

4 장

종합 실습 문제 답안 샘플

※ 다음 답안 샘플은 2026년 6월 25일부터 7월 7일 기간 동안 클로드 Fable 5로 테스트한 것입니다. 사용하는 모델과 환경에 따라 답은 달라질 수 있음을 참고하세요.

답안 활용 시 유의사항: 생성형 AI 결과의 변동성에 관하여

본 해설의 수치는 첨부 데이터로 재현 가능한 값이다. 그러나 4장 과제는 결과의 변동 폭이 크다. 회귀 계수처럼 닫힌 해가 있는 부분(단계 1~4, 8~10)은 어떤 AI를 쓰든 거의 동일하게 나오지만, 그리드 탐색(adstock의 λ), 비선형 최적화(Hill 함수), 제약하 예산 최적화(단계 5~7)는 AI가 선택한 탐색 범위, 초기값, 알고리즘, 통제변수 포함 여부에 따라 결과가 상당히 달라진다.

따라서 학생의 답안이 본 해설과 수치가 다르다고 해서 틀린 것이 아니다. 채점의 기준은 숫자의 일치가 아니라 (1) 각 분석의 원리(자기상관 진단, 잔존효과, 수확체감, 한계 ROAS 균등화, 무작위 배정의 논리)를 이해했는지, (2) 자신이 얻은 결과를 그 원리에 비추어 일관되게 해석했는지, (3) 결과의 한계를 스스로 지적했는지이다. 각 단계의 주황색 상자에 변동이 예상되는 지점을 구체적으로 명시하였다.

다만 모든 차이가 허용되는 것은 아니다. 예컨대 AI가 절편 없는 회귀모형을 선택하여 VIF를 89~107로 보고하는 경우가 있는데, 절편 없는 모형은 아주 특수한 경우에만 정당화되므로 이는 해석의 차이가 아니라 모형 설정의 오류이며 오답으로 처리한다(단계 3 참조).

■ 단계 1. 탐색적 데이터 분석

[문제 1의 답안 샘플]

프롬프트 예시

첨부한 smartbean_marketing_daily_2026.csv를 불러와 다음을 수행해줘. (1) 결측치 유무를 변수별로 확인, (2) 모든 수치형 변수의 평균·표준편차·최소·최대를 표로, (3) 4개 채널의 연간 총 광고비와 일평균 광고비, (4) 총 매출과 일평균 매출, (5) 월별 평균 매출 추이를 표와 선그래프로. 통화는 천 단위 구분 기호를 붙여 읽기 쉽게 표시해줘.

기대 결과

데이터는 365일(2026-01-01~12-31), 13개 변수로 구성되며 결측치는 없다(총 0개). 주요 변수의 기초 통계는 다음과 같다.

변수	평균	표준편차	범위
네이버 광고비	880,448원	73,103	70만~105만
인스타그램 광고비	649,278원	57,051	56만~78만
유튜브 광고비	477,145원	45,333	40만~58만
카카오톡 광고비	404,142원	38,442	34만~49만
일 매출	13,608,698원	823,891	1,000만~1,595만

〈표 1〉 광고비·매출 기초 통계

연간 총 광고비는 네이버 약 3.21억, 인스타그램 2.37억, 유튜브 1.74억, 카카오톡 1.48억 원이며 4채널 일평균 합계는 약 241만 원이다. 연간 총 매출은 약 49.7억 원, 일평균 1,361만 원이다. 월별 평균 매출은 겨울(2~3월, 약 1,417만~1,450만 원)과 연말(11~12월, 약 1,403만~1,409만 원)에 높고, 경쟁사 프로모션이 있던 6월(약 1,273만 원)에 가장 낮아 계절성과 외부 충격이 함께 관찰된다.

■ 단계 2. 상관관계 분석

[문제 1의 답안 샘플]

프롬프트 예시

4개 채널 광고비와 매출의 관계를 2×2 산점도(각 패널에 회귀선 포함)로 그리고, 5개 변수(4채널+매출)의 상관계수 행렬을 히트맵으로 시각화해줘. 채널 간 상관이 높아 다중공선성 우려가 있는지 평가해줘.

기대 결과

광고비-매출 단순 상관은 네이버 $r=0.283$, 인스타그램 0.202, 카카오톡 0.128, 유튜브 0.110 순이다. 모두 양(+)의 약~중간 상관으로, 네이버가 매출과 가장 강하게 연관된다. 채널 간 상관은 절댓값 기준 최대 0.091(네이버-유튜브)에 불과해 채널 간 다중공선성은 사실상 없다. 이 사실은 단계 3에서 VIF 해석의 기준점이 된다.

■ 단계 3. 다중 회귀분석 (원자료)

[문제 1의 답안 샘플]

프롬프트 예시

revenue를 종속변수로, 4개 채널 광고비를 독립변수로 하는 다중 회귀를 적합하고 (1) 각 계수와 p값, (2) VIF, (3) 잔차 플롯으로 선형성·등분산성, (4) Durbin-Watson으로 자기상관, (5) 조정된 R^2 를 보고해줘. 네이버 계수를 "광고비 1만원 증가 시 매출 변화"로 해석해줘.

기대 결과

채널	계수 β	p값	해석
네이버	3.18	<0.001	1원 추가 시 매출 3.18원 ↑ (1만 원당 약 3.2만 원)

인스타그램	2.91	<0.001	1만 원당 약 2.9만 원 ↑
유튜브	1.40	0.120	통계적으로 유의하지 않음
카카오톡	2.35	0.026	1만 원당 약 2.3만 원 ↑

〈표 2〉 원자료 다중 회귀 결과 ($R^2 = 0.143$, 조정 $R^2 = 0.133$)

모형의 설명력이 낮다는 점(조정 $R^2 = 0.133$)이 단계 4~6으로 넘어가는 핵심 동기다. 광고비만으로는 매출 변동의 14% 남짓밖에 설명하지 못한다.

진단 1 — 자기상관:

Durbin-Watson = 0.367로 강한 양의 자기상관이 있다. 이는 오차항이 독립이라는 가정의 위반으로, 광고 효과가 여러 날에 걸쳐 지속되는 잔존효과(adstock)와 매출의 계절성 때문이다. 단계 4~5에서 이를 모델링한다.

진단 2 — VIF:

단계 2에서 채널 간 상관이 최대 0.091로 낮았으므로, 개념적으로 다중공선성은 없다. 절편을 포함한 올바른 회귀모형에서 VIF를 계산하면 실제로 4개 채널 모두 $VIF \approx 1.0$ 이 나온다.

결과의 변동 가능성 (AI에 따라 다르게 나올 수 있는 부분)

AI에 따라 가장 크게 갈리는 수치가 VIF이다. 일부 AI는 절편 없는 회귀모형을 선택하거나(R 의 $\text{lm}(y \sim 0 + .)$, Python에서 상수항 없이 적합), 절편 없는 설계행렬에 statsmodels의 `variance_inflation_factor()`를 적용한다. 이 경우 VIF가 89~107로 매우 크게 나온다. 광고비가 모두 0에서 멀리 떨어진 큰 양수라서, 절편 없이 계산하면 비중심 R^2 가 부풀려지기 때문이다.

채점 기준: 절편 없는 모형에 근거한 VIF는 오답으로 처리한다. 절편 없는 회귀모형은 이론적으로 원점 통과가 강제되는 아주 특수한 경우에만 정당화된다. 본 과제는 광고비가 0이어도 기저 매출이 존재하는 상황이므로 절편이 반드시 필요하며, 절편을 제거하면 VIF뿐 아니라 계수 추정과 R^2 해석까지 모두 왜곡된다. 따라서 학생이 AI가 제시한 절편 없는 모형의 VIF(89~107)를 그대로 보고하고 "다중공선성이 심각하다"고 결론지었다면, 이는 해석의 차이가 아니라 모형 설정의 오류이므로 틀린 것으로 채점한다.

정답으로 인정하는 경우는 두 가지다. (1) 처음부터 절편 포함 모형에서 $VIF \approx 1.0$ 을 얻고, 단계 2의 낮은 채널 간 상관(최대 0.09)과 일치함을 확인한 답안. (2) 높은 VIF를 얻었더라도 그 원인이 절편 누락임을 스스로 진단하고, 절편 포함(또는 중심화) 후 재계산하여 $VIF \approx 1.0$ 으로 정정한 답안. 후자는 AI 출력의 오류를 비판적으로 검증한 사례이므로 오히려 우수 답안으로 평가할 수 있다.

■ 단계 4. 통제 변수 추가

[문제 1의 답안 샘플]

프롬프트 예시

단계 3 모형에 외생 변수를 통제변수로 추가해줘. 요일 더미(월요일 기준), 프로모션 여부, 경쟁사 프로모션 여부, 기온, 공휴일 여부를 넣고, 통제 전후 광고 계수가 어떻게 바뀌는지 비교해줘. 프로모션과 경쟁사 프로모션의 효과 크기를 원단위로 보고하고 조정 R^2 를 해석해줘.

기대 결과

통제변수를 넣으면 조정 R^2 가 0.133에서 0.828로 급등한다. 매출 변동의 대부분이 광고비가 아니라 계절성·요일·프로모션·경쟁 같은 외생 요인에서 비롯됨을 보여주는 결정적 장면이다. 통제 후 광고 계수는 네이버 3.86, 인스타그램 2.50, 카카오톡 2.01로 재추정되고, 원자료 모형에서 유의하지 않던 유튜브 계수도 유의해진다($\beta=1.63$, $p<0.001$). 누락변수 편익이 제거되면서 계수 추정이 안정된 것이다.

변수	효과	p값
프로모션	+1,299,919원/일	<0.001
경쟁사 프로모션	-1,094,483원/일	<0.001
기온	-48,131원/°C	<0.001
공휴일	-328,328원/일	<0.001

〈표 3〉 주요 통제변수의 효과

자체 프로모션은 하루 매출을 약 130만 원 올리고, 경쟁사 프로모션은 약 109만 원 깎는다. 기온이 오를수록 매출이 소폭 감소하는 것은 커피가 겨울 성수기 상품이라는 설정과 일치한다.

결과의 변동 가능성 (SI에 따라 다르게 나올 수 있는 부분)

요일 더미의 기준(월요일 vs 일요일), 공휴일 포함 여부, 월 더미 추가 여부에 따라 광고 계수와 조정 R^2 의 소수점 이하가 달라질 수 있다. 예컨대 월 더미까지 넣으면 R^2 는 더 오르지만 기온 효과가 계절성과 겹쳐 약해질 수 있다. 조정 R^2 가 0.8 부근으로 급등하고 유튜브가 유의해지는 정성적 패턴이 재현되었다면 올바른 답안이다.

■ 단계 5. 마케팅 믹스 모델링

[문제 1의 답안 샘플]

프롬프트 예시

각 채널 광고비에 geometric adstock 변환을 적용해줘. $adstock_t = x_t + \lambda \cdot adstock_{t-1}$. λ 를 0.1~0.9에서 0.1 간격으로 바꿔가며 각 채널의 최적 λ 를 R^2 기준으로 탐색하고, 각 채널의 반감기($= \ln 0.5 / \ln \lambda$)를 계산해줘. adstock 적용 전후 회귀 R^2 를 비교해줘.

기대 결과

데이터 생성 시 심어둔 실제 감쇠율(decay rate)은 네이버 0.3, 인스타그램 0.5, 유튜브 0.7, 카카오톡 0.4이며, 이에 대응하는 이론적 반감기는 각각 약 0.58일, 1.00일, 1.94일, 0.76일이다. 그리드 탐색으로 학생이 얻는 추정 λ 는 노이즈와 외생 요인 때문에 실제값과 정확히 일치하지 않는다. 광고비만 넣은 회귀에서 채널별로 탐색하면 대략 $\lambda = 0.5 \sim 0.7$ 부근의 추정치가 나오며, adstock 적용 시 R^2 는 원자료 0.143에서 약 0.21로 상승한다.

중요한 것은 정확한 λ 값이 아니라 (a) adstock을 적용하면 모형 적합도가 개선된다는 점, (b) 채널마다 잔존 강도가 달라 유튜브처럼 λ 가 큰 채널은 광고 효과가 오래 지속된다는 점($\lambda=0.7$ 이면 반감기 약 1.9일)을 발견하는 것이다.

우수한 답안은 λ 를 광고비만의 회귀 R^2 로 정하면 외생 요인 때문에 편향될 수 있으므로, 단계 4의 통제변수를 포함한 상태에서 λ 를 탐색하는 것이 더 타당하다는 점을 지적한다.

결과의 변동 가능성 (시에 따라 다르게 나올 수 있는 부분)

λ 추정치는 이 과제에서 변동이 가장 큰 수치 중 하나다. (1) 채널을 하나씩 바꾸며 탐색하는지 네 채널을 동시에 탐색하는지, (2) 통제변수를 포함하는지, (3) 탐색 간격(0.1 vs 0.05)과 기준(R^2 vs AIC)에 따라 채널별 최적 λ 가 0.4~0.8 사이에서 달라진다. 실제로 통제변수 포함 여부만 바뀌도 네이버의 최적 λ 가 0.5에서 0.6으로 이동한다.

따라서 학생의 λ 가 심어둔 실제값(0.3/0.5/0.7/0.4)이나 본 해설의 추정치와 달라도 감점 사유가 아니다. adstock 개념을 올바르게 구현했는지(재귀식), 반감기를 올바르게 환산했는지, 적합도 개선을 확인했는지를 평가한다.

■ 단계 6. Saturation 효과 모델링

[문제 1의 답안 샘플]

프롬프트 예시

각 채널의 광고비-매출 관계에 Hill 포화 함수 $f(A) = A^s / (K^s + A^s)$ 를 적합해줘. 채널별로 α (최대 기여), K (반포화점), S (형태 모수)를 추정하고 포화 곡선을 그려줘. 현재 평균 광고비가 K 보다 큰지 작은지 비교해, 각 채널이 아직 증액 여지가 있는지(비포화) 판단해줘.

기대 결과

채널별 평균 광고비는 네이버 88만, 인스타그램 65만, 유튜브 48만, 카카오톡 40만 원이다. 데이터는 각 채널이 반포화점 K 에 도달하기 전, 즉 수확체감이 본격화되기 전의 상승 구간에서 운영되도록 생성되었다. 따라서 모든 채널이 "아직 증액하면 매출이 늘어나는" 비포화 상태로 해석되는 것이 정상이다. 다만 한계 기여는 광고비가 늘수록 점차 둔화된다.

학생이 Hill 적합에서 K 와 S 추정이 불안정할 수 있는데(광고비의 관측 범위가 좁아 곡선의 포화 구간을 관측하지 못하므로), 이는 "관측 범위 밖으로 외삽할 때 신중해야 한다"는 실무 교훈과 연결된다.

결과의 변동 가능성 (AI에 따라 다르게 나올 수 있는 부분)

Hill 함수의 비선형 최소제곱 적합은 초기값과 옵티마이저에 민감하다. 실제로 이 데이터에 곡선을 적합하면 형태 모수 S 가 10~50 이상으로 발산하거나, 채널에 따라 K 가 관측 범위 밖의 값으로 추정되거나, 아예 수렴에 실패하는 경우도 있다. AI마다 (또는 같은 AI라도 실행마다) α , K , S 추정치가 크게 다르게 나오는 것이 오히려 정상이다.

채점 시 파라미터 수치 자체가 아니라 (1) 포화 함수의 개념(수확체감)을 이해했는지, (2) 적합이 불안정한 이유(포화 구간 미관측)를 설명했는지, (3) "현재 지출 수준에서는 비포화"라는 실질적 결론에 도달했는지를 본다. 로그 변환이나 지수 포화 함수($1 - e^{-\theta A}$) 같은 대안적 함수형을 쓴 답안도 동일한 결론에 도달했다면 정답으로 인정한다.

■ 단계 7. 예산 최적화

[문제 1의 답안 샘플]

프롬프트 예시

월 예산 8,000만 원을 4개 채널에 재배분해 매출을 극대화해줘. 제약: 총합 8,000만 원 고정, 각 채널 최소 10%(800만)~최대 40%(3,200만). MMM의 adstock·saturation 파라미터를 활용하고, 그리디 또는 라그랑지안으로 최적 배분을 구해줘. 현재 배분 대비 예상 매출 증가분과, 최적해에서 모든 채널의 한계 ROAS가 같아지는지 검증해줘.

기대 결과

핵심 원리는 최적 배분에서 모든 채널의 한계 ROAS(추가 1원당 매출 증가분)가 동일해진다는 것이다(등식 제약 하 극대화의 1차 조건, 라그랑주 승수법). 한계 기여가 높은 채널은 늘리고 낮은 채널은 줄이되, saturation 때문에 한 채널에 무한정 몰지 않는다.

단계 4의 통제 후 계수(네이버 3.86 > 인스타그램 2.50 > 카카오톡 2.01 > 유튜브 1.63)를 기준으로 한 방향성은 네이버 비중 확대, 유튜브 비중 축소다. saturation을 반영한 한 가지 권고 배분은 네이버 약 39%, 인스타그램 25%, 카카오톡 20%, 유튜브 16% 수준이다.

결과의 변동 가능성 (AI에 따라 다르게 나올 수 있는 부분)

권고 배분 %는 어떤 반응 함수를 쓰느냐에 따라 구조적으로 달라진다. 선형 계수만 사용하면 해는 항상 제약의 모서리에 위치한다: 계수가 큰 순서대로 상한을 채워 네이버 40%, 인스타그램 40%, 카카오톡 10%, 유튜브 10%가 나온다. saturation을 반영하면 수확체감 때문에 내부해로 이동하여 39/25/20/16과 같은 완만한 배분이 나온다. 두 결과 모두 각자의 모형 가정에서는 올바른 최적해다.

채점 기준: (1) 한계 ROAS 균등화 원리를 이해했는지, (2) 제약조건(총액 고정, 채널별 10~40%)을 올바르게 적용했는지, (3) 선형 외삽만으로 특정 채널에 몰아주는 답의 위험을 saturation 논리로 방어했는지. 정확한 % 배분의 일치 여부는 요구하지 않는다.

■ 단계 8. A/B 테스트 분석

[문제 1의 답안 샘플]

프롬프트 예시

smartbean_ab_test_2025_10.csv에는 처치 전(9월, period=pre)과 처치 후(10월, period=post) 데이터가 있어. 먼저 사전기간에 두 그룹의 일매출·광고비·신규고객이 통계적으로 다르지 않았는지 t-검정으로 확인해줘. 이후 사후기간에 대해 그룹별 비교표와 시각화를 만들어줘.

기대 결과

처치가 지역(서울/경기 vs 기타)과 겹치는 준실험 설계이므로, 사후 비교에 앞서 기저선 동등성을 반드시 확인한다. 사전기간(9월) 일매출은 처치군 1,204.3만 원, 통제군 1,202.0만 원으로 차이가 통계적으로 무의미하며 (Welch $t = 0.14$, $p = 0.888$), 광고비($p = 0.903$)와 신규고객($p = 0.827$)도 마찬가지다. 두 그룹은 처치 전 관측 가능한 결과변수에서 동등했다.

사후기간은 다음과 같다.그룹	일매출 평균	표준편차	일 광고비 평균	신규고객 평균
처치군(서울/경기)	14,173,546원	534,241	2,688,326원	165.1명
통제군(기타)	12,054,805원	657,344	2,682,088원	138.2명

〈표 4〉 처치군·통제군 기술통계 (2026년 10월, 각 31일)

두 그룹의 일 광고비는 약 268만 원으로 거의 동일하다(예산 총액은 같고 채널 간 배분만 달랐으므로). 처치군이 매출·신규고객 모두에서 일관되게 높고, 시계열·박스플롯에서도 분포가 뚜렷이 분리된다.

■ 단계 9. 가설 검정

[문제 1의 답안 샘플]

프롬프트 예시

처치군과 통제군의 일매출 차이가 통계적으로 유의한지 두 표본 t-검정으로 검정해줘. 귀무가설과 대립가설을 명시하고, $\alpha=0.05$ 에서 결론을 내리고, Cohen's d로 효과 크기를 해석해줘. 신규 고객 수에 대해서도 동일하게 검정해줘.

기대 결과

귀무가설 $H_0: \mu_{\text{처치}} = \mu_{\text{통제}}$, 대립가설 $H_1: \mu_{\text{처치}} \neq \mu_{\text{통제}}$.

일매출에 대한 Welch t-검정 결과 $t \approx 13.93$, $p \approx 4.3 \times 10^{-20}$ 로 유의수준 0.05에서 귀무가설을 기각한다. Cohen's $d = 3.54$ 로 매우 큰 효과다(관례상 0.8 이상이면 큰 효과). 신규 고객 수도 $t \approx 7.57$, $p \approx 2.7 \times 10^{-10}$, $d = 1.92$ 로 역시 강하게 유의하다. 두 지표 모두 처치(최적 예산 배분)가 통계적으로도 실질적으로도 의미 있는 개선을 가져왔다.

다만 처치가 무작위 배정이 아니라 지역 단위 배정이므로 이 차이만으로 인과 효과를 단정할 수 없다. 사전기간의 기저선 동등성(단계 8)이 결합될 때, 비로소 관찰된 개선을 예산 재배분의 효과로 해석할 근거가 마련된다.

결과의 변동 가능성 (SI에 따라 다르게 나올 수 있는 부분)

등분산 가정의 Student t-검정을 쓰면 t값과 자유도가 Welch 검정과 소폭 달라지지만(이 데이터에서는 결론 동일), R의 `t.test()`는 기본이 Welch, 일부 SI의 Python 코드는 기본이 Student(`equal_var=True`)라서 보고되는 t·p값이 다를 수 있다. Cohen's d도 합동표준편차식과 평균표준편차식 중 무엇을 쓰는지에 따라 소수점이 달라진다. 어느 방식이든 가설·검정·결론의 논리가 옳바르면 정답이다.

■단계 10. 비즈니스 의사결정

[문제 1의 답안 샘플]

프롬프트 예시

A/B 결과로 전사 확대 여부를 판단해줘. (1) 일·월·연 증분 매출(처치 평균 - 통제 평균), (2) 추가 광고비와 iROAS(=증분 매출/추가 광고비), (3) 마진 30% 가정 시 연 순이익 증가분과 투자 회수 기간을 계산하고 권고안을 제시해줘.

기대 결과

일 증분 매출은 약 211.9만 원, 월(30일 기준) 약 6,356만 원, 연간으로는 약 7.6억~7.7억 원이다. 두 그룹의 광고비가 거의 같았으므로(추가 광고비 일 약 6천 원에 불과) iROAS는 약 340배로 매우 크다. 즉 추가 비용 없이 예산을 재배분하는 것만으로 매출이 증가했다는 뜻이다. 마진 30%를 가정하면 연 순이익 증가분은 약 2.3억 원이다.

권고:

효과 크기가 크고($d > 0.8$), 통계적으로 강하게 유의하며($p < 0.001$), 추가 광고비가 사실상 0이므로 최적 배분의 전사 확대를 권고한다. 다만 학생이 (a) 실험이 10월 한 달·특정 지역 구분에 국한되어 외적 타당도에 한계가 있고, (b) 기저선 동등성을 확인했으나, 지역 단위 배정이므로 10월에 특정 지역에만 발생한 외부 충격(지역 이벤트, 날씨 등)과 처치 효과는 여전히 구분할 수 없고, 사전기간 1개월만으로 엄밀하게는 기저선 동등성을 담보할 수 없으며, (c) 장기적으로는 saturation 때문에 동일 효과가 지속되지 않을 수 있다는 점을 함께 지적하면 우수 답안이다.

결과의 변동 가능성 (시에 따라 다르게 나올 수 있는 부분)

연간 증분 매출은 계산 관행에 따라 두 값이 나온다: 일 증분 $\times 365$ = 약 7.73억 원, 월 증분(30일) $\times 12$ = 약 7.63억 원. 어느 쪽이든 계산 근거를 명시했다면 정답이다. 마진 30% 기준 연 순이익도 이에 따라 약 2.29억~2.32억 원 범위가 된다.

iROAS는 분모(추가 광고비 일 약 6,238원)가 0에 가까워 값이 매우 크고 불안정하다. 학생이 '수백 배'라는 정확한 배수 대신 "분모가 사실상 0이므로 iROAS 계산이 무의미하며, 비용 중립적 재배분으로 해석해야 한다"고 답했다면 오히려 더 정확한 해석으로 가산할 수 있다.