

회의 요약

1. 개인 경험에서 아이디어 제안

- 상준: 제품이 배달 시 파손되어 교환받은 경험 공유
- 현근: 악질 리뷰 및 유통매장 확보 관련 기사
- 한별: 리뷰를 보며 악플 유형 분류 가능성 제시

2. 주제 선정 과정 (ChatGPT 활용)

주제 후보

- 🎯 배송 지연 원인 분석 및 개선 전략 도출
- 📊 판매자 성과 분석과 평가 지표 수립

=> 결론: 이커머스 사용 시 가장 불만이 많았던 '배송'에 대한 분석을 진행하기로 결정

3. 역할 분담 및 조사 방법

- 각자 고민한 평가 방법 선정 및 개별 조사 진행하기로 합의

4. 개별 조사 내용 및 역할

한별

- 배송지연일자 분석
- 최대 188일, 평균 8.8일, 표준편차 13.94, 사분위수 10일
- 롱테일 분포, 3/4일 정도가 일반적 => 이상치가 있음을 추측함

상준

- 전체 평균 배송 지연일 분석
- 지역별 배송 지연 현황 파악 (상파울루 1800건 1위, 2위 리우)

민지

- 배송 지연이 별점에 미치는 영향 분석

재룡

- 역할 분배 및 일정 제안
- 불필요한 칼럼(이메일, 이름 등) 제거

현근

- 'boleto' 등 모르는 부분 조사
- 일정 관리 및 작업 분배

5. 분석 방안 및 피드백

- 상준: 기간별 비율로 데이터 시각화(지연배송 비율 조사에 초점)
- 한별: 배송지연일자 이상치 확인 및 기준 제시

- 민지: 배송 지연이 별점에 미치는 영향 및 별점별 반응 분석
- 재룡: 칼럼 정보 정리(누구나 이해하기 쉽게)
- 현근: 전체 일정 관리 및 작업 분배

추가 안내

각자 맡은 역할에 따라 조사 및 분석을 진행 / 다음 회의 전까지 개별 조사 결과 정리 및 공유
문의사항이나 추가 의견은 언제든지 공유해주세요!

0703작업 진행 예정 사항(이는 할당이 아닌 피드백에 따른 작업 제안 정리일 뿐 다른 작업을 제안하시거나 다른 분이 하셔도 되요!!!)

1. 민지: 평균 별점과 배송 지연 여부에 따른 별점의 차이 확인하기
=> 시간이 된다면 분야별로 사람들이 느끼는 배송 지연 여부와 평균별점 차이(쌀과 딸기의 배송 지연에 따른 불만족도는 다를 것)
2. 재룡: 칼럼에 담긴 정보 정리
3. 상준: 지역별 지연 배송 비율 정리하기
4. 현근: 지역에서 배송이 늦어지는 이상치 찾기 혹은 지역 외지 등을 찾아보기
5. 한별: 배송 지연 일자 이상치를 확인하고 배송지연에 대한 이상치 정의하기

얘기해볼 사항

- 현재 하고 있는 내용이 너무 정형화된 내용이 있어 다른 결과물들과 차이를 만들고 싶다면 어려울 것으로 보입니다. 이에 대해서 어떤 의견이든 환영입니다. 심지어 브라질에선 한국과 달리 김치를 먹지 않나봐요 같은 내용도 좋습니다.

저희의 목표는

1. 데이터 변수를 정의한다.
2. 데이터 변수를 가공하여 결과값을 냅니다.
3. 결과값을 통해 인사이트를 발견하고 이에 대해서 발표를 준비합니다.

0703

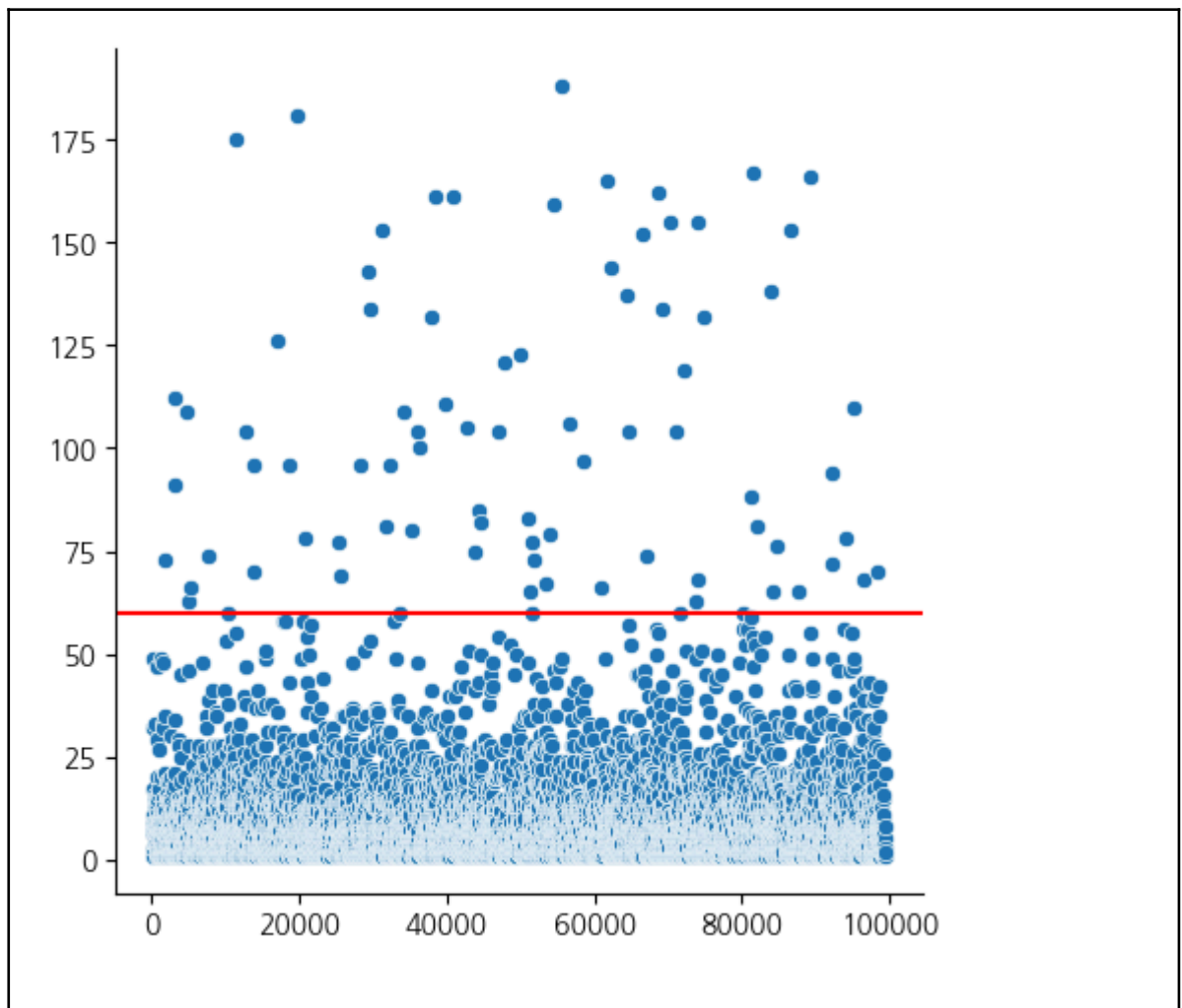
1. 오늘의 진행상황:
한별 :

한별

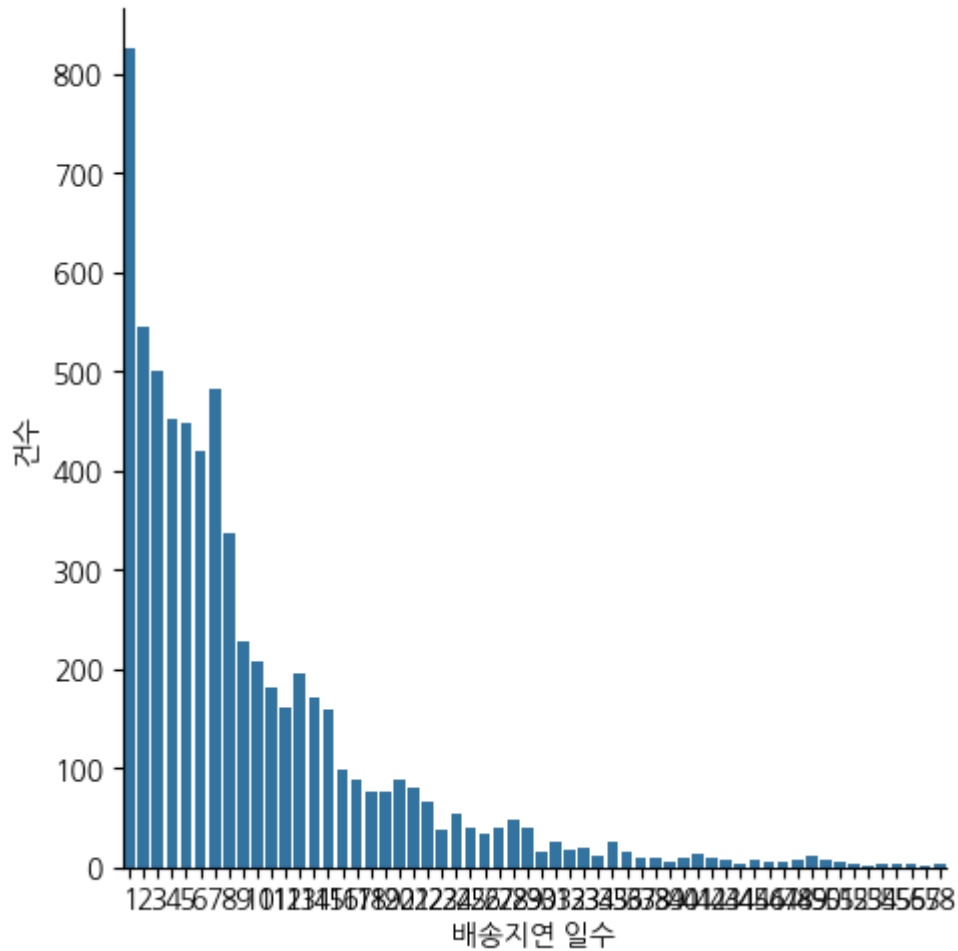
배송지연 수치

	이상치 포함	이상치 제거
배송지연 건수	6535건	6450건
배송지연 비율	6.57%	6.49%
평균 지연 일수	10.62일	9.4일
최대 지연 일수	188일	58일
표준편차	14.64일	9.23일
사분위수	10일	10일

-> 현근님 제안으로 이상치를 '**59일 이상**'으로 수정
(스캐터 플롯으로 확인 + 1건만 발생한 지연일만 선택)



배송지연 일수에 대한 건수



셀러별 지연 일수 Top 10

순위	셀러 ID	지연 건수 (수)	지연 일수 (수)	지역	파는 물건
1	df683dfda87bf71ac3fc63063bfa369d	1	167	farroupilha - RS	[가구/인테리어]
2	e83c76265fc54bf41eac728805e4da77	1	159	aracatuba - SP	[자동차용품]
3	4fb41dff7c50136976d1a5cf004a42e2	1	134	feira de santana - BA	[시계/선물]
4	3340ef1913fb70d28420f6ceb685c339	1	132	maringa - PR	[컴퓨터/엑세서리]
5	1782ce2db72c3dda627d5e7c898104e1	1	132	taboao da serra - SP	[스포츠/레저, 시계/선물]

					없음]
6	eebb3372362aa9a46975164b ed19a7e7	1	123	sao paulo - SP	[통신기기]
7	5d378b73ab7dd6f0418d743e5 dcb0bd1	1	100	sao paulo - SP	[없음]
8	7a91bf945c6fae0779f1c61ce9 7fe45c	1	97	mesquita - RJ	[뷰티/건강]
9	ad420dd0c4f92f8af951ac24b8 6d0cf5	2	96.5	fronteira - MG	[뷰티/건강, 장난감]
10	e09887ca8c7bf8a4621ce4818 20414ef	1	91	uba MG	[가구/인테 리어]
		합계 : 11건	평균 : 123.15일		

- 전체 주문 건수에서 Top 10 셀러 지연 일수의 비율은 0.168%에 불과하지만 지연 일수는 평균 123.15일에 달해 소수의 셀러가 긴 배송지연을 만드는 것을 알 수 있다.
- 지역과 지연 일수의 상관관계는 찾지 못했다.
- 불량스러운 셀러들은 뭘 팔길래 이럴까? 가구나 자동차 등 큰 물건을 팔고 있어서 그럴다.

셀러별 지연 건수 Top 10

순위	셀러 ID	지연 건수 (수)	전체 배송지연 건수에서 차지하는 비율 (이상치 포함함)(%)	평균 지연일 (일)	지역	파는 물건
1	4a3ca9315b74 4ce9f8e937436 1493884	172	2.63	12.04	ibitinga - SP	[침구/식 탁/욕실용 품, 집 꾸미기, 가구/인테 리어]
2	1f50f920176fa8 1dab994f90235 23100	125	1.91	12.36	sao jose do rio preto - SP	[정원용품 , 표지판 및 안전용품, 건축/공구 - 공구]

3	4869f7a5dfa27 7a7dca6462dcf 3b52b2	111	1.70	9.98	guariba - SP	[시계/선물, 기발한 상품, 뷰티/건강 , 스포츠/레저]
4	7c67e1448b00f 6e969d365cea 6b010ab	106	1.62	12.21	itaquaque cetuba - SP	[사무용 가구, 가구/인테리어, 스포츠/레저]
5	955fee9216a65 b617aa5c0531 780ce60	89	1.36	8.38	sao paulo - SP	[스포츠/ 레저, 건축/공구 - 조명]
6	6560211a19b4 7992c3666cc4 4a7e94c0	89	1.36	8.51	sao paulo - SP	[패션 가방/액세서리, 시계/선물 , 스포츠/레저]
7	ea8482cd71df3 c1969d7b9473f f13abc	85	1.30	9.35	sao paulo - SP	[통신기기 , 자동차용품]
8	da8622b14eb1 7ae2831f4ac5b 9dab84a	83	1.27	12.75	piracicaba - SP	[침구/식탁/욕실용품, 뷰티/건강 , 가구/인테리어]
9	cc419e0650a3 c5ba77189a18 82b7556a	76	1.16	8.93	santo andre - SP	[뷰티/건강, 통신기기, 향수, 식품]
10	8b321bb66939 2f5163d04c59e 235e066	72	1.10	11.10	sao paulo - SP	[전자제품 , 냉난방기기, 악기]
		합계 : 1008건	합계 : 15.42%	평균 10.56일		

- 전체 주문 건수에서 **Top 10** 셀러 자연 건수는 총 **1008건**이며 비율은 **15.42%**로 유의미하게 높아, 소수의 셀러가 다수의 자연 건수를 만들고 있음을 알 수 있다.
- 자연 건수가 많은 셀러가 긴 배송지연을 만드는 것을 알 수 있다.
- 자연이 다수 발생하는 주는 상파울로다.
- 불량스러운 셀러들은 뭘 팔길래 이럴까? 가구나 자동차 등 큰 물건을 팔고 있어서 그렇다.

해결방안

- 부피가 크고 무거운(가구, 전자제품, 자동차 등) 물건을 파는 셀러에게는 고객이 기다림을 예측할 수 있도록 예상배송기간을 더 늘려 고지하도록 안내한다.
 - 부피가 작은 불닭볶음면을 팔라고 안내해본다
- 소수의 셀러가 다수의 긴 배송지연을 만들고 있으므로, 배송지연 건수가 높은 상위 **10명**의 셀러에게는 아래의 패널티가 주어진다.
 - 마켓 별점을 떨어뜨리기
 - **10건 이상 배송 지연** -> 마켓 **3일** 정지
 - **20건 이상 배송 지연** -> 마켓 **일주일** 정지
 - **50건 이상 배송 지연** -> 마켓 **1달** 정지, 블랙리스트 명기 * 블랙리스트에 오를 시 추후 올리스트에서 새로운 마켓을 개장할 수 없음
 - **100건 이상 배송 지연** -> 마켓 정지
- 배송지연이 다수 발생하는 상파울로 도시의 상파울로 주에 빠르게 배송할 수 있는 물류센터를 찾아 삼양과 물류센터간 **B2B**협업을 한다.

1. "빠른 삼양" 배지 제도 도입

- 설명: 일정 기간 이상 배송지연이 없는 셀러 또는 배송정확도 **98%** 이상 셀러에게 "빠른 삼양 인증 배지" 부여.
- 효과:
 - 고객은 믿고 구매할 수 있음.
 - 성실한 셀러에게는 동기부여 제공.
- 삼양의 신뢰도 강화: 삼양 제품은 신뢰할 수 있는 셀러에게 배송됨을 각인.

2. 예상 배송시간 구간별 안내 강화 (카테고리별)

- 설명: 가구/전자제품/자동차처럼 배송지연 가능성이 큰 카테고리는 상품 상세페이지에 "예상 배송 기간: **5~20일 (부피 큰 제품)**" 식으로 구간 예측 제공.
- 효과:
 - 고객의 기대치를 조절 → 클레임 감소
 - 플랫폼과 셀러에 대한 불신 해소
- 삼양 제품은 "빠른 배송 구간" 강조: "당일/익일 발송" 표기 가능

3. 지연 고객 대상 리워드 프로그램

- 설명: 삼양 제품 구매 고객 중 예상일보다 **3일 이상** 배송 지연 시 포인트/할인쿠폰 제공.
 - 효과:
 - 고객 불만 완화
 - 브랜드 이미지 회복
 - 지연 셀러로 인한 브랜드 이미지 손상 최소화
-

4. 삼양 인증 풀필먼트 센터 또는 셀러 풀 관리

- 설명: 신뢰도 높은 셀러만 삼양 제품 판매 가능하게 하거나, 삼양 전용 풀필먼트 창고 도입 → **Olist의 Top 셀러**와 제휴
 - 효과:
 - 배송지연 최소화
 - 삼양 제품의 브랜드 가치 보존
 - 장기적으로 “올리스트 내 프리미엄 브랜드” 포지셔닝
-

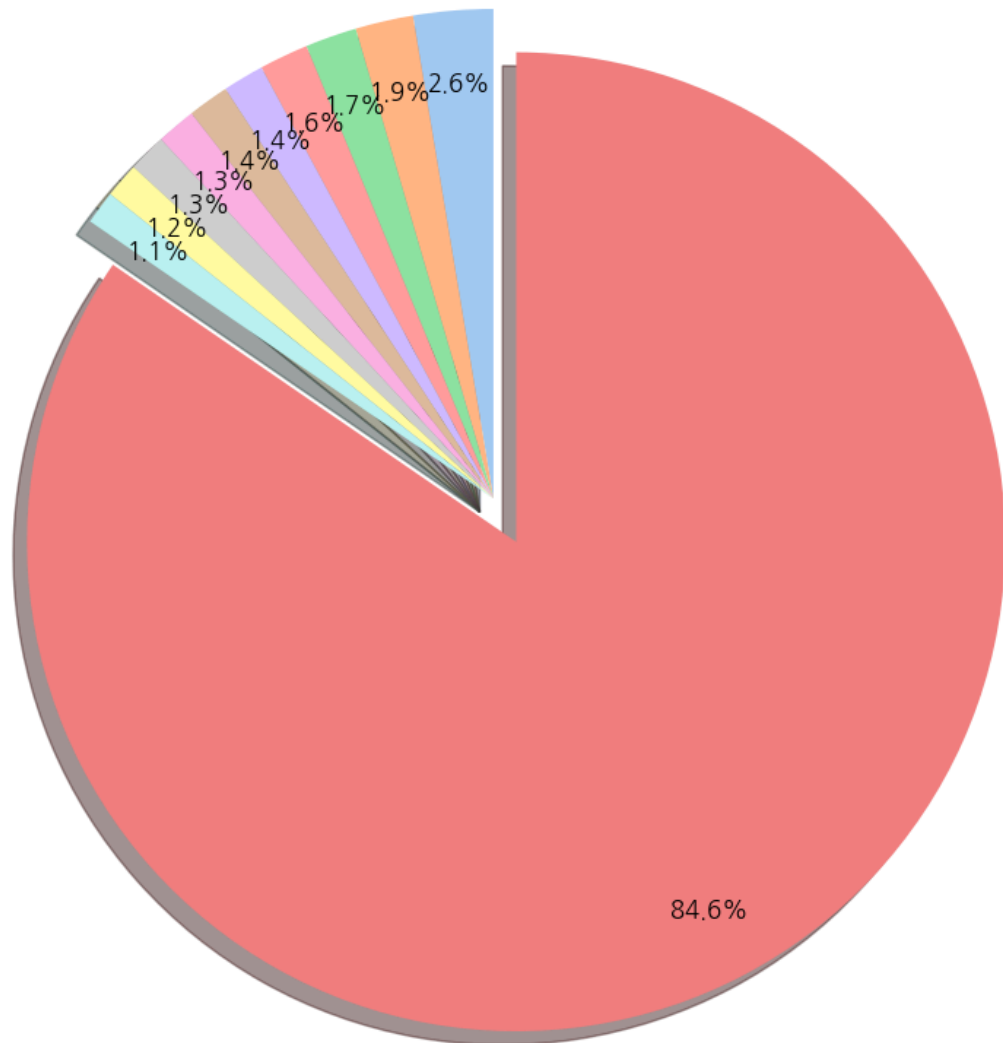
5. “불량 셀러 추적” 공동 감시 시스템 제안

- 설명: 고객 리뷰, 배송지연 데이터, 반품율 등을 기반으로 셀러 점수 시스템 개발하여 **Olist**에 제안
 - 삼양이 주도해서 개선방안 제안 → 플랫폼 파트너십 강화
 - 향후 **B2B** 협업/우선노출 등에 유리한 입지 확보
-

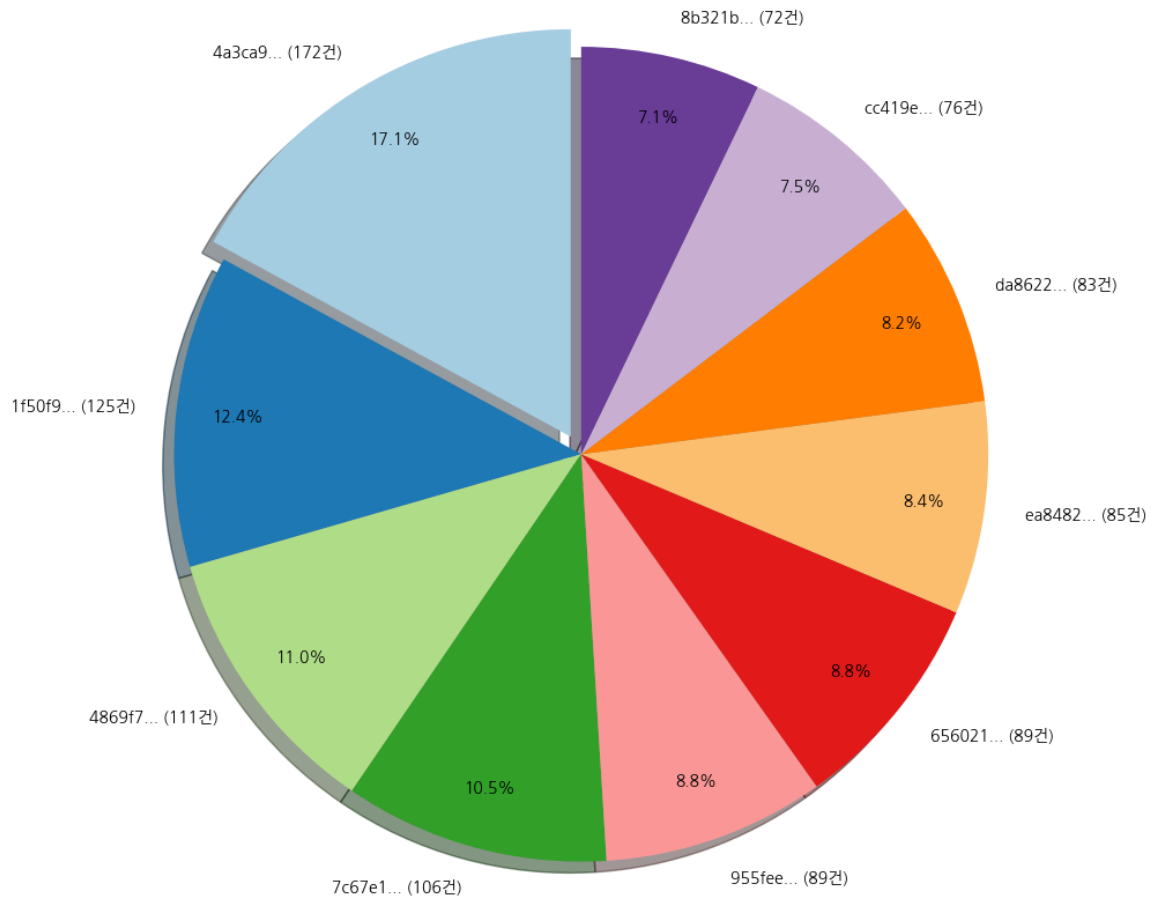
6. 고객 안심 메시지 시스템

- 설명: 지연 발생 시 자동화된 메시지로 배송 현황, 예상 추가 지연 기간, 사과 메시지 전달
- 삼양 전용 템플릿으로 구성해 신뢰성 강조 (예: "불닭볶음면은 지금 열심히 배달 중입니다!")

전체 6535건 중 각 항목별 지연 건수 비율



Top 10 셀러 ID별 지연 건수 비율 (총 1008건)



민지 :

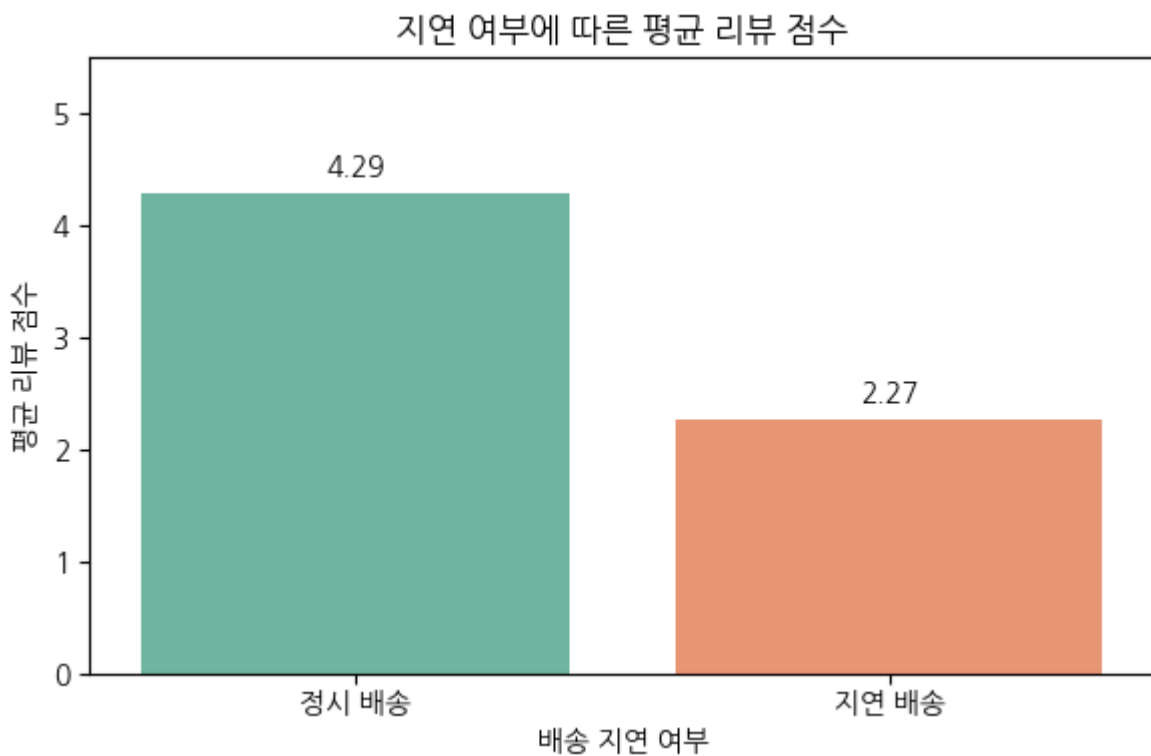
늦은 배송에 대한 리뷰 점수

* 늦은 배송 : ('order_estimated_delivery_date' +1 일) 보다 'order_delivered_customer_date'가 늦은 경우

리뷰 점수	개수
1 점	3,444
2 점	556
3 점	698
4 점	652
5 점	1,060
합 계	6,410

- "orders"와 "reviews"를 'order_id' 기준으로 병합한 DATA 기준
(숫자 디버깅 필요)

지연 여부에 따른 평균 리뷰 점수



- 상관계수 : -0.39144441509913175

▶ 지연 여부와 별점은 음의 상관관계를 갖는다.

p-value 통한 지연여부, 별점 간 상관관계 분석

예상 날짜 이후에 배송된 주문이 부정적인 리뷰(점수 ≤ 3)를 받을 가능성이 더 높은지 검토한다.

- 귀무가설(H_0): 배송 지연은 부정적인 리뷰 발생 가능성을 증가시키지 않는다.
- 대립가설(H_1): 배송 지연은 부정적인 리뷰 발생 가능성을 증가시킨다.

> 분석 결과

p-value : 0.00

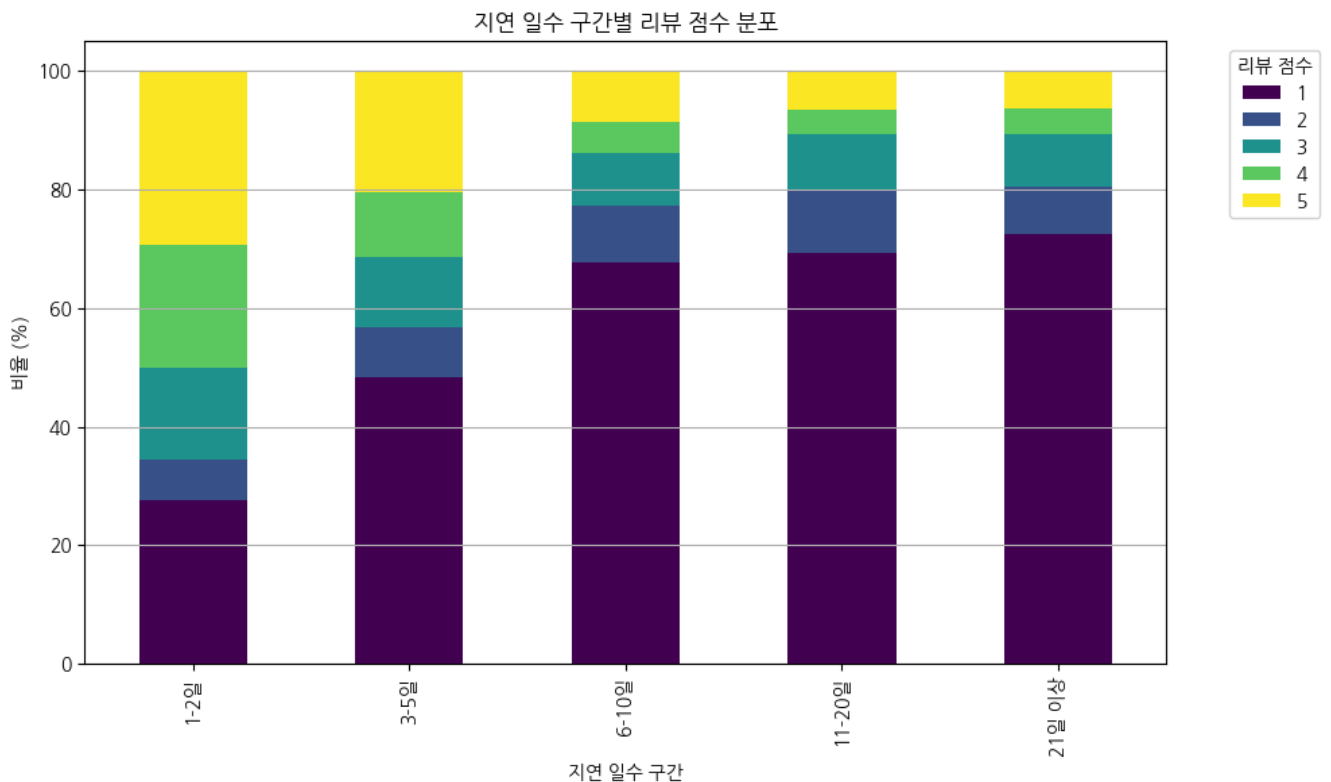
p-value (0.00)가 유의수준 (0.05)보다 작으므로 귀무 가설을 기각한다.

이는 납품 지연과 부정적 리뷰 사이에 통계적으로 유의미한 상관관계가 존재한다는 것을 의미한다.

▶ 즉, 납품 지연이 부정적인 리뷰를 받을 확률에 영향을 미친다고 볼 수 있다.

지연 일수 구간별 리뷰 점수 분포

- 지연 배송만 필터링 (1일 이상 58일 이하)



배송 지연 기간 구간별 건 수

- 리뷰 별점이 작성된 지연 배송 건만 카운트 (지연 기간 : 1일 이상 58일 이하)

배송 지연 기간	건 수
1-2일	537
3-5일	1,369
6-10일	1,642
11-20일	1,270
21-58일	688
합 계	5,506

브라질 시장에서 불닭볶음면 선택의 이유는?

출처 : KATI 농식품수출정보

* KATI : 한국농수산식품유통공사(aT)에서 운영하는 농식품 수출
관련 전문 정보 서비스

[브라질] 2024년 브라질 식품
트렌드

[https://www.kati.net/board/boardView.do?board_s
eq=100795&menu_dept=34](https://www.kati.net/board/boardView.do?board_seq=100795&menu_dept=34)

출처 : EBN 산업경제신문


삼양식품 불닭, 틱톡 팔로워 100만 돌파... 글로벌 Z세대 사로잡은 '매운맛
콘텐츠'

<https://www.ebn.co.kr/news/articleView.html?idxno=1666658>

결론

납품 지연이 부정적인 리뷰를 받을 가능성을 크게 높인다는 결론을 도출할 수 있다. 이러한
분석 결과는 다음 사항들의 중요성을 강조하며 고객 만족을 위한 개선이 필요하다.

- 정확하고 시기적절한 배송 : 고객과의 신뢰이자 고객 만족의 핵심
- 고객 기대치 관리 : 현실적인 배송 일정을 제시하여 배송 지연 관련 실망을 줄여야 함
- 지연 발생 시 알림 및 그에 대한 보상 제공 : 문제가 생겼을 때 고객과 소통하고 적절한
조치를 취할 것

 코드 정리

[https://colab.research.google.com/drive/1WfJKxdPWOJOGowGo57t9hKywJlzUaK0q?usp=
sharing](https://colab.research.google.com/drive/1WfJKxdPWOJOGowGo57t9hKywJlzUaK0q?usp=sharing)

안재룡

[1] " 배송 단계별로 (준비, 운송) 어디에서 지연이 주로 발생하는가? "

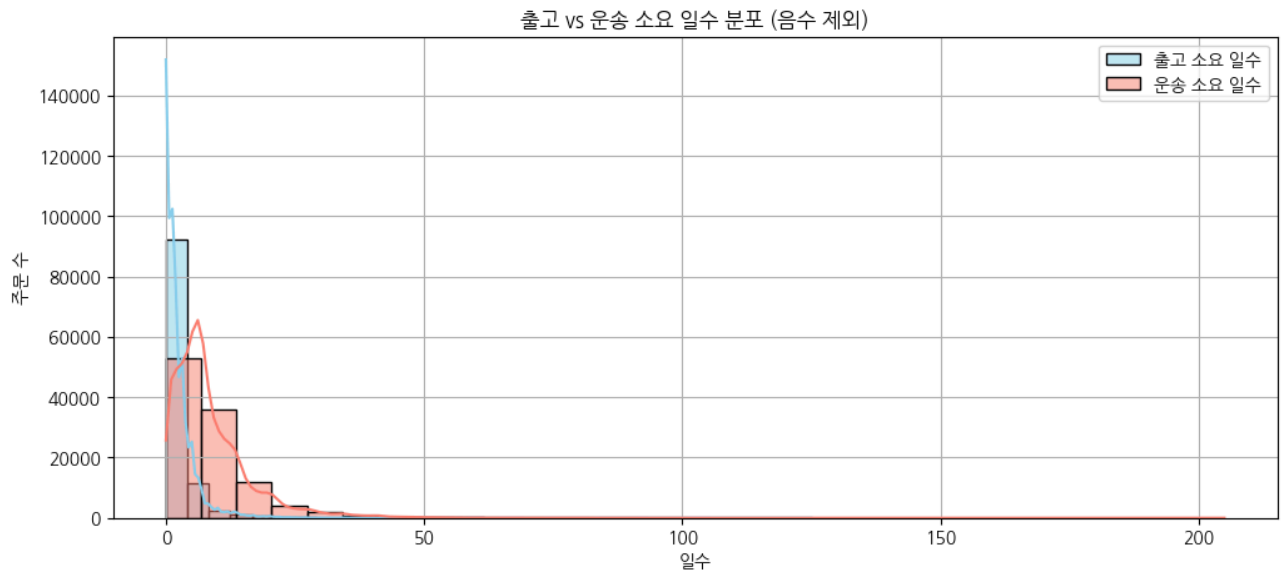
▼ 활용 데이터

(orders, orders_items)

< 출고, 운송 소요 일수 분포 및 지연률 >

→ 출고 지연률: 9.32% (출고 예정일을 넘긴 비율)

운송 지연률: 7.91% (운송 예정일을 넘긴 비율)



< 더 오래 걸리는 출고 vs 운송 >

📌 출고 vs 운송 지연 일수 통계 비교

	shipping_delay_days	delivery_delay_days
count	110180.000000	110180.000000
mean	2.342294	8.733954
std	3.590279	8.612080
min	-172.000000	-17.000000
25%	0.000000	4.000000
50%	1.000000	7.000000
75%	3.000000	11.000000
max	125.000000	205.000000

< 각 단계별 '이상 지연' 건수를 확인 >

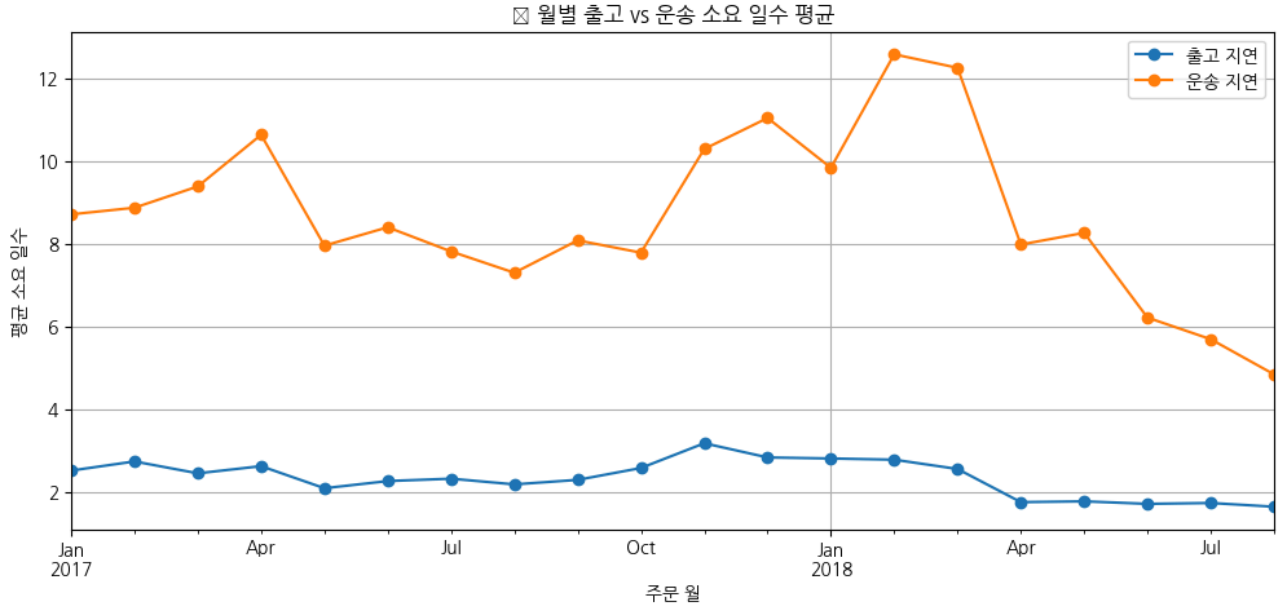
📦 출고 지연 5일 이상 건수: 11380건

🚚 운송 지연 5일 이상 건수: 66776건

📦 출고 지연 비율 (>5일): 10.33%

🚚 운송 지연 비율 (>5일): 60.61%

< 시간흐름에 따른 추세 분석 : 월별 >



(*현근님 요청 : 2017- 1 이후로 표시)

< 출고지연과 운송지연의 상관관계 >

	shipping_delay_days	delivery_delay_days
shipping_delay_days	1.000000	0.026224
delivery_delay_days	0.026224	1.000000

상관계수	해석
0.0 ~ ±0.1	전혀 상관없음 or 매우 약함
±0.1 ~ ±0.3	약한 상관관계
±0.3 ~ ±0.7	중간 정도 상관관계
±0.7 ~ ±1.0	강한 상관관계

< 결론 >

“ 고객 불만의 주요 원인 중 하나인 배송 지연 문제의 원인을 파악하기 위해, 출고 지연과 운송 지연이 어떤 관계에 있는지 정량적으로 분석함.

출고 지연 일수와 운송 지연 일수 간 **상관계수: 약 0.026 (2.6%)**

→ 선형적인 상관관계는 거의 없음

→ 둘은 독립적인 변수일 가능성 있음

하지만, 단순히 일수 기준의 상관관계만으로 두 변수 간의 관계를 판단하기는 어렵기 때문에,

출고 지연 여부와 운송 지연 여부 간의 관계를 확인하기 위해 **카이제곱 독립성 검정**을
수행하기로함.

출고지연과 운송지연이 각각 독립적으로 존재한다면 거기에 대한 심화내용
결과적으로 seller 혹은 운송업체가 문제가 있는것인지 (뒷받침 이유) * **현근님 제안**

0703 작업 내용 및 회의 요약

상준: 주별에 따른 데이터 분석을 통해 지연율이 높은 주들과 낮은 주들의 그룹간의 관계가 있음을 확인

- 지연율이 높은 주
주 자체가 대규모이고, 빈부격차가 심한 경우가 많으며 동북부 쪽에 몰려 있음
또한 해당 지역들에서 많은 양의 물류량이 특징
- 지연율이 낮은 주
주 자체가 소규모이고, 사람이 살 수 있는 공간이 작아 유인 지역의 경우 인구 밀도가 높은 경우가 많음
=> 이후 결 내용을 맡아 현재 작업 진행중

민지: 지연 일수 구간별 리뷰 점수 분포에 따라 사람들의 배송 지연에 따른 만족도를 비교함

- 이전의 예측 대비 당일 도착까지 늦은 것으로 판정하고 있던 것을 확인하고 당일 도착까지 지연 배송으로 데이터 작업을 수정토록함
- 이후 수치 표준화 작업에 선도적으로 진행하고 있음
- p-value 통한 지연여부, 별점 간 상관관계 분석

한별: 어제 진행했던 지연에 대한 비율에 대한 데이터 분석을 지속

- 이상치의 경우 상위 25%기준에서 59일(빈도수가 1회가 되는 기준)으로 조장과 논의 후 결정
- 배송 지연에도 셀러에 따른 2080법칙이 적용됨을 확인하고자 하였고 이를 상위 지연 셀러 10명이 차지하는 비율로 확인함=> 다만 이러한 셀러들이 자동차, 매트릭스등 대형 화물인 경우가 많아 어려움이 있었던 것으로 확인됨

재룡: 내부적으로 생기는 문제 기초 현황 분석 및 해결 접근 방향

- 출고 지연과 배송지연을 나눠서 조사
- 출고/배송 지연 정의 및 파생변수로 지연률 요약, 지연 분포 시각화, 상관관계 분석
- 다만 현재 출고 지연과 배송 지연의 경우 약한 상관관계가 나와 다른 접근이 필요할 것으로 보임
- 카이제곱 검정으로 출고/배송 간의 유의미한 상관관계 증명 - 셀러의 신뢰도를 바탕으로 한 배송시스템에 대해 분석

현근: 각자 조원들이 느끼는 기술적 어려움을 해결하는 방향으로 접근함

- 조원들의 각자 접근하고 조사한 내용들을 정리하고 이를 융화시키는 방향으로 진행함
- 발표자료에 대한 기승전결을 정리하여 초안으로 작성함
- 모두의 아이디어를 버무려 삼양사의 불닭볶음면 공장 혹은 물류공장을 짓는 방향을 잡아봄

0704 작성 안재룡

[2] " 출고/배송 정말 상관관계가 없는가? 에 대한 검증 "

< 카이제곱 검정 >

1.1 교차표와 기대값 비교

출고 지연 여부	운송 지연 X	운송 지연 O	총합
정상 출고	93,675건	6,240건	99,915건
출고 지연	7,791건	2,474건	10,265건

“출고 지연과 운송 지연이 완전히 무관하다면, 이런 결과가 나오는 게 자연스러운건가?”

→ 이를 평가하기 위해 “기대값” 을 계산

출고 지연 여부	운송 지연 X (기대값)	운송 지연 O (기대값)
정상 출고	약 92,013건	약 7,902건
출고 지연	약 9,453건	약 812건

하지만 실제로는 운송 지연이 출고 지연된 주문에서 **2,474건**이나 발생했기 때문에, 기대값보다 무려 **3배** 이상 많은 지연이 발생한 것이고, 이는

출고 지연과 운송 지연이 무관하지 않다는 강력한 통계적 증거가 됨.

1.2 p-value = 0.0

검정 결과로 도출된 p-value는 **0.000... (≈ 0)** 으로,
이는 유의수준 **0.05**보다 훨씬 작기 때문에

귀무가설(H_0 : 둘은 무관함)은 기각되고,
출고 지연과 운송 지연 사이에 유의미한 관계가 존재한다고 판단.

1.3 지연률과 지연 정도 비교

- 출고 지연은 전체 주문의 ****9.3%****에서 발생하며, 평균 지연일은 **2.6일**
- 운송 지연은 **7.9%** 발생했지만, 평균 지연일은 **8.7일**로 체감은 훨씬 심각

즉, 알 수 있는 데이터는

- 출고 지연은 자주 발생
- 운송 지연은 상대적으로 적게 발생하지만, 피해는 더 큼

지역별 분석 결과

고객의 지역(customer_state)을 기준으로 운송 지연률을 분석한 결과,
일부 지역에서 평균보다 월등히 높은 운송 지연률이 관찰됨
이는 물류 인프라, 배송 거리, 지역별 파트너 차이 등의 구조적 문제가 영향을 미쳤을 가능성이 있음.

< 종합 결론 >

1. 시간 상의 단순 상관관계는 낮지만, 출고 지연과 운송 지연은 통계적으로 연관이 있음
2. 기대값과 실제값 간 차이를 통해 출고 지연이 운송 지연을 유의미하게 증가시킨다는 사실을 검증
3. 출고 지연은 빈도 중심의 문제, 운송 지연은 심각도 중심의 문제로 각각 다르게 대응할 필요
4. 추가적으로 지역별 운송 문제도 구조적으로 확인되고 있어 병행 개선이 필요

0704_작성(2)_안재룡

(1)

“어디서 배송이 늦어지는가?”

내부적으로 배송 단계별 어디에서 지연이 주로 발생하는가?

출고 지연 / 배송 지연의 값 (일로 파악) 파악 후 상관관계파악

지연일 수 들로 상관관계를 파악했을 때 선형적인 상관관계 X (독립적)

(2)

“출고 ↔ 배송, 연관 있는가?”

출고 / 배송 연관이 있는가?

심화 검증을 위해 카이제곱의 독립성 검정을 수행함

기대값과 실제값을 구해서 비교를 함

통계적 수치로 무관하지 않다는 것을 입증함

검정 결과로 도출된 **p-value**값으로 귀무가설(독립)은 기각

출고 지연과 운송 지연 사이 유의미한 관계가 존재한다고 판단

출고 지연이 배송 지연으로 이어지는 구조 확인

← 여기부터 진행

(3)

“지역별 물류 신뢰도 분석 필요성 확인”

입지선정 - 둘의 지연률 고려한 신뢰도 기반 접근 방식 필요성

배송 지연율만 보는게 아닌 출고 지연율도 고려해 셀러 신뢰도 판단

출고 지연률과 배송지연률 간의 상관계수는 **0.3818** 로 양의 상관관계 존재함 확인

p-value도 통계적으로 유의미

결과적, 출고 지연이 배송 지연으로 이어지는 경향이 있음을 보여주며, 배송지연만이

아닌 출고과정 역시 고객 만족도에 중요한 영향을 끼친다는 것을 보여줌

(4)

“셀러 신뢰도 파악을 바탕으로 한 활용”

셀러별 출고/배송 지연률을 계산한 결과로 셀러별 신뢰도 지수 값을 파악 및 활용

이것을 토대로 셀러 지역정보와 병합해서 지역별 신뢰도 분석

이것을 토대로 셀러 평균 리뷰 평점을 도출 및 상관관계를 파악.

[3] " 지역별 물류 신뢰도 분석 필요성 확인 "

[4] " 셸러 신뢰도 파악을 바탕으로 한 활용 "

0704 대본 초안

(1페이지)

안녕하세요 오늘할 발표는 올리스트라는 이커머스 기업을 통해 현재 k엔비디아로 사람들의 관심을 한번에 받고 있는 두유노의 주인공 불닭볶음면을 좀 더 효율적으로 판매하기 위한 전략을 소개하고자자리에 선 안재룡(상큼하게), 이한별(발랄하게)입니다.

(2페이지)

저희 조원들의 모습입니다.

저희 조는 현근짱, 재룡짱, 민지짱, 상준짱, 한별짱과 함께 진행하였습니다!!!!

모두 요로시쿠네~~~~

(3페이지)

저희는 기승전결을 동기, 데이터 분석 방안, 결론, 이후 해결방안이라는 구성으로 구상했습니다!
추가적으로 이후 다음 단계와 출처 코드와 자료 정리를 진행했어영~~~

(4페이지)

이번에는 모두들 왜 우리가 이런 주제를 선정했는 지 궁금할 거 같아요!

갑자기 이커머스에서 먹을거? 그것도 불닭? 너무 MZ한데 왜 이걸 하게되었을 까요? 그 이유부터 알려드리겠습니다.

(5페이지)

이치방! 생산을 위한 거점이 아닌 공장이 아니라 왜 물류일까요? 왜 불닭을 만드는 게 아니라
물류가 중요하다고 할까요?

그것은 바로바로!!! 불닭은 공장에서 쓰이는 기계 부품이 아니라 우리같이 소중한 사람들이 먹는
최종소비재 이기 때문이에요. 특히 식자재의 경우 빠른 배송이 만족도의 중요한 요소구
말이에요!(최대한 콧소리와 애교를 섞어서)

거기다가 생산 공장을 제작하는 것이 더 큰 효과를 가지는 경우는 자동차 부품, 엔진등과 같이
대형 자재들, 즉 부피가 크거나 무게가 무거워 운송이 어려운 경우지 라면 같이 개별 작은
제품이나 가벼운 것들은 별로 효과가 나이데스!

또한 밀양에 새 공장이 생겨버려서 자칫 공장을 확장한다면 그냥 파리만 왜앵왜앵 날리는 거예요
그렇다면 왜 브라질일까요? 브라질은 매운 거 먹냐고요? 아니요 잘 안먹어요 저 **이도 매운거
먹으면 아야하는데 나악한 서양사람들이 버티긴? 너무 어렵죠 하지만 젊은 사람들, 특히 틱톡등의
숏폼에서 매운맛이란 소재가 널리 광고됨에 따라 점차 시장이 확대될 것으로 보이고 있어요! 이는
남미 지역과 북미 지역에서도 나타나고 있는 상황이고요!

(6페이지)

그럼 왜 배송 지연을 중심으로 우리 **이가 봐왔을까요? 그건 우선 이커머스의 평균 10%가 넘게
성장세가 아주 빠르다는 것이고, 저희 조에서 개인 경험을 상의해 보았어요. 어떤 게 있었을 까요?
요즘 시대는 다양성의 시대!! 그래서 모두의 개인이야기에서 출발하고자했어요! 그랬더니
상준님이 배달 시 파손되어 교환받은 경험을 필두로 배송과 관련된 불만족 경험이 많았어요
그래서 이를 중심으로 접근했습니다

그래서 이를 좀 더 자세하게 보구 시퍼써서! 상관관계를 돌리고 지연 배송 여부에따라서 사람들의
리뷰 만족도를 수치화했더니 어떻게 된지 아세요? 보이세요? 바로 지연 배송을 했더니 매우 높은
연관성을 보이며 음의 상관계수가 나왔어요!!!! 사람들이 늦게 오면 싫다는 걸 확인한 거예요!

(7페이지)

데이터 분석 방안을 정리해 드릴 게요! 저희가 어떻게 데이터를 전처리했는 지를 중심으로 설명해
드릴 거예요!

(8페이지)

배송지연 데이터 전처리에 사용한 코드예요! 저희는 이렇게 배송 지연율을 계산하는 데 판다스 코드와 이를 위해서 유효한 배송만 필터링 하는 등의 전처리 과정을 거쳤답니다!

(9페이지)

이제 이상치를 제거할 시간! 아직은 사람이 만져줘야 한다구요~ 딱보고 딱 눈치를 챌 수 있는 사람의 눈치가 필요한단 말이죠! 그리고 2080법칙이라고 아시나요? 연준님이 잘아실거 같은데 대부분의 사고는 소수의 사고뭉치들이 다친다는 말이죠! 그래서 사고뭉치들이 있나 딱 뽑아 봤어요! 그랬더니 이렇게 사고뭉치들을 알아낼 수 있었고 사고뭉치 10명이 지연 배송의 15.4프로나 된다고요!

(10페이지)

배송 지연 데이터 이상치는 뽁뽁한 곳에서 밀도가 낮은 곳으로 추적했어요! 그래서 빈도가 1이 나오는 59일을 기준으로 결.정.했답니다!!!!

(11페이지)

데이터 분석 결론이에요! 여기서는 데이터 분석과정을 거쳐 어떤 결론에 다다랐는 지 설명해 드릴 거예요!

(12페이지)

이제까지 배송 지연이 고객들에게 얼마나 문제가 되는 지 알게 되었습니다. 그렇다면 저희가 이를 바탕으로 할 수 있는 것들은 중 공급 유통망의 본질에 접근하여 믿을 수 있는 셀러가 누구인지 찾아보고자 했습니다. 물론 그전에 믿을 만한 셀러를 찾는 것이 왜 중요한가를 정의하자면 보통 셀러가 배송에서 가장 큰 영향을 미치는 것이 무엇인지 알고자 했고, 그중 출고가 그 기준이 될 수 있다고 판단했습니다.

그래서 기대값과 실제값의 괴리를 파악하는 카이제곱 독립성 검증을 진행하여 이에 대한 저희의 판단이 맞는 지 검증하였습니다. 검증 결과 출고 지연에 따라 운송 지연이 더 많이 나타나게 된다는 것을 알게 되었고, 결국 배송지연이 적은 셀러는 출고지연도 적을 테니 이러한 신뢰할 수 있는 셀러가 많은 곳에 저희의 물류센터를 지어야한다고 생각했습니다. 이와 별개로 배송 지연율이 높은 곳에 배송 전문 물류센터를 짓는 것으로 셀러들의 지연을 줄일 수 있을 것이라 판단하였습니다.

(13페이지)

여기는 지역별 배송지연 분석입니다. 배송 지연률이 높은 곳의 경우 브라질 동북부 지역에 위치한 경우가 많았고, 빈부격차과 더불어 농업, 어업, 관광업 중심의 경제구조를 가지고 있었습니다. 반대로 배송 지연율이 낮은 곳의 경우 주문량 자체가 적은 곳이 많고, 북부에 대다수 있으며 낮은 인구밀도가 특징입니다. 저희 조에서 집중한 것은 바로 같은 인프라의 낙후에도 다른 차이를 보이는 것이었습니다. 이에 대해 지연율이 낮은 곳은 반대 경우와 달리 대다수의 인구가 주도에 집중해서 살고 있다는 것을 알게 되었습니다.

(14페이지)

결, 즉 해결방안 부분인 지역 선정 및 셀러 관리 방안 제안입니다.

(15페이지)

같은 낙후 인프라에도 다른 결과가 나온 것을 확인한 저희 조는 접근 방식도 한가지로 접근해서는 안된다고 생각했습니다.

첫번째 소비자 친화 물류센터를 조성한다. 이는 코스트코와 같이 물류 중심이긴 하지만 최종소비자들이 접근할 수 있도록 하는 매장 겸 물류센터를 의미합니다. 이러한 이유는 이 지역들의 경우 한 지역에 몰려살고 있는 높은 인구밀도로 셀러뿐만 아니라 최종 소비자인 고객들이 접근하기 좋다고 생각했기때문입니다.

두번째 불닭볶음면 전용 물류전문센터를 조성한다.

이 지역들의 경우 향만이 있는 도시들이며 교통인프라가 잘되어 있으나 배송 지연율이 높은 지역들로 물류 거점이 있다면 큰 효과를 볼것이라 추정되는 지역들입니다.

(16페이지)

세번째 셀러들을 어떻게 관리할 것인가 입니다.

아무리 좋은 지역에 간다고 하더라도 더 좋은 성과를 위해서는 셀러들을 관리하고 이들의 역량을 성장시키는 것이 필수적입니다. 이를 위해서 빠른 삼양 인증 배지와 같은 상벌 시스템 올리스트와의 협업으로 셀러들의 고충 해결, 마지막으로 고객들에게 서비스를 제공하는 형태로 지원을 주는 것입니다.

(17페이지)

저희 조에서 제안하고자 한 내용은 이것으로 마칩니다. 이제 다음단계 및 실행 계획을 통해 저희 조에서 과정 중 부족했던 점을 소개하고 이에 대한 개선방향을 말씀드리겠습니다.

(18페이지)

저희 조에서는 지연이 빈번하게 나타난 상위 10개 셀러들을 조사하면서 이들이 카테고리별로 특징이 나타난다는 것을 알게 되었습니다. 가구, 인테리어, 자동차 용품과 같이 화물의 크기가 크고 무겁고 식품과 달리 시간에 영향을 덜 받는 제품들의 주문에서 이러한 영향이 나타난다는 걸 알게되었습니다. 차후 조사에서는 데이터의 카테고리를 식료품등으로 국한하여 작업의 정밀성을 더하겠습니다.

(19페이지)

나머지는 저희 조가 작업한 자료들의 출처와 코드들을 정리한 부분입니다.

발표자료 구성(초안)

기 (동기)

- 브라질 **B2B** 전자상거래 시장 성장 (연 15.9%)
- 불닭볶음면 **B2B** 진출 필요성
- 고객 만족도 및 지연율에 대한 상관관계를 데이터 분석하여 지연율이 얼마나 사람들의 만족감을 낮추는 지 확인하여 지연율이 낮은 지역에 공장 혹은 물류 센터를 지을 수 있도록 한다.
- => 여기에서 왜 적절한 셀러들이 있는 곳에 장사를 해야하는 지 확인한다.

승 (데이터 분석)

- 올리스트 셀러별 배송 성능 분석
- 지역별 배송 지연율 분석
- 공장과 물류센터 중 물류센터로 접근하는 이유는?
- 1. 현재 취급하려는 불닭볶음면의 경우 최종소비자 중심의 제품임
- 2. 공장에서 바로 출고하게 되는 제품들의 경우 보통 규모가 크고 무게가 나가는 경우가 많음
- 3. 그렇기에 물류센터 혹은 허브로 접근하는 것이 좋을 거 같음
- => 여기에서 출고 지연이 늦어지면서 배송지연이 나타나게 된다는 것을 보이고, 이를 바탕으로 믿음직한 셀러들이 많은 곳을 선정하는 것이 중요함을

전 (접근법)

- 배송 성능 우수 지역 선별 기준 설정
- 신뢰할 만한 셀러 정의 (지연율 기준)
- 지역별 시장 특성 고려
- 여기서는 데이터 분석의 결과를 총체하는 느낌으로 가야겠다.

결 (지역 선정 및 셀러 관리 방안 제안)

- 분석된 결과를 바탕으로 적절한 지역들에 대해서 조사하고, 이에 대한 추천 이유를 제
- 다른 지역 고려 시 참고 요소:
 - 지역별 배송 지연율

- 셀러 신뢰도 지수
- 시장 성장 잠재력
- 물류 인프라 현황

다음 단계 및 실행 계획

- 카테고리 분석 결과 배송 지연 셀러들을 분석결과 차량 부품, 핸드폰등 카테고리들의 문제인 경우가 보임 그렇기에 차후 분석시에 식료품에 국한한 데이터로 작업한다거나 심한 지연 발생 셀러들을 꼽는 다던지 하는 방법으로 발전시키는 것을 희망함

0706 재롱 발표 대본

1페이지: 표지

안녕하세요, 5팀의 발표를 맡은 안재롱입니다. 저희는 브라질 대표 이커머스 멀티플랫폼인 Olist 데이터를 분석하면서, 한국의 대표 매운맛 불닭볶음면을 브라질 시장에 어떻게 진출시킬 수 있을지를 고민해봤습니다. 지금부터 '데이터 기반 전략 제안'이 어떻게 이루어졌는지 설명드리겠습니다.

2~3페이지: 팀 소개 및 목차

먼저 저희 팀은 총 5명으로 구성되어 있고,

현근님께서 팀장을 맡아 모두가 진행한 데이터 분석 내용을 취합 선별해 전체적인 틀을 잡고 일정 조율 및 진행을 해주셨습니다.

민지님과 상준님은 데이터 분석 위주로 진행해주시면서 데이터를 좀 더 심도있게 분석 파악해주셨고

한별님께서서는 데이터분석과 더불어 ppt작성 및 발표로 중반부 데이터 분석 부분을 설명해주시는 역 맡아주셨습니다.

마지막으로 저는 데이터 분석한 내용을 바탕으로 여러분께 저희의 동기부분 설명과 데이터 분석을 통해 해결방안과 전략을 설명하겠습니다.

4페이지~5페이지

오늘 발표는 다음과 같은 흐름으로 구성되어 있습니다. 먼저, 프로젝트를 수행하게 된 동기를 설명드린 뒤, 데이터 분석 방법과 결론 그리고 그에 따른 지역선정 및 셀러 전략, 마지막으로 실행계획과 제안으로 마무리하겠습니다.

📄 6~7페이지: 브라질 시장에서 불닭볶음면 선택의 이유 (기)

삼양의 불닭볶음면은 전 세계적으로 '매운맛' 열풍을 일으킨 대표적인 K-스낵입니다.

브라질 소비자들의 매운맛에 대한 관심은 최근 SNS와 한류문화 확산을 통해 빠르게 확산 중입니다. 불닭볶음면은 K-스낵 중에서도 글로벌 인지도가 높고, 도전적이고 유쾌한 이미지로 젊은 층에게 인기가 많습니다. 저희는 브라질에 기반한 'Olist'사의 내부 데이터'를 분석하는 과정에서 더 나아가 브라질 이커머스 시장의 높은 성장성과 빠르게 변화하는 소비 트렌드를 발견했고, 그 시장에 삼양사의 불닭볶음면이 진입한다면 충분히 경쟁력 있는 상품이 될 수 있다고 판단했습니다.

📄 8페이지: 왜 배송 지연을 중심으로 조사했는가?

브라질 B2B 전자상거래 시장의 성장률은 코로나 팬데믹을 기점으로 시장에 소비자들의 대규모 유입으로 전례없는 상승폭을 보이며 현시점에서는 연 15.9%의 달하는 성장률이 앞으로 더 높은 수치를 보일 것이라고 전망되고 있습니다.

여기서 이제 저희는 왜 배송 지연을 중심으로 조사했는가?에 대해 설명드리고자 합니다. 브라질 이커머스의 핵심 문제 중 하나는 '배송'이었습니다.

저희 팀원들의 실제 경험에서도, 제품이 파손되거나 배송이 지연되어 리뷰가 악화된 사례를 접했고, 이 문제를 'Olist'사 데이터'로 확인해보니 고객의 별점과 배송 지연 간에 상관관계가 -0.39로 나타났습니다.

📄 9페이지

즉, 배송 지연이 발생할수록 곧 고객 만족도를 하락시키며, 이는 브랜드 이미지에 직결된다는 점을 파악할 수 있었습니다.

처음 시장에 진입하는 삼양에게 이 문제는 절대로 간과할 수 없는 중요한 요소입니다. 그래서 앞으로의 데이터 분석 내용을 바탕으로 이 문제를 해결하는 방안까지 제안하고자 합니다.

뒤 이어서는 한별님에게 발표 부탁드립니다.

20페이지: 삼양과 Olist의 협업 전략 (결)

이어서 분석결과를 바탕으로, 삼양과 Olist가 함께 풀어나갈 수 있는 협업 전략을 말씀드리겠습니다.

21페이지: 추천 물류센터 위치 (결)

같은 낙후 인프라에도 다른 결과가 나온 것을 확인한 저희 팀은 접근 방식도 한가지로 접근해서는 안된다고 생각했습니다. 그래서 먼저 2가지 방안을 제시합니다.

첫번째 소비자 친화 물류센터를 조성합니다. 이는 코스트코와 같이 물류 중심이긴 하지만 최종소비자들이 접근할 수 있도록 하는 매장 겸 물류센터를 의미합니다. 이러한 이유는 이 지역들의 경우 한 지역에 높은 인구밀도로 셀러뿐만 아니라 최종 소비자인 고객들이 접근하기 좋다고 생각했기때문입니다.

두번째 불닭볶음면 전용 물류전문센터를 조성합니다.

이 지역들의 경우 항만이 있는 도시들이며 교통인프라가 잘되어 있으나 배송 지연율이 높은 지역들로 물류 거점이 있다면 큰 효과를 볼것이라 추정되는 지역들입니다.

추가로, 저희가 제안한 물류센터 위치는 단순히 배송지연율이 높은 지역이라는 이유뿐 아니라, 그 지역이 실제로 '주 전체 인구의 절반 이상이 집중된 중심 도시'이며, 항만, 고속도로 인프라가 갖춰져 있어 물류거점으로 적합하다는 점까지 고려한 것입니다. 예를 들어 '포르탈레자'는 세아라 주의 수도로 항구를 보유하고 있고, '마세이오' 역시 해상 물류가 가능한 지역입니다. 이러한 지역은 단순 인프라뿐 아니라 Olist 플랫폼 내에서도 주문량과 셀러 활동이 활발한 지역으로 파악되어 저희가 제안한 불닭 전용 물류망 시범 운영지로 적합하다고 판단했습니다.

22페이지: 삼양과 Olist의 협업 전략 (결)

세번째로, 차후 관리방안에 대해서 설명드리겠습니다.

먼저, 신뢰도 기반 프리미엄 셀러 우대 정책입니다. 출고 지연이 적고 배송 만족도가 높은 셀러에게는 '프리미엄 셀러 배지'를 부여하고, 제품 공급에 우선권을 주는 방식으로 운영합니다.

또, 추후 배송을 진행하며 브라질의 날씨, 교통 등 외부 요인을 반영한 배송 지원 시스템을 구축하여, 실질적인 지연 문제를 완화시키는 것입니다.

22페이지: 고객 대상 서비스 방안 (결)

삼양 입장에서는 고객과의 접점을 개선하는 것도 중요합니다.배송지연과 리뷰와의 상관관계를 시각화 했을때, 3일 이후 지연건에 대해서 가장 큰 기울기를 확인할 수 있었습니다. 그러므로 예상 배송일보다 3일 이상 지연될 경우 포인트나 쿠폰을 제공하고, 실시간 문자 알림 서비스를 제공함으로써 고객의 불만을 완화시키는 것도 좋은 방법이라 생각합니다.

이러한 정성적 대응은 단순한 상품 공급을 넘어, 고객과의 관계를 지속시키는 신뢰 기반을 다지는 것으로 이어질 수 있습니다.

22페이지: 셀러 관리 및 시스템 정비 (결)

또한 삼양은 장기적으로 Olist와 함께 셀러 관리 체계를 정립할 수 있습니다.

우선, 출고 신뢰도가 낮은 셀러에 대해서는 교육 및 가이드를 제공하고, 이와 같은 삼양 주도하의 개선방안 제안은 플랫폼 파트너십을 강화할 수 있으며, 좋은 이미지를 형성해 향후 B2B 협업, 우선 노출 등에 유리한 입지를 확보 할 수 있습니다.또, 브라질의 특수한 노동환경—예를 들면 월 2회 임금 지급 시스템에 따라 발생하는 물류 인력 이슈 등—을 반영해 출고 시점을 유연하게 조정해 나가는 방식도 고려해볼 수 있겠습니다.

23페이지: 다음단계 및 실행 (결)

마지막으로 이번 분석을 통해 저희가 파악한 보완점 개선 부분과 향후 진행시키면 좋을 계획에 대해서 설명 드리겠습니다.

24페이지: 보완점 제시 (결)

이번 분석에서 확인한 또 하나의 시사점은,배송 지연이 셀러들의 대부분이 식품이 아닌 핸드폰이나 차량 부품 등 취급한다는 것과도 관련이 깊다는 점입니다.

향후에는 식품 데이터에 국한해서 지연 이슈를 더 정밀하게 분석하거나, 심각한 지연을 일으키는 셀러를 별도로 추적하여 관리하는 방식으로 개선을 이어나갈 수 있을 것입니다.

추가적으로, 현재 셀러 신뢰도는 출고/배송 지연률로만 산출했지만, 셀러의 활동 기간, 누적 판매량 등도 함께 고려하면 더 안정적인 셀러 분류가 가능할 것입니다.

시간대별 배송 지연 패턴이나 시즌별 차이도 추후 분석 범위에 포함시킬 계획입니다. 예를 들어 월말, 명절 등 특정 기간에는 지연률이 급격히 높아질 수 있으니, 이에 대한 사전 예측 시스템이 유용할 것입니다.

 25페이지: 마무리 인사

지금까지 저희는 브라질 Olist 데이터를 기반으로, 배송이라는 이슈를 해결하며 삼양이 어떻게 성공적으로 시장에 진입할 수 있을지 전략을 제안드렸습니다.

이 전략이 전해질지 모르겠지만 (?) 브라질 내에서 불닭볶음면이 사랑받는 브랜드로 자리 잡는 데 도움이 되길 바랍니다. 감사합니다.

한별 대본 정리

Part 2

이제 데이터 분석 방안에 대해서 설명 드릴 건데요.

올리스트와 브라질에 불닭볶음면을 안정적으로 팔기 위해 물류센터를 짓는 것이 목표라고 말씀 드렸죠.

그럼 왜 하필 물류센터를 짓는 것을 선택했는지, 어느 곳에 물류센터를 지을 것인지를 설득하기 위해 데이터 전처리 과정 즉 EDA 결과를 보여드리면서 설명드리겠습니다.

보시는 코드는 배송지연 데이터 전처리에 사용한 코드예요!

저희는 우선 배송지연이 중요한 지표라고 생각해서 배송지연을 중심으로 데이터를 살펴보았습니다.

특히 배송지연을 자주 일으키는 셀러가 따로 있지 않을까? 하는 가정을 가지고 분석해보았습니다.

그랬더니 따라란 이런 결과가 나왔는데요.

혹시 파레토의 법칙, **2080법칙**이라고 아시나요? 전체 결과의 **80%**가 **20%**의 사고뭉치들에게서 발생한다는 현상을 설명합니다. 그래서 사고뭉치들이 있나 딱 뽑아 봤어요! 그랬더니

이렇게 사고뭉치 탑 **10**을 알아낼 수 있었고 무려 평균 **10.56**일의 지연을 일으켰고, 이들이 지연배송 건수의 합이 지연 배송의 **15.4%**나 된다는 것도 알아냈습니다. 결론적으로 사고뭉치 **10**명이 배송지연을 유의미하게 일으키고 있다고 볼 수 있습니다.

그 다음으로 그럼 배송지연이 고객 만족도 즉 리뷰 점수에 영향을 주는지 살펴보았습니다. 너무 당연한 얘기 같지만, 정확한 근거를 마련하기 위해서 **p-value**를 통해서 지연여부와 별점 간 상관관계를 분석해보았는데요.

귀무가설로는

대립가설로는

상관분석 결과 **p-value**가 **0**으로 유의수준인 **0.05**보다 작아 귀무가설을 기각합니다.

그래서 저희의 데이터 분석 결론을 말씀드리겠습니다.

저희가 배송지연을 중점적으로 살펴보았잖아요. 그럼 이때 배송지연이라는 것이 과연 어디서 시작되는가를 살펴보았습니다. 출고에서 지연이 생길 수 있고, 운송 자체에서 지연이 생길 수도 있으니까요. 하지만 저희는 출고지연과 운송지연에 상관관계가 있을 것이라고 가정했습니다. 그래서 각각의 단계의 평균 소요 일수를 비교해 출고지연과 운송지연의 상관관계를 분석해보았습니다.

출고 지연은 전체 주문의 ****9.3%****에서 발생하며, 평균 지연일은 **2.6**일

운송 지연은 **7.9%** 발생했지만, 평균 지연일은 **8.7**일로 체감은 훨씬 심각

출고 지연은 자주 발생

운송 지연은 상대적으로 적게 발생하지만, 피해는 더 큼

그래서 기대값과 실제값의 괴리를 파악하는 카이제곱 독립성 검증을 진행하여 이에 대한 저희의 판단이 맞는 지 검증하였습니다. 검증 결과 출고 지연에 따라 운송 지연이 더 많이 나타나게 된다는 것을 알게 되었고, 결국 운송지연이 적은 셀러는 출고지연도 적을 테니

출고 지연을 시키지 않는 셀러를 구하는 것이 중요하다고 판단했습니다. 결론적으로 이러한 신뢰할 수 있는 셀러가 많은 곳에 저희의 물류센터를 지어야겠죠. 이와 별개로 배송 지연율이 높은 곳에 배송 전문 물류센터를 짓는 것으로 셀러들의 지연을 줄일 수 있을 것이라 판단하였습니다.

그래서 지역별 배송지연률을 분석해 보았는데요. 배송 지연률이 높은 곳은 브라질 동북부 지역에 위치한 경우가 많았고, 이 지역은 빈부격차가 심하고 낙후된 인프라를 가지고 있었습니다. 더불어 농업, 어업, 관광업 중심의 경제구조를 가지고 있었습니다. 반대로 배송 지연율이 낮은 곳의 경우 주문량 자체가 적은 곳이 많고, 북부에 대다수 위치해있으며 인구 자체가 작아 인구밀도가 낮지만 그 소수의 인구가 일부 지역에 몰려 살고 있는 것이 특징입니다.

저희 조에서 집중한 부분은 바로 같은 인프라의 낙후에도 다른 차이를 보이는 것이었습니다. 또한 배송지연율이 낮은 곳은 반대 경우와 달리 대부분의 인구가 주도에 집중해서 살고 있다는 것을 알게 되었습니다.

이 결과를 가지고 재룡님을 다시 불러보겠습니다.

250708 회고

1. 아이스브레이킹

상준: 공부하다가 갔음. 집 가니 11시. 씻자마자 잠

재룡: 뭐 한 것도 없는데 새벽 1시 었음

민지: 괜찮아요. 술 보기만 해도 토나올 것 같아. 기분 좋아요. 속이 안 좋아요.

한별: 자소서가 안 써져서 기분이 안 좋아요.

2. 데이터 수집

좋았던 거

주제가 참신했다. 불닭!

처음에 막막했는데 훌륭하신 팀원들이 잘 길을 만들어주신 것 같다.

트러블이 생겼어도 잘 중재했다.

분업이 잘 됐다.

팀장이 잘 리드해서 혼선이 덜 됐다. 팀장 리더십 짱었다.

속전속결. 다른 팀은 마감시간 맞추는 게 촉박했는데 무리한 일정이 아니었음에도 주어진 시간을 잘 활용했다.

아쉬웠던 거

피드백을 각자에게 주었어야 했는데 모두가 있는 곳에서 주어서 무거운 시간이 길었다.(사실 아님... 다른 조는 우리가 현근님한테 빠따맞은 줄 앎)

쓸데없이 데이터 전처리에 시간이 오래 걸렸다. -> 주제와 피피티를 잡는 데 시간이 적게 걸렸다. 전처리가 느렸다가 보다는 주제 선정과 피피티 제작이 빨랐다.

챗지피티랑 작업할 때 지피티와 대화를 해서 이 코드가 어떤 코드인지 이해하면서 작성하는 것이 필요하다. 개인의 성장을 위해서라도.

처음에 어떤 걸 조사해야하는지 잘 몰랐다. 각자 하고 싶은 걸 조사한 건데, 만약 빈 부분이 있었다면 어땠을까. -> 현근이 생각하기에 서로 떨어져 있는 것 같아 보여도 합치는 걸 스스로 잘한다고 생각해서 걱정하지 않았다.

지역별 위치 찾을 때 좀 더 디테일하게 못 찾은 것 같아서 데이터 수집 단계에서 더 자세히 찾을 걸 그랬다.

좀 더 브라질 시장에 대한 더 자세한 조사를 했으면 좋았을 것 같다(연준 피드백 관련)

4L

Liked (좋았던 점)	Learned (배운 점)	Lacked (부족했던 점)	Longed For (바랐던 점)
- 주제가 참신했다 (불닭!)	- 프로젝트가 진행되다가 주제를 정해도 된다. 주제를 정하기 위해 다른 작업을 못할 정도로 너무 많은 시간을 쓰지 않는 것이 중요함.	- 피드백을 모두가 있는 자리에서 진행해 분위기가 무거워짐	- 지역 정보 수집의 디테일함
- 처음에 막막했지만 팀원들이 잘 길을 만들어주었다	- 챗GPT와 작업 시, 코드를 이해하려는 대화가 중요하다는 것을 배움	- 브라질 시장에 대한 조사 부족	- 브라질 시장에 대한 더 깊은 조사
- 트러블 발생 시 중재가 잘 되었다	- 피드백은 상황과 방식에 따라 신중히 전달해야 함을 느낌	- 데이터 전처리에 시간이 오래 걸린 듯 보였음	- 조사 항목 분담에 대한 사전 계획
- 분업이 잘 되었다		- 지역별 데이터를 디테일하게 찾지 못함	- 데이터 작업 중 카테고리별 분별이 필요
- 팀장의 리더십이 뛰어나 혼선이 없었다		- 조사 초기, 무엇을 조사할지 명확하지 않았음	
- 주어진 시간을 효율적으로 활용했다 (속전속결)			

인사이트 도출 (15분)

개선할 점에 대한 아이디어를 브레인스토밍합니다.

- 비슷한 내용의 포스트잇을 그룹으로 묶습니다.
- 그룹별로 토론하며 문제의 원인을 분석하고, 성공 요인을 파악합니다.
- "왜?"라는 질문을 반복하며 근본 원인(Root Cause)을 찾아냅니다.
- 가장 중요하고 영향력이 클 것이라 생각되는 아이디어를 2~3개 이내로 투표하여 선정합니다.

1. 작성된 의견들을 돌아가며 간략히 설명합니다. (비판이나 토론 X)

=> 요기는 날것의 이야기에서 키워드를 뽑아놓은 것으로 작업해라

2. 비슷한 내용끼리 그룹핑하고, 각 그룹에 제목을 붙입니다.

1. 에서 작업한 내용을 포괄하는 그룹명 붙이기

3. 논의가 필요한 가장 중요한 그룹에 투표합니다. (1인당 2~3표)

=> 인당 중복투표로 진행한다.

4. 투표를 많이 받은 주제부터 원인과 해결 방안에 대해 토론합니다.

=> 여기에 대해서 해결 방안 토론하기

구분	소제목	내용	투표
✖ Lacked	피드백 전달 방식	- 피드백을 모두가 있는 자리에서 진행해 분위기가 무거워짐	현근
✖ Lacked	조사 및 데이터 수집의 아쉬움	- 조사 초기, 무엇을 조사할지 명확하지 않았음 - 지역별 데이터를 디테일하게 찾지 못함 - 브라질 시장에 대한 조사 부족 - 지역 정보 수집의 디테일함 필요 - 브라질 시장에 대한 더 깊은 조사 필요	민지 상준 한별 재룡
✖ Lacked	작업 과정의 시간 배분	- 데이터 전처리에 시간이 오래 걸린 듯 보였음 (실제로는 다른 작업이 빨랐던 것일 수 있음)	
💡 Longed for	역할 분담의 사전 계획	- 조사 항목 분담에 대한 사전 계획이 있었다면 더 좋았을 것	현근, 한별,재룡 민지,상준

실행 계획(Action Item) 수립 (15분)

목표: 다음 기간에 '무엇을 시도해볼 것인지' 구체적이고 실행 가능한 계획(Action Item)을 세웁니다.

- 개선할 점에 대한 아이디어를 브레인스토밍합니다.
- 가장 중요하고 영향력이 클 것이라 생각되는 아이디어를 2~3개 이내로 투표하여 선정합니다.
- 선정된 아이디어에 대해 누가(Owner), 무엇을(Action), 언제까지(Due Date) 할 것인지 명확하게 정합니다.

1. 조사 및 데이터 수집의 아쉬움

1. 이유

- 브라질에 대한 관심이 별로 없었다? 아는 거 삼바랑 브릭스, 펠레, 월드컵밖에 모르는 디 갑자기 찾으라하니까 뭔가 어려움. EDA진행하면 해당 시장에 대한 관심이 필요했을 거 같음
- 데이터 분석에 치중하느라 '브라질' 자체에 관심을 덜 두었음.
- 데이터 분석을 하는 것이 중요하고 시간이 많이 걸리기 때문에 브라질 시장을 조사하는 것이 덜 중요하게 느껴졌음.
- 정보를 찾을 때 더 자세하게 정보를 찾아야 했는데 필요한것만 보고 지나친거 같아 나중에 질문을 받을때, 부족한 것을 느낌
- 조사 초기 각자 관심있는 분야를 조사하는 방식이 좀 문제가 있었던 것 같음. 겹치거나 재릉님 말대로 조사에서 빈 곳이 발생할 수 있었기 때문. 이것을 막기 위해서 사전에 역할분담을 더 제대로 하는 게 필요할듯.

2. 이를 개선할 방법

2. 역할 분담의 사전 계획

1. 이유

- 데이터 분석 첫 프로젝트(경험의 부족)이며 방대한 데이터 양에 놀라 정확한 방향 설정에 어려움이 있었음
- 팀장과 각 조원의 수집내용 교류는 잘 이루어졌지만, 각 조원들끼리는 조사내용 교류가 조금은 부족했던 것 같다.
- 팀장이 작업 방법을 노하우 방면으로는 아나 이를 문서화하지 못해 진행과정에서 설명을 해주었어야했다.
- 다 같이 조사 내용을 검토하고 조사 내용을 세분화한 뒤에 각자 어떤 부분을 맡을지 사전에 이야기했으면 좋았을 것 같다. 그렇게 하지 못해서 역할분담이 명확하지 않은 상태에서, 그러니까 불확실한 상태에서 계속 조사를 진행했던 것 같다.
- 조사 때 말고 그 후 피피티 제작이나 발표 역할을 정하는 건 잘 되었다. 속전속결로!

2. 이를 개선할 방법

1. 조사 및 데이터 수집의 아쉬움

구분	내용
이유 (도출된 내용)	<ul style="list-style-type: none"> - 브라질에 대한 기본 정보나 관심이 부족했다 (삼바, 펠레 외에 아는 게 별로 없었음) - EDA에 집중하느라 시장에 대한 이해가 소홀 - 필요한 정보만 수집해 전반적인 맥락이 약했음 - 조사 방식이 개별적이라 겹치거나 빈 구간이 생김
개선 방법 (팀 도출)	<ul style="list-style-type: none"> - 역할 분담을 사전에 더 정밀하게 계획 => 이를 위한 작업 테이블 작성 - 관심과 이해를 바탕으로 데이터 분석 진행

개선 방법 (ChatGPT 제안)	<ul style="list-style-type: none"> - 프로젝트 시작 전 시장/주제 브리핑 세션을 진행하여 관심 유도 및 배경 지식 공유 - 리서치 템플릿 제공: 조사 항목을 세분화해 누락 방지 - 조사 내용 간단 정리본을 슬랙/노션 등에 공유하여 서로 파악 가능하게 하기 - 데이터 분석 전, 시장 구조/현황을 요약한 5분 발표를 팀원이 돌아가며 진행하게끔 운영
-----------------------	---

2. 역할 분담의 사전 계획 부족

구분	내용
이유 (도출된 내용)	<ul style="list-style-type: none"> - 첫 데이터 분석 프로젝트라 방향 설정 어려움 - 팀장 ↔ 조원 간 교류는 잘 됐으나, 조원 ↔ 조원 간 공유 부족 - 팀장의 노하우는 있었지만 문서화가 부족했고, 설명을 그때그때 전달해야 했음 - 조사 전에 세분화된 분담 없이 각자 조사하다 보니 불확실한 상태로 진행됨
개선 방법 (팀 도출)	<ul style="list-style-type: none"> - 조사 시작 전에 조사 항목을 세분화하고, 명확하게 역할을 나누는 계획 수립 - 발표나 PPT 제작 역할은 나중에 잘 나뉘었기에, 조사 초기 단계도 그렇게 했어야 함 - 자신들의 과정을 독스등에 기록하는 것을 좀 더 습관화하여 작업하도록 장려 - 회의마다 이를 녹음하거나 공유 문서에 공유 - 쉬는 시간과 회의시간을 좀더 명확히
개선 방법 (ChatGPT 제안)	<ul style="list-style-type: none"> - 초반 회의에서 "조사 분야별 책임자" 지정 (예: 시장 규모, 소비자 특성, 경쟁사 등) - 조사 진행 중간에 점검 미팅을 1~2회 넣어 자료 공유 및 누락 체크 - 팀장의 작업 방식이나 진행 순서를 체크리스트로 문서화하여 모두가 기준점 공유 - 조사 마감 전, 각자 조사한 내용을 1~2문단 요약으로 슬랙/노션에 공유하는 방식 도입

1. 조사 및 데이터 수집의 아쉬움	
- 역할 분담을 사전에 더 정밀하게 계획 => 이를 위한 작업 테이블 작성	재 룡 상 준 한 별 민 지
- 관심과 이해를 바탕으로 데이터 분석 진행	현 근
- 프로젝트 시작 전 시장/주제 브리핑 세션을 진행하여 관심 유도 및 배경 지식 공유	민 지 상 준
- 리서치 템플릿 제공: 조사 항목을 세분화해 누락 방지	한 별 재 룡
- 조사 내용 간단 정리본을 슬랙/노션 등에 공유하여 서로 파악 가능하게 하기	현 근
- 데이터 분석 전, 시장 구조/현황을 요약한 5분 발표를 팀원이 돌아가며 진행하게끔 운영	
2. 역할 분담의 사전 계획 부족	
- 조사 시작 전에 조사 항목을 세분화하고, 명확하게 역할을 나누는 계획 수립	재 룡
- 발표나 PPT 제작 역할은 나중에 잘 나뉘었기에, 조사 초기 단계도 그렇게 했어야 함	
- 자신들의 과정을 독스등에 기록하는 것을 좀 더 습관화하여 작업하도록 장려	

- 회의마다 이를 녹음하거나 공유 문서에 공유	
- 쉬는 시간과 회의시간을 좀더 명확히	한별
- 초반 회의에서 "조사 분야별 책임자" 지정 (예: 시장 규모, 소비자 특성, 경쟁사 등)	현근 한별 재룡 민지 상준
- 조사 진행 중간에 점검 미팅을 1~2회 넣어 자료 공유 및 누락 체크	상준
- 팀장의 작업 방식이나 진행 순서를 체크리스트로 문서화하여 모두가 기준점 공유	
- 조사 마감 전, 각자 조사한 내용을 1~2문단 요약으로 슬랙/노션에 공유하는 방식 도입	현근 민지

선정된 개선 방안

1. 조사 및 데이터 수집의 아쉬움

- 역할 분담을 사전에 더 정밀하게 계획 => 이를 위한 작업 테이블 작성

2. 역할 분담의 사전 계획 부족

- 초반 회의에서 "조사 분야별 책임자" 지정 (예: 시장 규모, 소비자 특성, 경쟁사 등)

마무리 (5분)

결정된 액션 아이템을 다시 한번 확인하고 회고를 마칩니다.

- 결정된 실행 계획을 다시 한번 모두에게 공유하고 확인합니다.
- 회고가 어땠는지에 대한 간단한 피드백을 받습니다. (예: "오늘 회고에서 좋았던 점은?")
 - 한별: 생각보다 개선할 점이 많진 않았다.
 - 민지: 좋았던 점이 개선할 점 보다 많은 듯 정리한 표 보면 전반적인 진행이 잘 된듯 회고에서도 다양한 의견 좋았다.
 - 상준: 첫프로젝트이다 보니 함께하는 팀원들의 역량에 감사하고 많은 걸 배우고 스스로 부족한 부분들을 알 수 있었다.
 - 재룡: 다음 프로젝트 시에 어떤 걸 주안점으로 두고 진행할 지가 보였다.
 - 현근: 강한 피드백에 잘 받아줘서 고맙고, 다들 잘따라줘서 고맙습니다.
- 서로에게 감사를 표하며 마무리합니다.

포트폴리오용 프로젝트 정리

- 프로젝트 개요
 - 프로젝트 이름 : 불닭볶음면 브라질 진출을 위한 물류센터
 - 기간·팀 규모 : 4일, 5인
 - 목표·문제 정의 : 배송지연에 따른 불닭
 - 담당 역할
 - 한별: ppt제작 및 데이터 분석, 발표
 - 민지: 데이터 가공 및 분석
 - 재룡: 데이터 분석 및 메인 발표
 - 상준: 데이터 분석 및 자료 출처 및 근거 정리
 - 현근: 코드 수정 및 리뷰, 프로젝트 진행
- 동기 & 배경
 - 주제 선정 이유
 - 데이터 분석 및 개인들의 경험에서 배송 지연이 미치는 불만 사항을 알게됨
 - 기존(시장·사용자) 문제점 분석
 - 대다수의 배송지연을 차지하는 일부 셀러가 존재함
 - 목표 설정 과정
 - 서로의 대화에서 접근 방법을 찾고 이후 차별화 및 깊은 인상을 주기 위해 키워드 “불닭볶음면” 선정
- 요구사항 분석
 - 핵심 기능 목록
 - 주별 배송 지연율 및 소수의 셀러가 대다수의 배송 지연을 차지함을 확인
 - 사용자 시나리오·페르소나
 - 브라질에 처음 진출해서 시장 점유율 확대를 노리는 삼양식품
 - 성공 지표(KPI) 정의
 - 딱히 진행안했음
- 기술 스택 & 아키텍처
 - 사용 언어·프레임워크
 - pandas, matplotlib, seaborn 등
 - 인프라·배포 환경
 - 배포는 진행안하고 ppt로 발표 진행
 - 시스템 구성도
 - ppt로 발표 진행
- UI/UX 디자인
 - 와이어프레임·프로토타입
 - 사람들의 기억에 잘 남도록 호치를 다꾸처럼 넣음
- 협업 & 프로세스
 - 역할 분담·커뮤니케이션 도구
 - 구글 독스로 본인들이 한 작업물을 공유하였음
 - 애자일/칸반 등 개발 프로세스
 - 개인적으로 조원들이 조사하고 이를 조장이 결함함
- 결과 및 성과
 - 출시 결과(사용자·매출·트래픽 지표)

- 삼양이 좋은 셀러를 선정할 수 있을 것
 - 주요 학습·개선점
 - 해결방안을 이용할 대상을 선정하는 것이 작업 시작 전 가장 중요하다.
 - 사용자 피드백·리뷰 하이라이트
 - 제안 해결방안을 소비자 친화 매장(코스트코 느낌)과 불닭볶음면 전문(쿠팡 물류센터)으로 나눈 것이 생각보다 이해하기 어려워 하는 이들이 많음
- 회고 & 다음 단계
 - 개인적으로 얻은 인사이트
 - 주제를 약간의 변형으로 신선한 느낌을 줄 수 있다.
 - 아쉬웠던 점·보완 계획
 - 데이터를 카테고리별로 정리하여 접근하는 것이 중요하다는 걸 알게되었음
 - 문제가 되는 셀러들의 카테고리가 국한되어 있어 차후 데이터는 카테고리별로 분류후 작업하고자 함
 - 확장 아이디어·로드맵
 - 삼양이 국내에서 축적한 데이터와 올리스트의 데이터를 결합하여 전인류 공통 데이터 축적

