

6 장

종합 실습 문제 답안 샘플

※ 다음 답안 샘플은 2026년 6월 25일부터 7월 7일 기간 동안 챗지피티 5.5 Pro로 테스트한 것입니다. 사용하는 모델과 환경에 따라 답은 달라질 수 있음을 참고하세요.

요약 결론

- 1. 분석 종속변수는 환불을 반영한 순매출(net_revenue)로 두었다. 전체 총매출은 185.47억 원, 순매출은 167.78억 원로, 환불/반품 반영 후 매출은 총매출의 90.5% 수준이다.
- 2. 분석 단위는 상품-주로 고정했다. 일 단위는 리뷰 수가 희소해 텍스트 신호가 불안정하고, 상품-주는 리뷰 신호와 퍼널 지표를 결합하기에 충분한 표본 안정성을 제공한다.
- 3. 개선 대상 제품의 평균 별점은 H1 4.235에서 H2 4.483로 상승했고, 주간 평균 순매출도 883.3만 원에서 1012.8만 원으로 증가했다.
- 4. 상품 및 주차 고정효과와 가격·할인·광고·품질을 통제한 회귀에서 평점 0.3점 상승은 주간 순매출 약 15.8% 증가와 연관된다. 평균 상품-주 기준으로는 약 100.4만 원 규모다. 단, 이는 인과 효과가 아니라 통제 회귀 기반의 연관성 추정이다.
- 5. 퍼널 관점에서는 조회율보다 장바구니율과 구매전환율이 평점 변화에 더 민감했다. 저평점 불만 토픽 중 환불률은 배송 지연/미도착, 제품 효과/피부반응, 포장/파손 순으로 높았다.

데이터 개요

데이터셋	행 수	기간/구성	핵심 요약
제품 데이터	50	5개 카테고리 / 5개 브랜드	개선 대상 제품 10개
리뷰 데이터	13923	2024-01-02~2025-01-05	평균 별점 4.01
매출 데이터	18250	2024-01-01~2024-12-30	총 순매출 167.78억 원

■ 단계 1. 질문을 측정 가능한 형태로 정리

[문제 1에 대한 답안 샘플]

순매출(net_revenue)을 사용한다. CHAPTER 06의 질문은 리뷰와 평점이 실제 구매 경험과 매출에 어떻게 연결되는지 보는 것이므로, 반품과 환불이 반영된 순매출이 고객 경험 악화의 비용을 더 직접적으로 반영한다. 총매출은 환불 전 판매 규모를 보여주지만, 부정 리뷰와 환불이 동시에 증가하는 경우 성과를 과대평가할 수 있다.

표 1-1. 매출 지표 선택 근거

지표	값	해석
총매출	185.47억 원	환불 전 매출 규모
환불/반품금액	17.68억 원	총매출의 9.5%
순매출	167.78억 원	총매출 대비 90.5%

[문제 2에 대한 답안 샘플]

분석 단위는 상품-주(product-week)로 고정했다. 매출 데이터는 일 단위지만, 리뷰 텍스트는 매일 모든 상품에 충분히 쌓이지 않는다. 상품-주 단위로 묶으면 리뷰 수, 평균 평점, 감성 점수, 토픽 비중을 안정적으로 집계하면서도 가격·할인·광고·품질 같은 일별 교란 요인을 주간 평균 또는 주간 합계로 통제할 수 있다.

회귀 분석에서는 더 엄격하게, 해당 주의 매출을 설명하는 리뷰 신호를 그 주 시작일 이전 30일 동안의 리뷰만으로 계산했다. 이는 같은 주 매출 이후에 작성된 리뷰가 설명변수에 섞이는 정보 혼입을 줄이기 위한 처리다.

[문제 3에 대한 답안 샘플]

관찰 구간은 products 데이터의 is_improving=True인 10개 개선 대상 제품으로 정하고, 2024년 1~6월(H1)과 2024년 7~12월(H2)을 비교했다. 이 구간에서 평균 별점은 4.235에서 4.483으로 상승하여, 실습 시나리오의 “4.1에서 4.4로 상승”에 가까운 개선 패턴을 보인다.

단순 전후 비교만으로는 평점 상승이 매출을 유발했다고 단정할 수 없다. 같은 기간에 광고비, 할인율, 품질률, 시즌성, 제품 자체 개선, 카테고리 트렌드가 동시에 바뀔 수 있기 때문이다. 실제로 H2에는 개선 제품의 할인율과 광고비도 소폭 상승했고, 이러한 변수는 매출과 함께 움직일 수 있다.

표 1-2. 개선 대상 제품과 기타 제품의 전후 비교

제품군	기간	평균 별점	리뷰 수	조회율	장바구니율	구매전환율	일평균 순매출	평균 할인율	일평균 광고비	품질일 비율
기타 제품	H1	3.809	4,782	6.88%	14.17%	3.16%	78.6만 원	1.90%	11.5만 원	2.25%
기타 제품	H2	3.920	5,271	6.89%	14.36%	3.32%	82.4만 원	1.86%	11.2만 원	1.79%
개선 대상 제품	H1	4.235	1,710	7.09%	16.28%	4.30%	125.9만 원	1.77%	11.3만 원	2.25%
개선 대상 제품	H2	4.483	2,061	7.19%	17.38%	4.94%	149.4만 원	2.12%	11.8만 원	2.51%

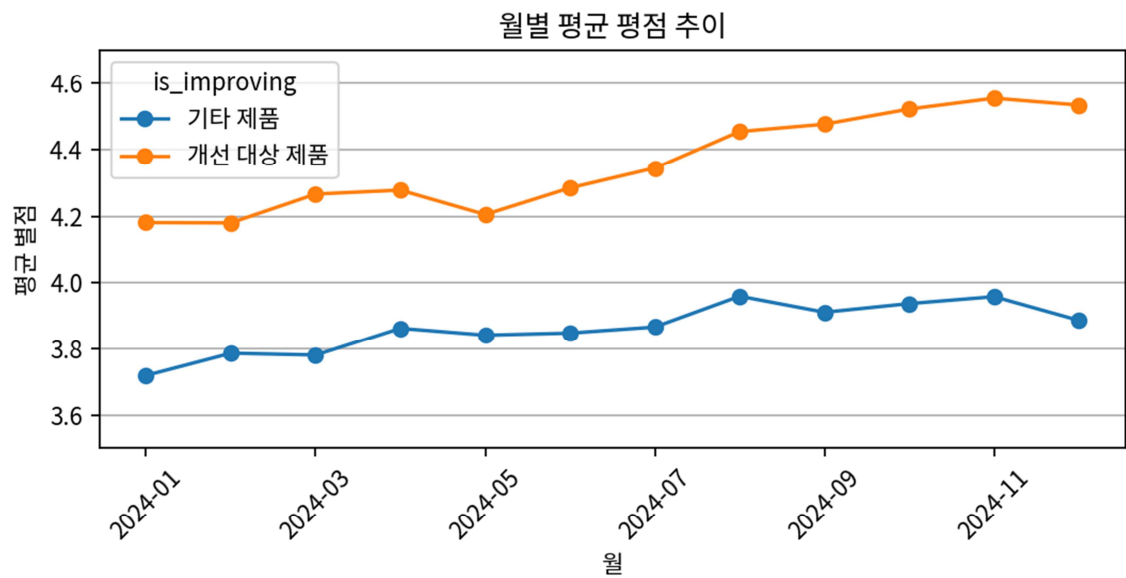


그림 1-1. 개선 대상 제품은 하반기로 갈수록 평균 평점이 상승하며, 기타 제품은 상대적으로 완만한 상승을 보인다.

■ 단계 2. 퍼널 관점에서 리뷰가 영향을 주는 지점 찾기

[문제 1에 대한 답안 샘플]

상품-주 단위로 조회율(product_views/impressions), 장바구니율(add_to_cart/product_views), 구매전환율(purchases/product_views)을 계산했다. 전체 상품-주 패널은 2,650건이며, 최근 30일 리뷰 신호가 있는 회귀 가능 표본은 2,595건이다.

표 2-1. 상품-주 단위 퍼널 지표 요약

퍼널 지표	평균	표준편차	최소	최대
조회율	6.94%	0.17%	6.37%	7.32%
장바구니율	14.75%	1.81%	8.91%	21.26%
구매전환율	3.50%	0.97%	0.00%	7.48%

[문제 2에 대한 답안 샘플]

각 상품-주의 주간 퍼널 지표에 직전 30일 평균 평점, 리뷰 수, 평균 감성확률, 부정 감성 비율, 토픽 비중을 결합했다. 평점 구간별 단순 평균을 보면 평점이 높을수록 장바구니율과 구매전환율이 특히 뚜렷하게 높아진다.

표 2-2. 최근 30일 평점 구간별 퍼널·매출 지표

최근 30일 평점 구간	상품-주 수	조회율	장바구니율	구매전환율	평균 주간 순매출	평균 할인율	품질일 비율
<4.1	1,771	6.85%	13.89%	3.04%	502.3만 원	1.86%	2.02%
4.1~4.4	533	7.09%	16.28%	4.29%	882.9만 원	1.88%	2.44%
>=4.4	291	7.18%	17.28%	4.93%	981.1만 원	2.12%	1.87%

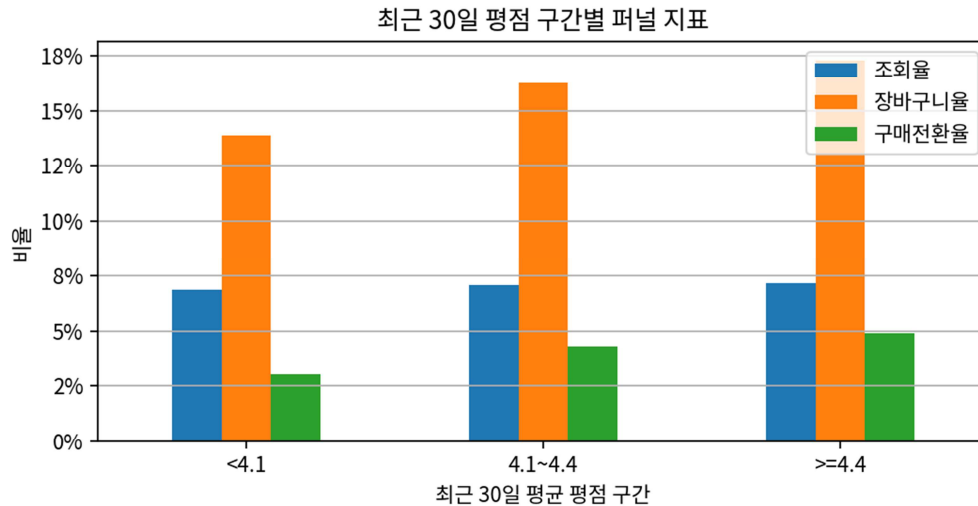


그림 2-1. 평점 구간별 퍼널 지표. 장바구니율과 구매전환율의 기울기가 조회율보다 크다.

[문제 3에 대한 답안 샘플]

상품 고정효과, 주차 고정효과, 할인율, 광고비, 가격, 품질일 비율을 함께 통제한 회귀에서 평점 0.3점 상승은 조회율보다 장바구니율 및 구매전환율의 개선과 더 크게 연결되었다. 이는 리뷰가 첫 노출보다는 “상세페이지 진입 이후 구매 의사결정” 구간에서 더 강하게 작동한다는 해석과 맞다.

표 2-3. 평점 상승과 퍼널 지표의 통제 회귀 결과

종속변수	평점 계수	p값	평점 +0.3p 효과	평균 대비 상대 변화	조정 R ²
조회율	0.00225	<0.001	0.067pp	1.0%	0.899
장바구니율	0.02246	<0.001	0.674pp	4.6%	0.902
구매전환율	0.01317	<0.001	0.395pp	11.3%	0.898

교란 요인 점검 결과, 평점 구간별 평균 할인율과 품질률에도 차이가 존재한다. 따라서 “평점이 높아져서 모든 퍼널이 좋아졌다”가 아니라, 광고·가격·품질·시즌성을 통제한 뒤에도 장바구니율과 구매전환율에서 긍정적 연관성이 남는다고 정리하는 것이 적절하다.

■ 단계 3. 리뷰 텍스트를 모델 입력으로 만들기

[문제 1에 대한 답안 샘플]

표 3-1. 전처리 규칙

항목	처리 방침	근거
특수문자/문장부호	!, ..., … 등은 공백으로 정규화하되 감정 강도 표식으로 과도하게 반복된 경우 별도 토큰화 가능	본 데이터는 “만족! 재구매”, “배송 늦음.. 건조해요”처럼 구두점이 구분자 역할을 함
이모지/반복문자	ㅋㅋ, ㅋㅋ가 있을 경우 삭제하지 않고 <웃음>, <울음>으로 치환; 반복 문자는 2회로 축약	감성 분석에서는 감정 강도 신호가 될 수 있음
숫자/옵션 코드	가격·용량·옵션 번호는 <NUM> 또는 <OPTION>으로 정규화하되, 가격 불만 문맥은 보존	“가격 대비”는 가격 토픽 핵심 신호
불용어	상품, 구매 같은 지나치게 일반적인 단어는 제거 후보. 단, 안/못/별로/없음은 보존	부정어 제거 시 감성 방향이 뒤집힘
토큰화 단위	형태소 또는 업무 사전 기반 어절/구문 토큰. 본 데이터는 구문이 짧아 “배송 늦음”, “박스 찌그러짐”을 사전 구문으로 보존	토픽 라벨을 업무 언어로 연결하기 쉬움

[문제 2에 대한 답안 샘플]

텍스트 표현 방식은 TF-IDF 기반 문서-단어 행렬에 1~2gram을 결합하는 방식을 선택했다. 이 데이터의 리뷰는 짧고 반복 표현이 많아 문장 임베딩보다 TF-IDF/n-gram 방식이 해석 가능성이 높다. 특히 “배송 늦음”, “박스 찌그러짐”, “효과 없음”, “CS 불친절”처럼 2gram이 업무상 불만 유형과 직접 연결된다.

감성 분석에는 TF-IDF+로지스틱 회귀가 적합하다. 어떤 단어/구문이 긍정 또는 부정 확률을 높이는지 계수로 설명할 수 있기 때문

이다. 토픽 모델링에는 형태소·구문 기반 카운트 행렬과 업무 사전 라벨링을 병행하는 것이 더 안정적이다.

[문제 3에 대한 답안 샘플]

최소 빈도는 min_df=2 또는 5로 두어 매우 드문 오탈자성 토큰을 제거하고, 최대 빈도는 max_df=0.95로 두어 모든 문서에 가까이 등장하는 일반 단어를 제거한다. 어휘 사전 크기는 max_features=1,000으로 제한했다. 실제 데이터는 고유 리뷰 텍스트가 124개 수준으로 짧고 정형적이므로, 이 제한은 과적합과 과도한 희소성을 막는 안전장치로 충분하다.

■ 단계 4. 감성 분석 설계 및 평점과의 불일치 점검

[문제 1에 대한 답안 샘플]

지도 학습 기반 감성 분석을 선택했다. 별점 4~5점을 긍정, 별점 1~2점을 부정으로 약한 라벨링하고 별점 3점은 중립으로 제외하여 TF-IDF+로지스틱 회귀 기준선 모델을 학습했다. 장점은 설명 가능성과 재현성이고, 단점은 별점이 라벨로 쓰이므로 텍스트 감성의 독립 검증이 약하다는 점이다.

표 4-1. 감성 분석 기준선 성능

항목	결과	해석
학습 라벨	별점 4~5=긍정, 별점 1~2=부정, 별점 3=중립으로 제외	약한 지도학습 기준선
검증 표본	2,248건	긍정 비율 95.7%
ROC-AUC	1.000	텍스트와 별점이 매우 일관적임
Average Precision	1.000	불균형 라벨에서도 기준선 성능 양호
Accuracy / F1	1.000 / 1.000	합성 실습 데이터 특성상 거의 완전 분리

검증 혼동행렬은 부정 96건/긍정 2152건을 모두 맞춘 형태였다. 이는 데이터가 실습용으로 정형화되어 있고, 부정 텍스트와 낮은 별점이 매우 일관적으로 구성되어 있음을 의미한다.

[문제 2에 대한 답안 샘플]

별점별 감성확률 분포를 비교하면 별점 1~2점은 평균 감성확률이 약 0.01, 별점 4~5점은 약 0.99로 분리된다. 별점 3점은 중립 텍스트가 대부분이어서 약 0.60 수준으로 나타난다.

표 4-2. 별점별 감성확률 분포

별점	리뷰 수	평균 감성확률	표준편차	최소	최대
1	23	0.009	0.005	0.003	0.024
2	458	0.010	0.005	0.003	0.027
3	2,687	0.595	0.000	0.595	0.595
4	7,013	0.990	0.003	0.983	0.995
5	3,742	0.990	0.003	0.983	0.995

표 4-3. 평점-텍스트 감성 불일치 점검

불일치 점검 기준	건수	해석
별점 4~5인데 감성확률 < 0.40	0건	없음. 데이터상 긍정 별점과 긍정 표현이 일치
별점 1~2인데 감성확률 > 0.60	0건	없음. 데이터상 부정 별점과 불만 표현이 일치
실무 점검 필요 유형	정성 샘플링 권장	반어법, 선물 목적 리뷰, 배송은 긍정이나 품질은 부정인 혼합 리뷰

실무 데이터였다면 별점 5점이지만 “배송은 빠르는데 향이 너무 독함”처럼 혼합 감성이 있는 리뷰, 별점 1점이지만 “제품은 좋은데 배송이 안 옴”처럼 원인별 감성이 다른 리뷰를 반드시 샘플링하여 따로 분류해야 한다.

[문제 3에 대한 답안 샘플]

각 상품-주에는 그 주 시작일 이전 30일간의 평균 감성확률을 집계했다. 리뷰 수가 적은 상품-주의 평균 감성이 불안정해지는 문제를 줄이기 위해 최소 리뷰 수 기준을 두거나, 전사 평균과 상품-주 평균을 리뷰 수로 가중하는 베이지안 스무딩을 적용할 수 있다. 예를 들어 $\text{smoothed sentiment} = (n \times \text{상품-주 평균} + k \times \text{전체 평균}) / (n + k)$ 로 계산하고 k는 5 또는 10으로 둔다.

■ 단계 5. 토픽 모델링 기반 불만 내용 구조화 및 토픽 수 정당화

[문제 1에 대한 답안 샘플]

낮은 별점 리뷰(1~3점)를 대상으로 CountVectorizer와 LDA를 탐색적으로 적용하고, 동시에 업무 사전 기반 라벨링으로 토픽을 안정화했다. 리뷰가 매우 짧아 LDA만 사용하면 “보통/나쁘지 않음” 같은 중립 표현이 토픽을 흐리는 문제가 있었다. 따라서 퍼플렉서티, 간이 일관성, 시드 안정성뿐 아니라 토픽의 업무 운영 가능성을 함께 고려하여 K=6을 채택했다.

표 5-1. 토픽 수 K 탐색 요약

K	퍼플렉서티(낮을수록 좋음)	간이 UMass(높을수록 좋음)	시드 안정성	해석
4	10.84	-4.375	0.414	배송·포장·효과 불만이 섞임
5	10.98	-4.120	0.467	가격/CS가 부분적으로 분리
6	10.81	-4.202	0.462	배송, 포장, 피부반응, 가격, CS, 중립을 업무 언어로 분리 가능
7	10.94	-3.975	0.465	중립 리뷰가 과도하게 쪼개짐
8	12.78	-3.954	0.488	퍼플렉서티 악화, 토픽 과분할

K=6은 배송 지연/미도착, 포장/파손, 제품 효과/피부반응, 가격 대비 가치, CS 응대, 중립/일반평으로 나뉘어 실제 개선 과제와 연결하기 쉽다. K=7~8은 일부 지표가 개선되어 보여도 중립 토픽이 불필요하게 쪼개져 실행 언어로서의 장점이 낮았다.

[문제 2에 대한 답안 샘플]

저평점 리뷰에서 중립/일반평을 제외하면 포장/파손, 배송 지연/미도착, 제품 효과/피부반응이 가장 많이 나타난다. 환불률은 배송 지연/미도착 토픽이 58.5%로 가장 높아 매출 손실과 직접 연결될 가능성이 크다.

표 5-2. 저평점 리뷰 토픽별 지표

토픽 라벨	리뷰 수	저평점 내 비중	평균 별점	환불률	평균 도움돼요
중립/일반평	2,687	84.8%	3.00	0.0%	0.0
포장/파손	138	4.4%	1.93	37.7%	9.3
배송 지연/미도착	123	3.9%	1.97	58.5%	9.5

제품 효과/피부반응	112	3.5%	1.96	42.0%	9.7
CS 응대	61	1.9%	1.97	32.8%	10.2
가격 대비 가치	47	1.5%	1.94	0.0%	9.7

저평점 리뷰(1~3점) 기준 토픽별 리뷰 수

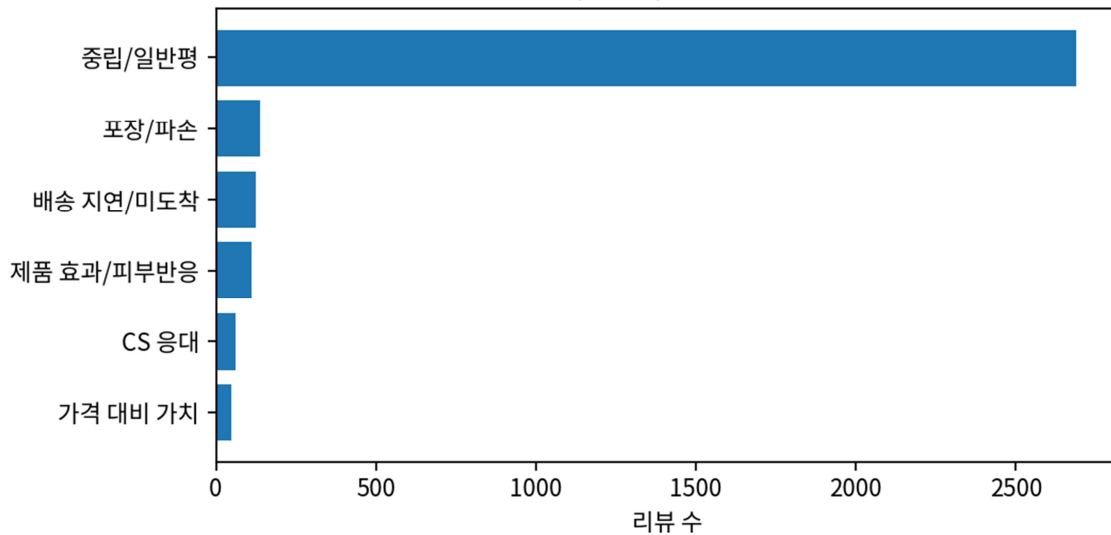


그림 5-1. 저평점 리뷰 토픽 분포. 중립/일반평을 제외하면 포장/파손, 배송 지연/미도착, 제품 효과/피부반응이 핵심 불만이다.

표 5-3. 개선 대상 제품의 전후 토픽 비중 변화

기간	사용감/향 긍정	배송/가성비/재구매 긍정	중립/일반평	배송 지연/미도착	포장/파손	제품 효과/피부반응	가격 대비 가치	CS 응대
H1	52.9%	34.2%	11.6%	0.5%	0.1%	0.4%	0.1%	0.3%
H2	56.6%	37.8%	5.1%	0.1%	0.0%	0.2%	0.0%	0.2%

개선 대상 제품에서는 H2에 사용감/향 긍정 및 배송/가성비/재구매 긍정 토픽 비중이 높아졌고, 중립/일반평 및 배송·포장 관련 불만 비중은 낮아졌다. 이는 평점 상승이 단순 숫자 변화가 아니라 긍정 경험 언어의 증가와 일부 불만 언어의 감소를 동반했음을 보여 준다.

■ 단계 6. 토픽과 감성을 매출로 연결

[문제 1에 대한 답안 샘플]

종속변수는 $\log(1+\text{주간 순매출})$ 로 설정했다. 주요 설명변수는 직전 30일 평균 평점, $\log(\text{리뷰 수})$, 평균 감성확률, 주요 토픽 비중이다. 통제 변수로 평균 가격, 평균 할인율, $\log(\text{광고비})$, 품질일 비율, 주차 고정효과, 상품 고정효과를 포함했다.

표 6-1. 통제 변수와 이유

통제 변수	통제 이유
가격	가격이 높으면 구매전환과 매출이 동시에 바뀔 수 있음
할인율	할인은 전환율을 높이는 동시에 평점·리뷰 작성 유인을 바꿀 수 있음
광고비	광고는 노출과 매출을 직접 증가시키므로 리뷰 효과와 구분해야 함
품질 여부	평점이 좋아도 품질이면 구매가 발생하지 않음
주차 더미	시즌성, 월별 행사, 외부 수요 변화를 통제
상품 고정효과	브랜드, 카테고리, 기본 가격대, 출시 시점 등 상품 고유 차이를 통제

표 6-2. 주간 순매출 회귀 결과

변수	계수	p값	해석
최근 30일 평균 평점	0.4897	<0.001	평점이 높을수록 주간 순매출 증가 방향
$\log(\text{최근 30일 리뷰 수})$	0.1289	<0.001	리뷰량 자체도 신뢰/노출 신호로 작용
평균 감성확률	-0.6279	0.002	평점과 상관이 매우 높아 단독 해석 주의
평균 할인율	0.1648	0.934	본 패널 고정효과 모델에서는 별도 효과 뚜렷하지 않음
$\log(\text{광고비})$	0.1716	<0.001	광고비가 높을수록 순매출 증가 방향
품질일 비율	-2.5829	0.051	품질은 매출 감소 방향
조정 R^2	0.875		상품 및 주차 고정효과 포함

[문제 2에 대한 답안 샘플]

회귀계수 기준으로 평점 0.3점 상승은 $\exp(0.4897 \times 0.3) - 1 = 15.8\%$ 의 주간 순매출 증가와 연관된다. 분석 표본의 평균 상품-주 순매출을 적용하면 약 100.4만 원 수준이다.

표 6-3. 평점 상승 시나리오 해석

항목	계산/관찰	결과	주의
평점 0.3점 상승	$\exp(0.4897 \times 0.3) - 1$	15.8%	평균 상품-주 기준 100.4만 원 증가와 같은 규모
개선 대상 제품 H1→H2 관찰값	평균 별점 4.235→4.483	+0.248점	동시에 평균 주간 순매출 881.3만 →1,012.8만 원으로 상승
해석 범위	회귀+고정효과 기반 연관성	인과효과 아님	광고, 가격, 품질, 시즌성을 통제했지만 미관측 품질개선·캠페인 동시 변동 가능

표 6-4. 개선 대상 제품의 상품-주 관찰값 변화

기간	최근30일 평점	최근30일 리뷰 수	조회율	장바구니율	구매전환율	평균 주간 순매출	평균 할인율	주간 광고비	품질일 비율
H1	4.222	25.8	7.09%	16.30%	4.32%	883.3만 원	1.79%	79.3만 원	2.23%
H2	4.451	33.4	7.19%	17.38%	4.95%	1012.8만 원	2.18%	80.1만 원	2.43%

[문제 3에 대한 답안 샘플]

이 결과는 통제 회귀 기반의 연관성이지만 인과 효과가 아니다. 첫째, 제품 품질 개선, 패키지 리뉴얼, 프로모션, 광고 소재 교체처럼 관측되지 않은 변화가 평점과 매출을 동시에 움직였을 수 있다. 둘째, 매출이 증가하면 구매자가 늘고 리뷰 수도 증가하여 역인과 가능성이 있다. 셋째, 리뷰 작성 고객은 전체 구매자를 대표하지 않을 수 있다. 따라서 이 결과는 “평점 상승이 매출 증가와 함께 움직인다”는 근거로 사용하되, 실제 인과 효과는 A/B 테스트나 단계적 롤아웃, 차분의 차분 설계로 추가 검증해야 한다.

최종 실행 제안

표 7-1. 리뷰 분석 기반 실행 과제

우선 과제	데이터상 신호	실행안	KPI
배송 지연/미도착	환불률 최고	물류 SLA 점검, 미도착 자동 보상, 배송 안내 메시지 개선	배송 불만 리뷰 비중, 배송 관련 환불률, 구매 전환율
포장/파손	저평점 불만 수 상위	포장재 강도 개선, 파손 사진 자동 수집, 출고 전 검사	포장/파손 토픽 비중, 환불률, 도움돼요 높은 부정 리뷰 수
제품 효과/피부반응	환불률 높음	피부 타입별 상세페이지 문구 개선, 사용법 콘텐츠, 민감성 주의 표기	피부반응 토픽 감성, 상세페이지 체류, 반품률
가격 대비 가치	환불은 적지만 불만 강도 높음	용량/가격 커뮤니케이션, 번들 구성, 가치 설명 강화	가격 토픽 비중, 장바구니율, 쿠폰 사용 후 재구매율
긍정 토픽 확대	사용감/향·재구매 언어 증가	후기 하이라이트, 베스트 리뷰 노출, 추천 카피 강화	긍정 토픽 비중, 리뷰 도움돼요, 구매 전환율

분석 운영 시에는 스냅샷 원칙을 유지해야 한다. 매출을 설명하는 리뷰 지표는 반드시 해당 매출 기간 이전까지의 리뷰로만 계산하고, 토픽과 감성 모델은 신조어·이모지·제품명 변경에 맞춰 정기적으로 재검증한다. 최종 보고는 감성 점수 자체가 아니라 “무엇을 고치면 어떤 토픽이 줄고 어떤 퍼널 지표가 개선되는가”로 연결되어야 한다.