡점

만두



최적 : 1개

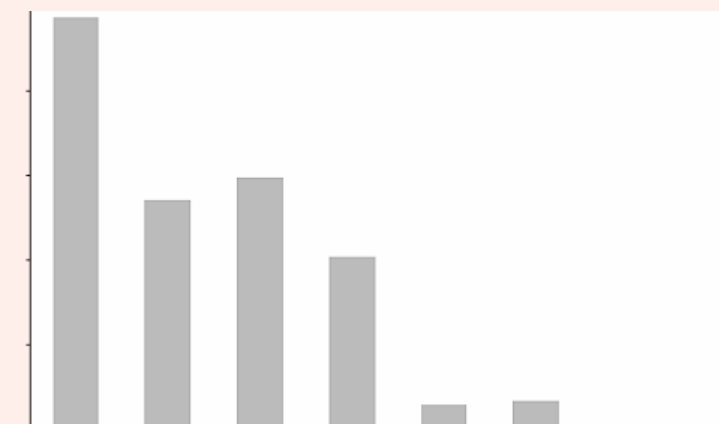
국



최적 : 1 > 3 > 18개

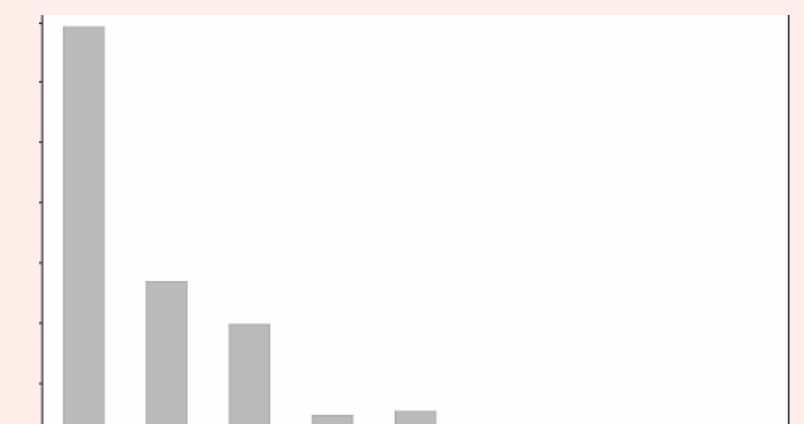
소비자 폭이 넓다

반찬



최적 : 1개

밥



최적 : 1개

햇반

밥, 조리된 밥

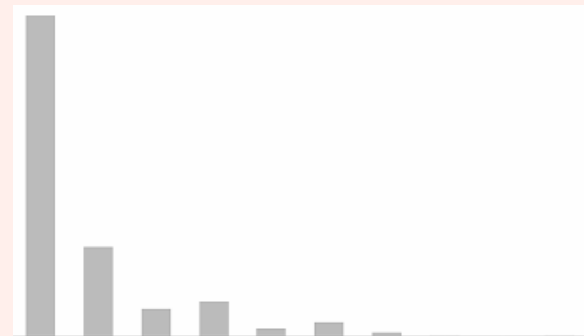
(2) - 3 | 햇반 묶음수량 변곡점

밥



최적 : 24 > 1개

조리된 밥



최적 : 1개

- 즉석밥을 구매하는 소비층은 평소에도 즉석밥을 자주 먹을 가능성이 크다는 것을 유추할 수 있다. 그렇기 때문에 한 번 살 때 많이 사는 것
- 유통기한도 9개월로, 길다는 것이 장점

백설

기름, 조미료, 가루, 면, 반찬

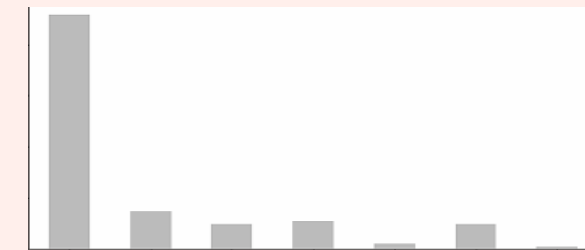
(2) - 2 | 백설 묶음수량 변곡점

반찬



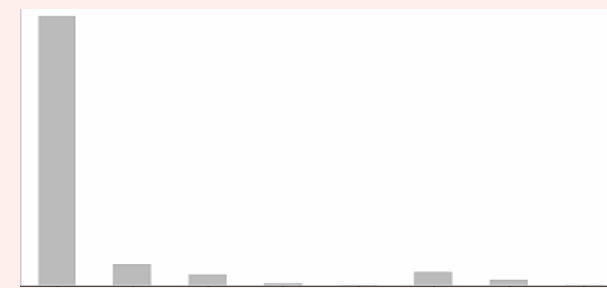
최적 : 1개

조미료



최적 : 1개

가루



최적 : 1개

기름



최적 : 2 > 3개

면



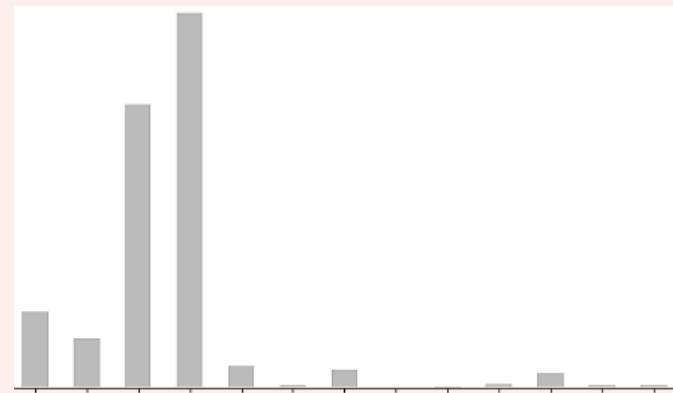
최적 : 1개

- 단기간 소비량 多
- 미개봉 상태에서는 5년까지도 소비O

다담

양념

(2) - 4 | 다담 묶음수량 변곡점



최적 : 4 > 3개

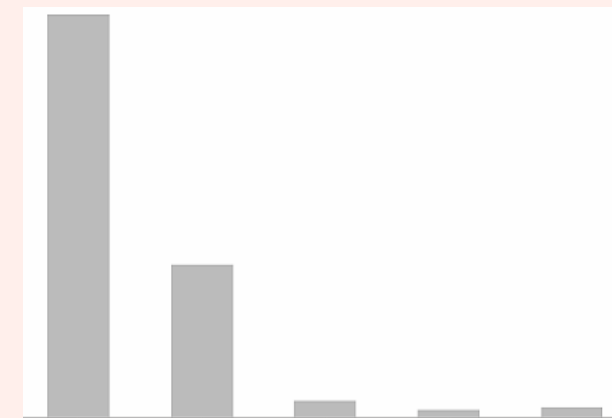
- 찌개, 국 등 우리 밥상에 자주 올라오는 음식의 양념
--> 한 번 구매할 때 여러 개 구매한다.
- 된장찌개 등 일반의 가정에서 직접 만드는 음식을
소비자들이 구매한다는 것
= 그 맛을 인정했다는 뜻
=> 이미 그 맛을 알기 때문에 믿고 여러 개 구매

고메

피자, 튀김, 중화, 베이커리, 양식

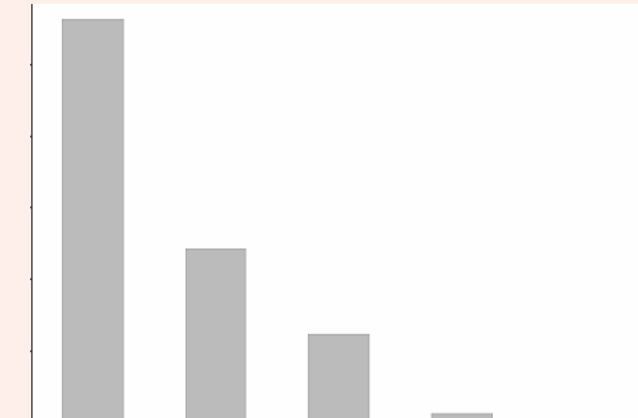
(2) - 5 | 고메 묶음수량 변곡점

피자



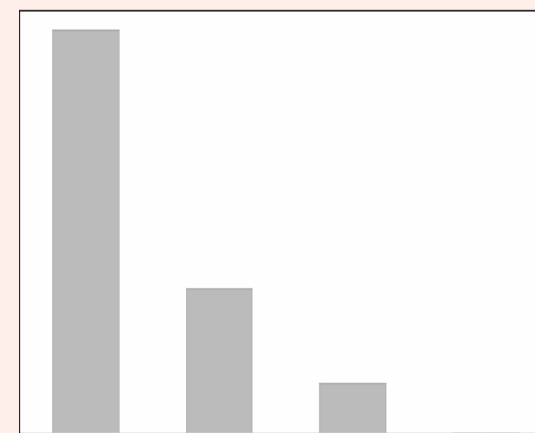
최적 : 1개

튀김



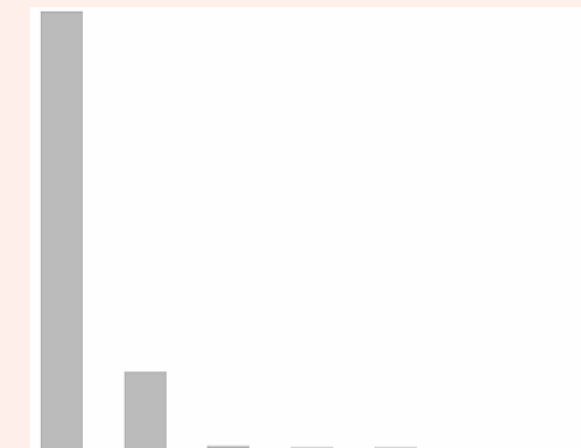
최적 : 1개

중화



최적 : 1개

베이커리



최적 : 1개

양식



최적 : 4 > 1개

4) 마감구분

< 상온 : 네이버가 더 높음 >

- 네이버 판매량 : 햇반 >> 비비고
- 상온 상품의 대부분은 햇반이다.
- 밥을 직접 지어먹는 핵가족이 아닌 1인 혹은 2인 가구라는 뜻이다.

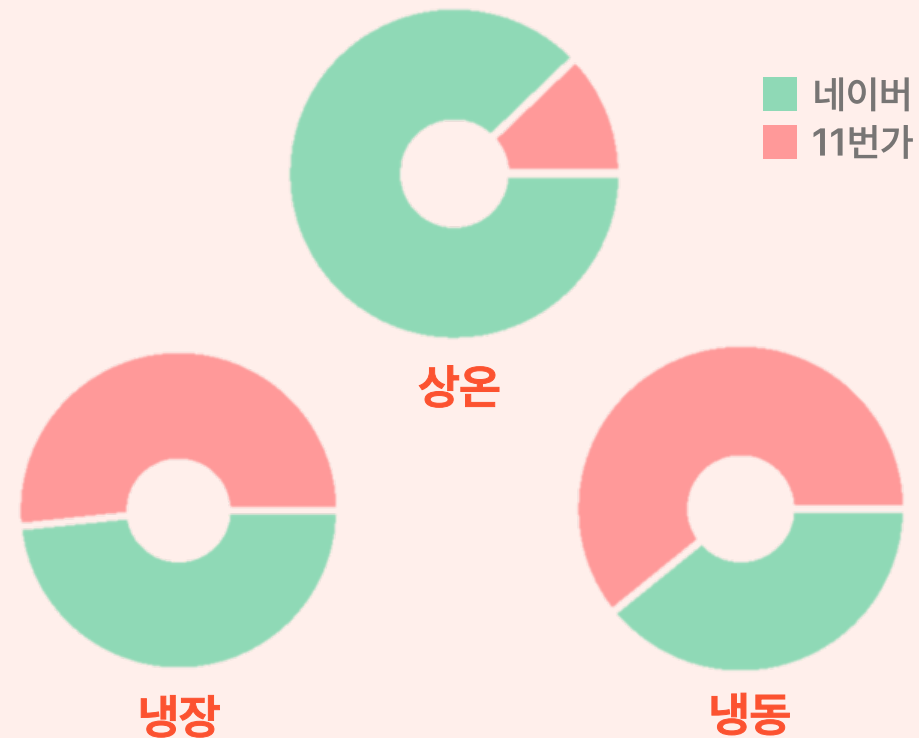
이는 네이버의 주 고객층이 젊은 2-30대이기 때문에 도출된 결과라고 할 수 있다.

< 냉동 & 냉장 : 11번가가 더 높음 >

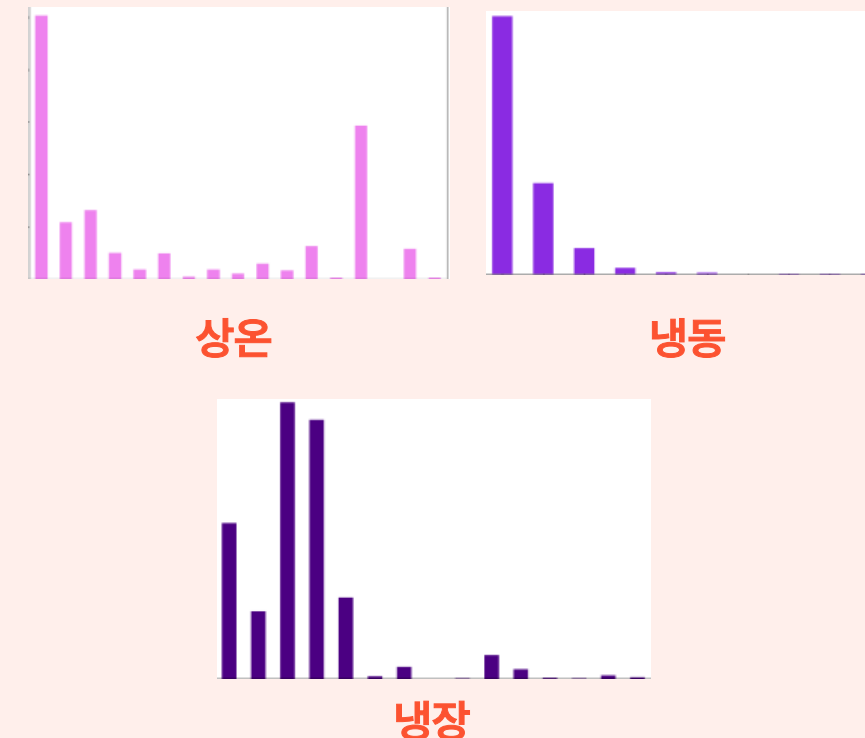
- 11번가 판매량 : 비비고 >> 다담
- 냉동, 냉장 식품 중 비비고와 다담이 차지하는 비중이 가장 높다.
- 자녀가 있는 4인 이상의 가구에서는 만두, 너겟, 찌개 양념 등 집밥을 만들기 위한 상품을 구매한다.

이는 11번가의 주 고객층이 자녀가 있는 3-40대이기 때문에 도출된 결과이다.

(1) 거래처별 매출액 비교



(2) 묶음수량 변곡점



상온 : 1 > 24개

- 단일상품 구매 건수가 제일 많다
- 24개 묶음상품 : 상온에서 가장 큰 비중을 차지하는 햇 반의 영향이 크다고 유추할 수 있다. 햇반 > '밥' 의 묶음수량 변곡점이 24개이다.

냉동 : 1개 = 단일상품

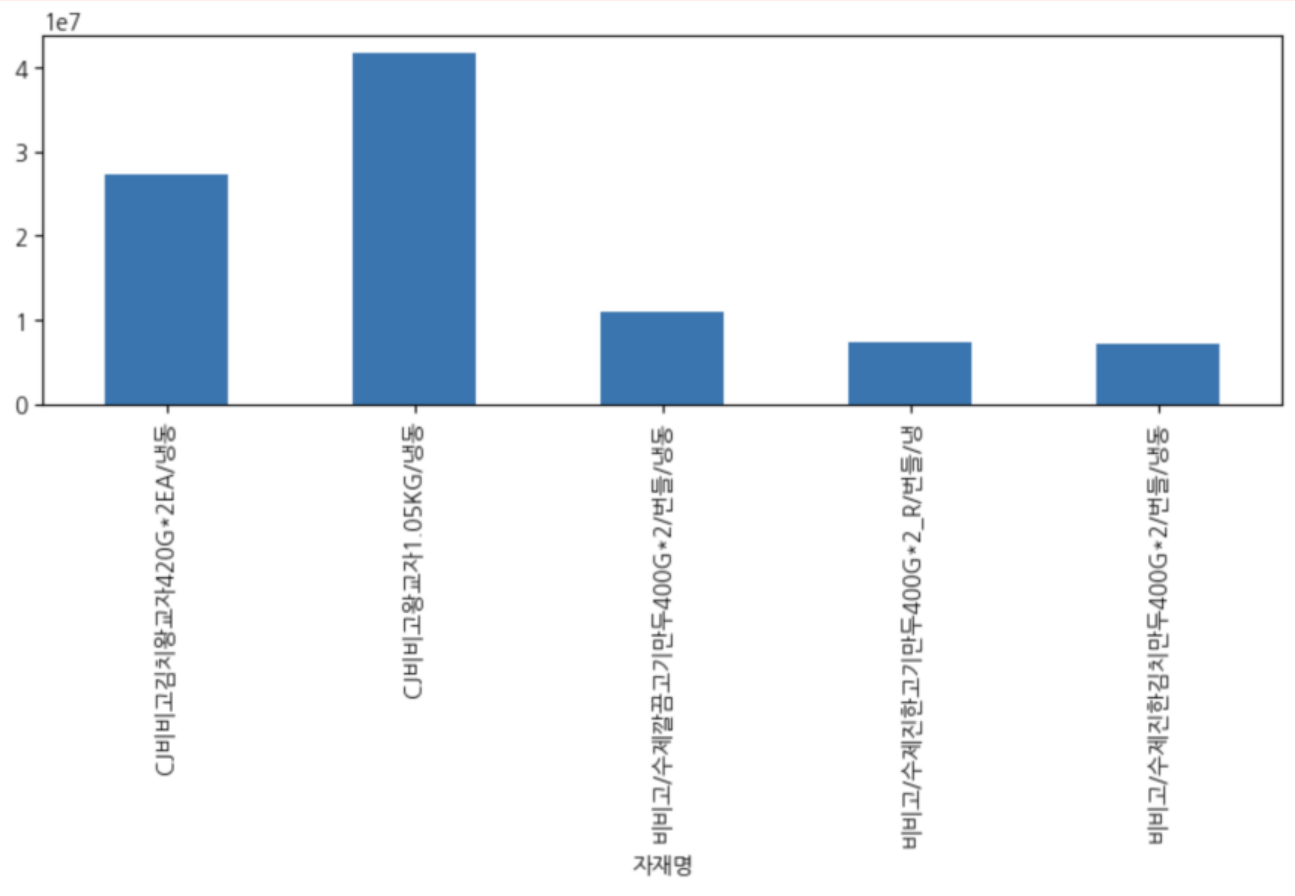
- 단일상품 구매건수가 제일 많다

냉장 : 3 > 4개

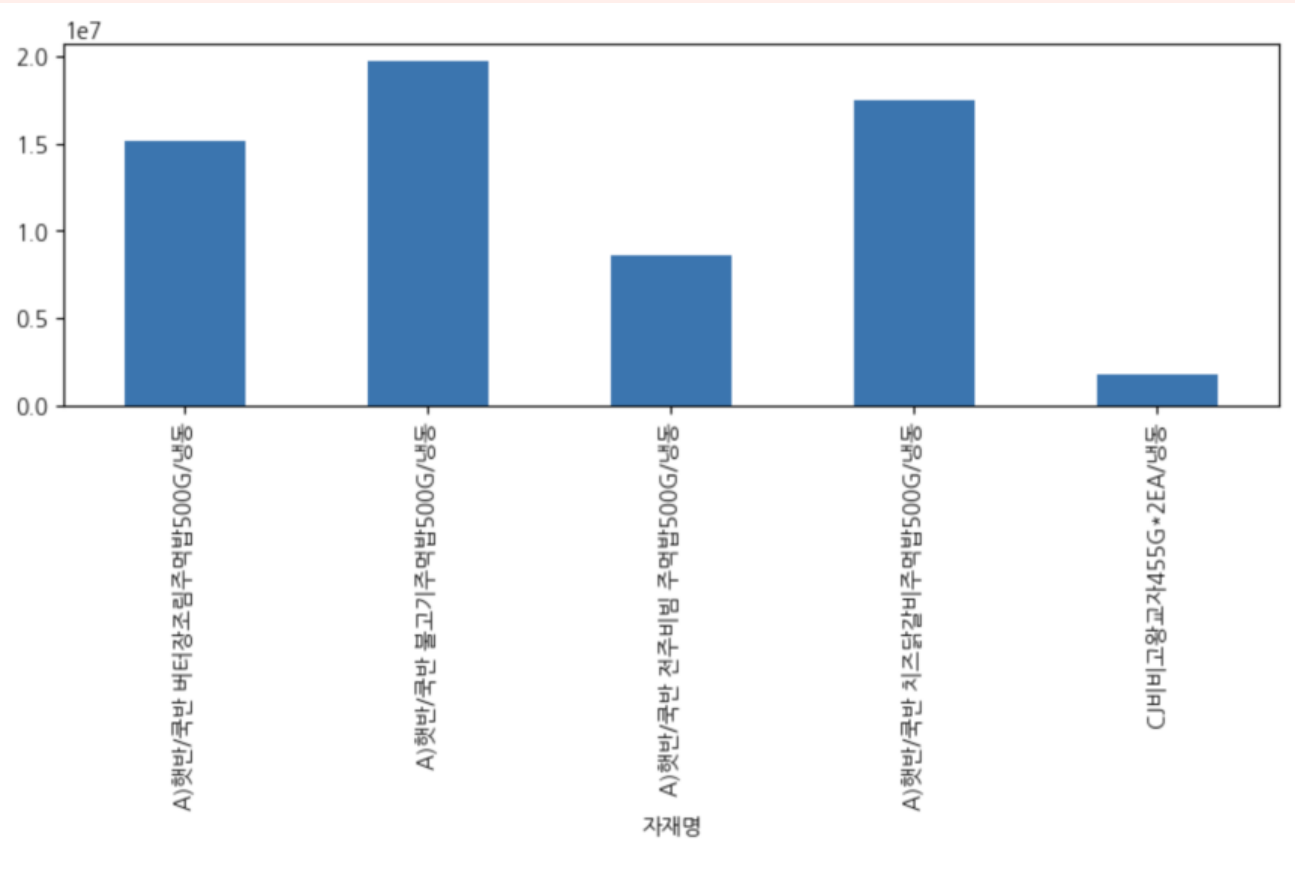
- 다른 마감구분과 다르게 3개가 변곡점
- 냉장 상품에서 가장 큰 비중을 차지하는 다담 때문에 그러하다. 다담의 묶음수량 변곡점이 3~4개이다.

5) 거래처주문번호

동일 구매 판매량 상위 자재와 함께 구매된 다른 상품들



비비고



햇반

함께 구매되는 자재들은 복합상품으로 묶어서 판매하면 더 큰 이익을 불러올 수 있다.

6) 사은품 여부

(1) 매출 평균 비교

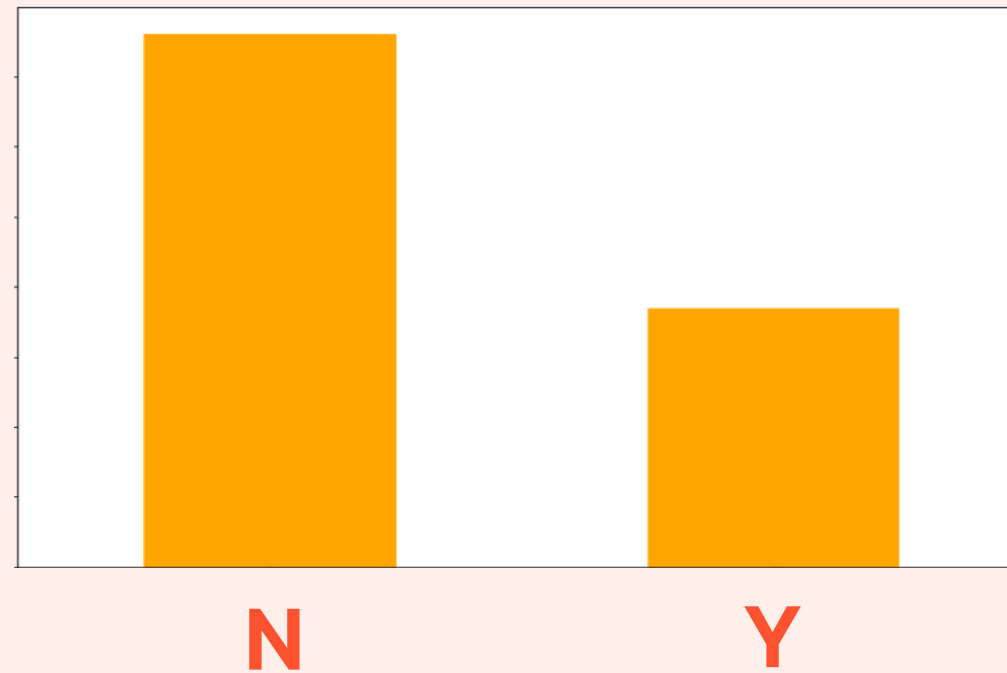


사은품 없는 상품

v

사은품이 동봉되는 상품

(2) 사은품과 묶여서 판매된 적이 있는 제품의 매출 평균 비교



사은품을 포함시키는 것은
오히려 부정적인 효과를 가진다

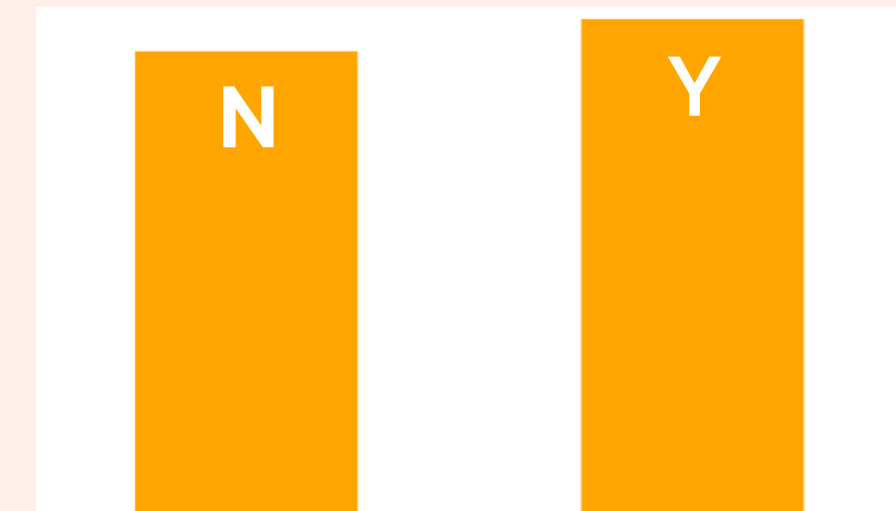
(3) 판매량 차이에 따른 매출 평균

판매량이 적은 상품 : 시럽



F-value: 0.001
p-value: 0.97

판매량이 많은 상품 : 비비고



F-value: 5.39
p-value: 0.02

판매량이 적은 상품에 사은품을 붙였을 때
매출에 유의미한 영향을 끼쳤다.

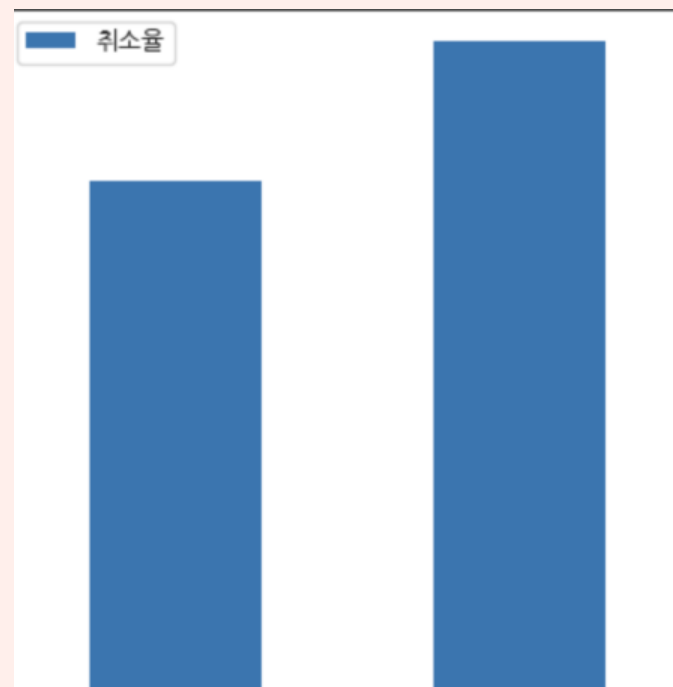
track1 : 이쿠조

3. Sub Data 데이터 분석 .

Sub Data : 매출이 발생하지 않은 데이터

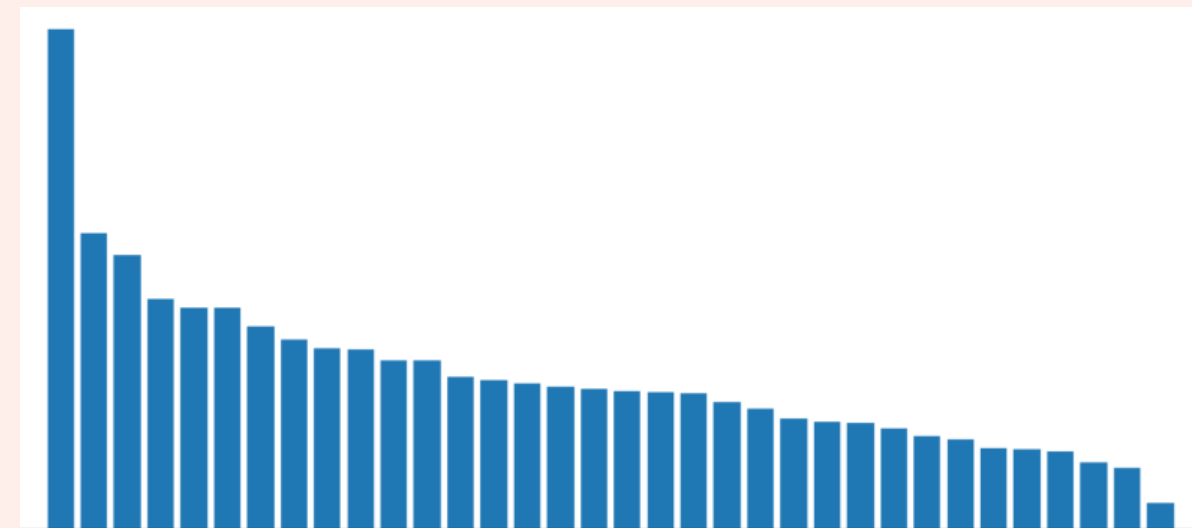
1) sub1 : 주문취소&취소완료

(1) 거래처 주문 취소 비율



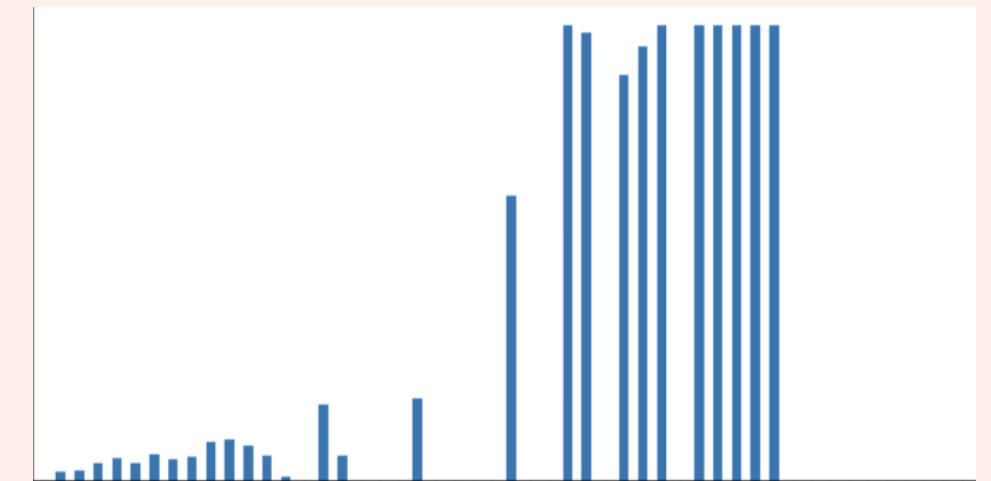
네이버 > 11번가

(2) 자재명_분류 주문 취소 비율



백설 > 미정당 > 맥스봉 > 소금

(3) 예상출고소요기간별 주문 취소 비율

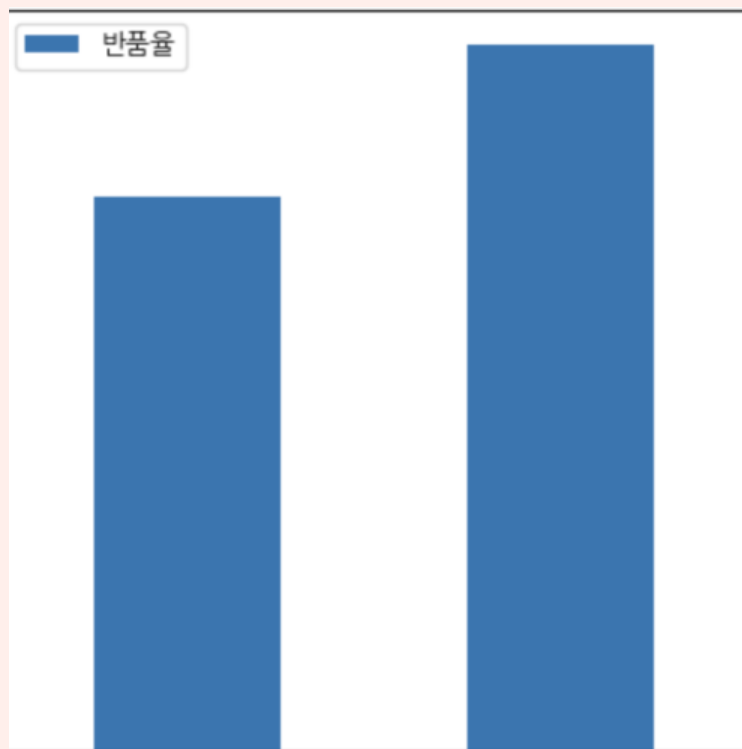


max : 62 days

예상출고소요기간이 길어질수록
주문취소 비율이 높아진다

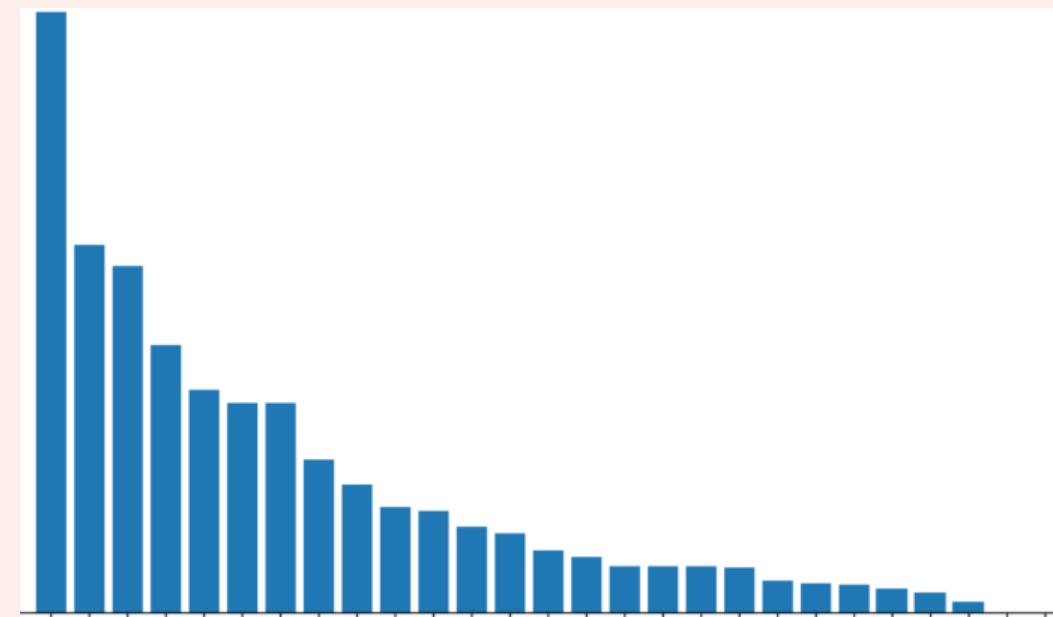
2) sub2 : 반품처리중 or 반품완료 or 반품신청

(1) 거래처 반품율



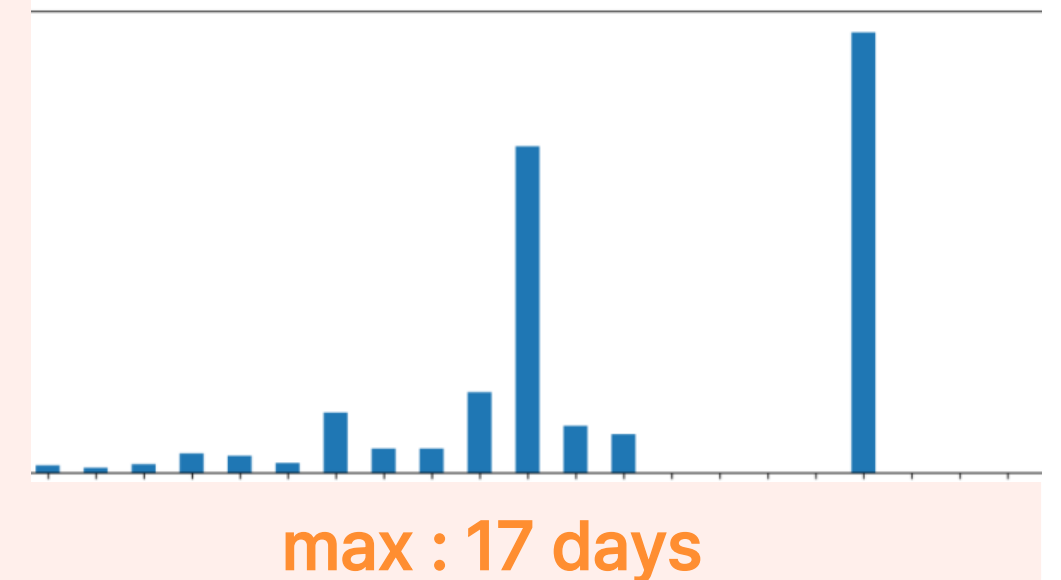
네이버 > 11번가

(2) 자재명_분류 반품율



행복한콩 > 삼호 > 믹스 > 쿡킷

(3) 실제출고소요기간별 반품율



실제출고소요기간이 길어질수록
반품율이 높아진다

마케팅 전략 1

거래처마다 다른 상품에 포커싱하라

분석 결과

- 거래처 판매량 순위
- 마감구분별 거래처 비율
- 마감구분별 자재명_분류 판매량 순위

네이버 : 핫반

- 네이버의 주 고객층이 젊은 2-30대

- 네이버 판매량 : 핫반 >> 비비고 >> 백설
- 상온 상품 판매량 1위는 핫반이다.
- 밥을 직접 지어먹는 핵가족이 아닌 1인 혹은 2인 가구라는 뜻

11번가 : 비비고, 다담

- 11번가의 주 고객층이 자녀가 있는 3-40대

- 11번가 판매량 : 비비고 >> 다담
- 냉동 식품 중 비비고가 차지하는 비중이 가장 높다.
- 자녀가 있는 4인 이상의 가구에서는 만두, 너겟 등 집밥을 구성하기 위한 상품을 구매한다.

마케팅 전략 2

할인행사 진행 시 행사 간격과 할인 금액 적극 고려

분석 결과

- 주문일별 핫반 평균 가격 그래프
- 주문일별 핫반 매출액 그래프
- 마감구분별 거래처 비율
- 거래처 판매량 순위

네이버 핫반 할인 프로모션

할인금액은 18,000원대로, 행사 간격은 24일에 한번으로!

< 할인 금액 >

핫반 평균 가격 그래프를 살펴보면 가격이 18000원대 초반까지 내려갔을 때 매우 유의미한 매출액 증대 효과

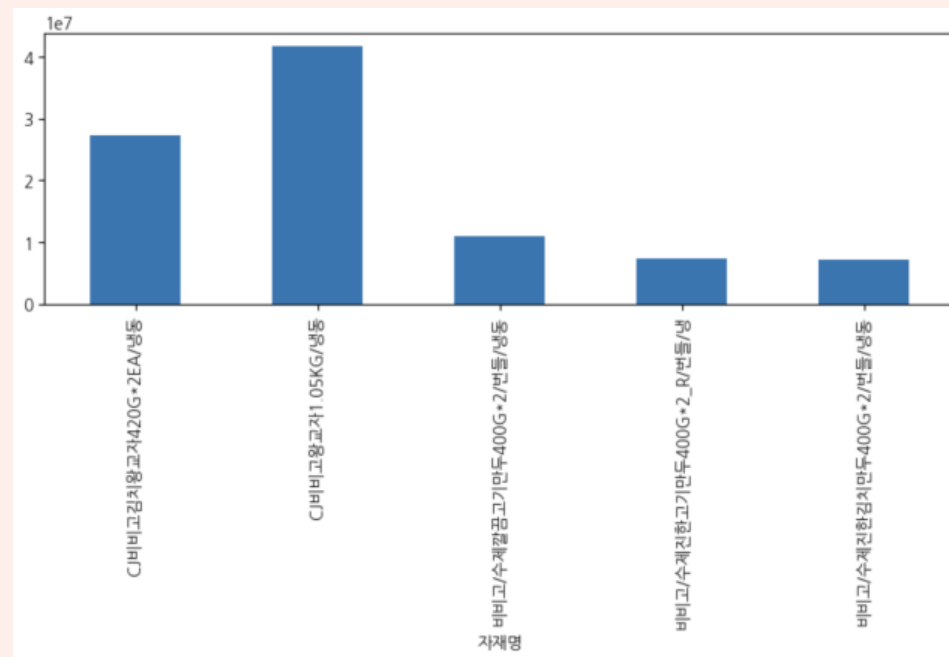
< 행사 간격 >

핫반 가격 그래프를 살펴보면 할인행사가 4~7일 간격으로 진행되었지만, 앞선 분석의 3일에 비하면 현저히 저조한 실적을 보였다. 따라서 행사 간격은 분석 결과의 24일로 둔다.

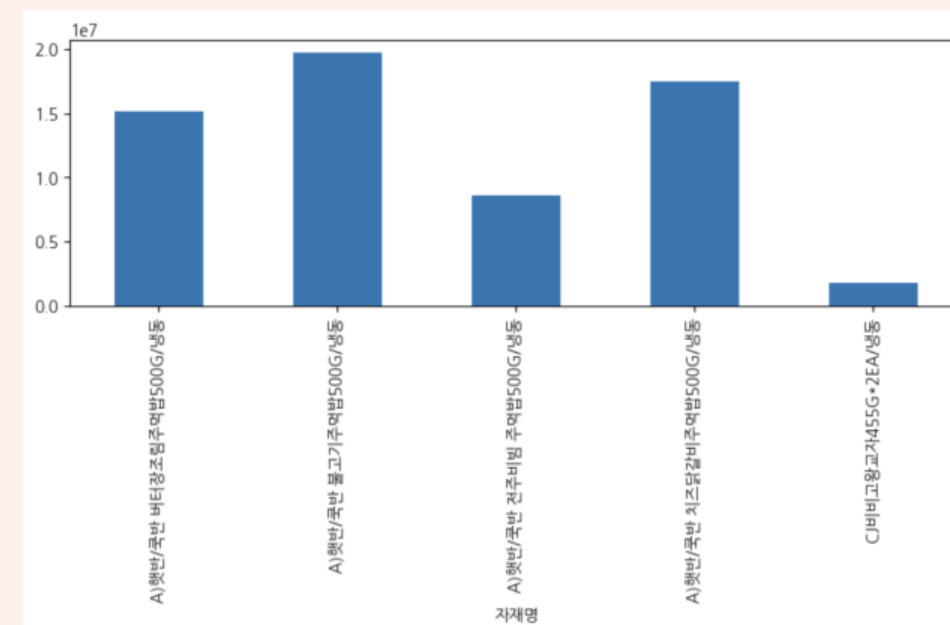
마케팅 전략 3

동시 구매 데이터를 통한 복합상품 마케팅 전략

동일 구매 판매량 상위 자재와 함께 구매된 다른 상품들



비비고



햇반

한 사람이 여러 상품을 동시 구매한 경우를 '동시 구매' 데이터라고 칭할 때, 우리는 동시 구매 데이터를 통해 어떤 상품이 다른 상품과 동시에, 즉 함께 구매되었는지를 살펴볼 수 있다.

네이버와 11번가의 거래처 특징을 고려해 봤을 때 비비고와 햇반, 두 가지 자재를 예시로 들었다.

사람들이 함께 구매하는 제품들이라는 것 으로부터 해당 상품들을 복합상품으로 묶었을 때의 수요를 확신할 수 있다.

마케팅 전략 4 판매량이 높은 상품에 사은품을 붙이자

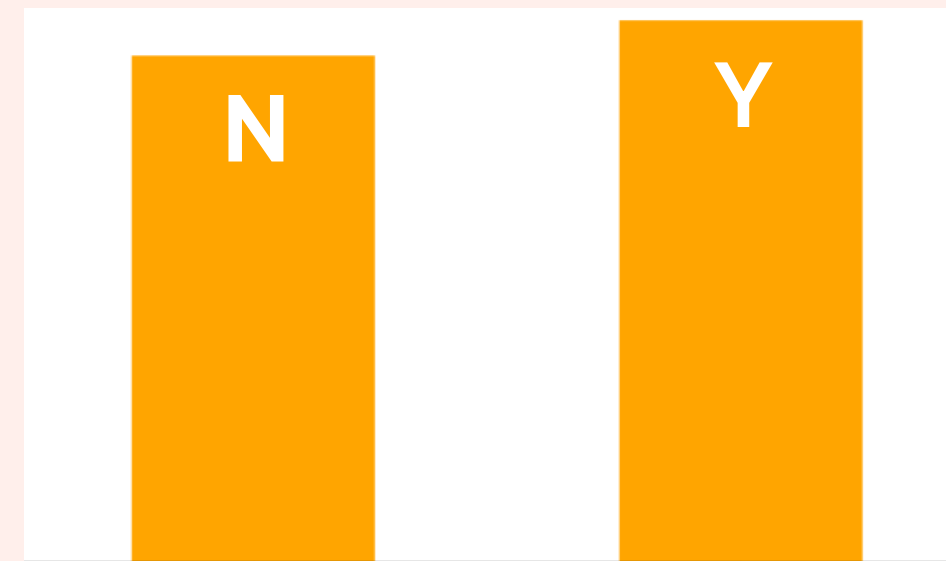
판매량 차이에 따른 매출 평균

판매량이 적은 상품 : 시럽



F-value: 0.001
p-value: 0.97

판매량이 많은 상품 : 비비고

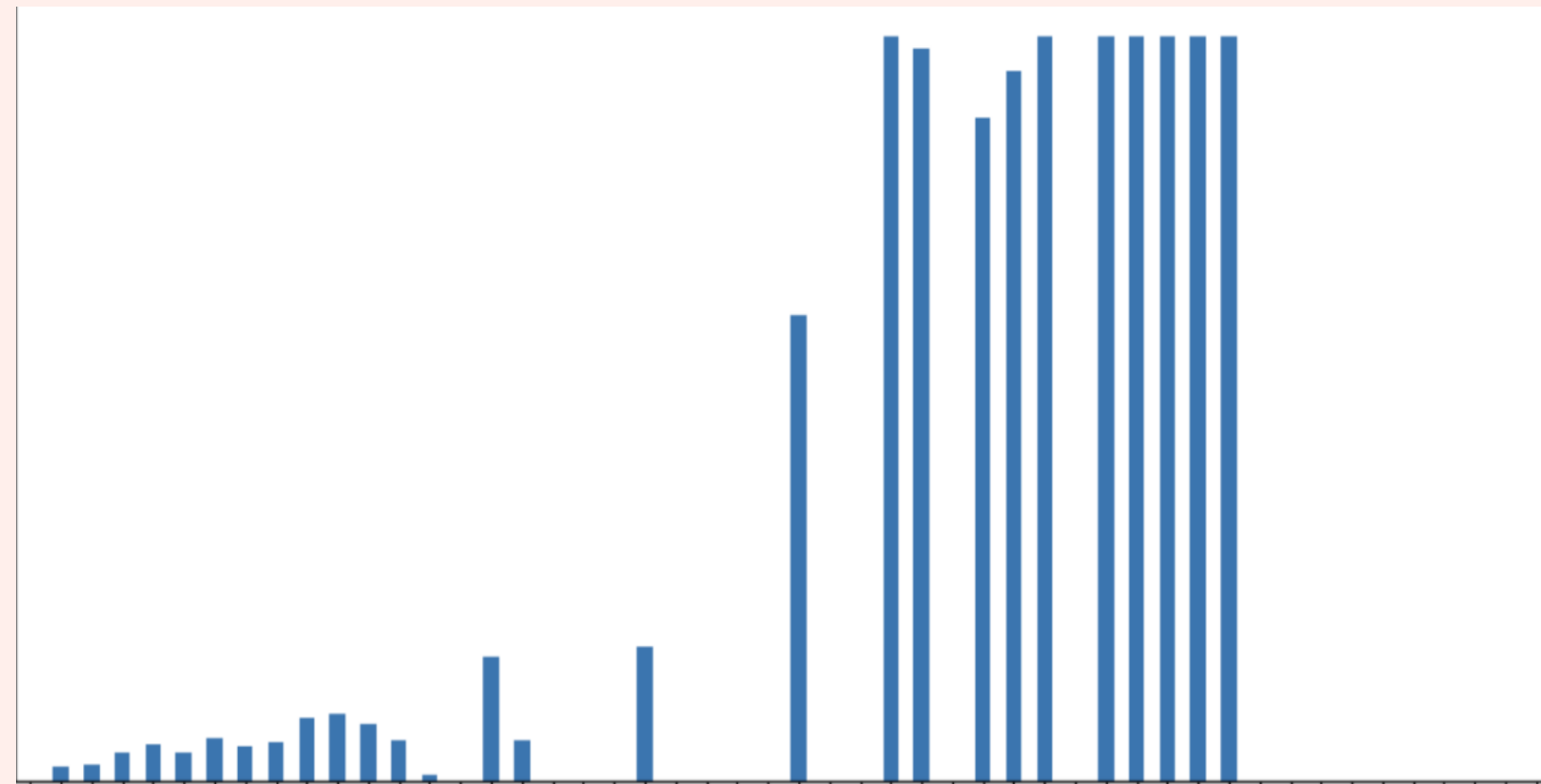


F-value: 5.39
p-value: 0.02

세일즈 전략 5

예상출고소요기간을 최대한 줄일 수
있는 배송 방법을 고안하자

예상출고소요기간별 주문 취소 비율



max : 62 days

4. 제언 및 향후 방향성 .