

1. 데이터 개요

TV, 라디오, 소셜미디어 광고비와 이에 따른 매출 정보를 포함한 총 4,546건의 광고 집행 기록으로 구성됨.

본 분석은 광고비와 매출 간 관계를 다각도로 검토하고, 예측 모델 구축 및 시나리오 분석을 통해 효율적인 예산 배분 전략을 도출하는 것을 목표로 함.

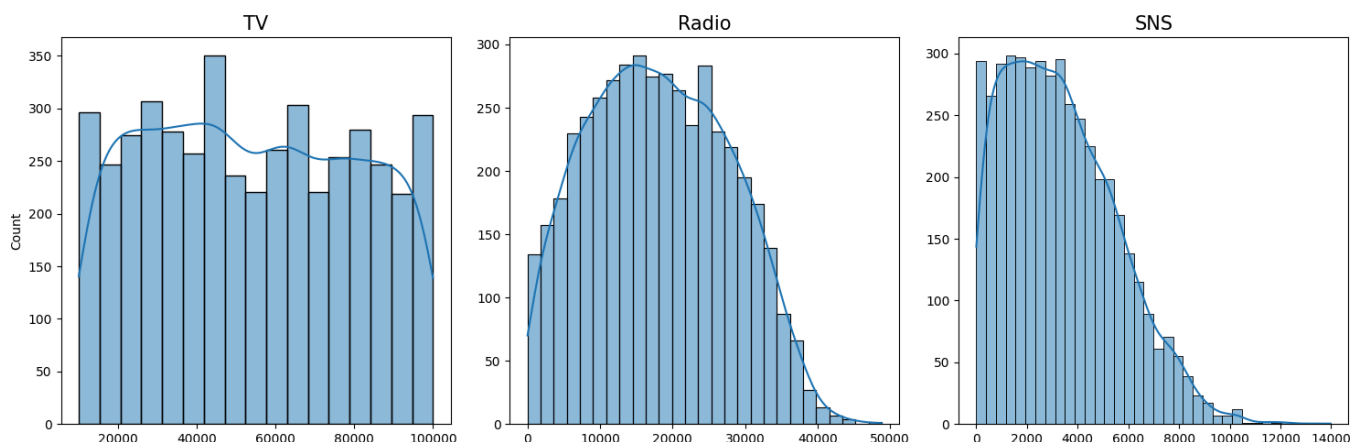
- 데이터 구성: TV 광고비, 라디오 광고비, SNS 광고비, 인플루언서 유형, 매출

2. 기초 통계 및 분포 확인

지표	평균	표준편차	최소값	최대값
TV 광고비	54,062.91	26,104.94	10,000.00	100,000.00
라디오 광고비	18,157.53	9,663.26	0.68	48,871.16
SNS 광고비	3,323.47	2,211.25	0.03	13,981.66
총 매출	192,413.33	93,019.87	31,199.41	364,079.75

- 광고비 평균: TV > 라디오 > SNS, 매출 평균: 약 192,413
- 변수 전반에 걸쳐 넓은 범위와 큰 표준편차를 보이며, 매출은 약 3만~36만 사이에 분포함

매출 분포 시각화

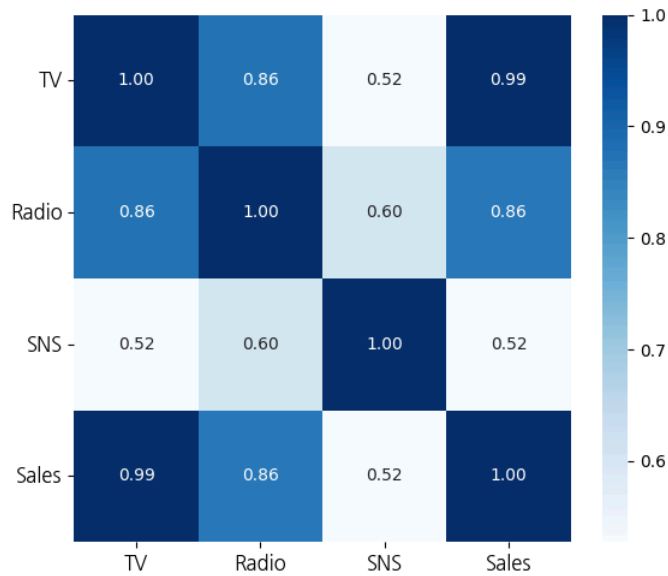


- 매출 분포는 전반적으로 완만하게 퍼진 형태로, 일부 고액 매출이 존재함

- 극단적인 이상값 없이 비교적 연속적 분포를 가짐

3. 광고비와 매출 간 관계 분석

3-1. 상관관계 분석



- TV 광고비와 매출: 0.99 → 가장 높은 영향력
- 라디오 광고비: 0.86
- SNS 광고비: 0.53 → 중간 수준 영향력
- TV와 라디오는 핵심 채널로, SNS는 보조적 효과로 해석 가능

3-2. 다중 회귀 분석

항목	결과값	해석
R ²	0.999	세 광고비 모두 포함해도 매출의 거의 모든 변동을 설명함
TV 계수	3.563	광고비가 1 증가할 때 매출이 평균 3.56만큼 증가함
라디오 계수	-0.004	매출과 통계적으로 유의한 관계가 없음
SNS 계수	0.005	매출에 유의미한 영향 없음
오차 진단	이상 없음	잔차의 정규성·독립성 조건 모두 충족

→ TV 광고비는 매출과 매우 강한 선형 관계를 가지며, 유의한 독립 변수로 작용

- 다중공선성(VIF)

변수	VIF	해석
TV	18.566	다중공선성 의심 수준(10 초과)
라디오	21.142	다중공선성 강하게 의심됨
SNS	4.961	허용 범위 내 (다중공선성 문제 없음)

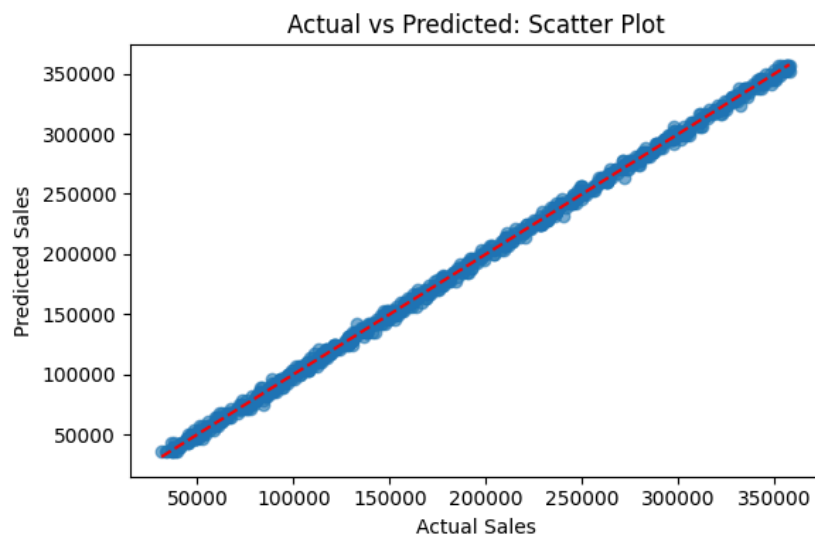
다중공선성(VIF): TV와 라디오 간 다중공선성 가능성이 있음

특히 라디오 광고비는 회귀 계수가 통계적으로 유의하지 않아 제거 고려 가능

4. 예측 모델 구축 및 평가

- 학습 데이터(80%)로 선형 회귀 모델 구축, 테스트 데이터(20%)로 예측 수행
- 모델은 TV, 라디오, 소셜미디어 광고비를 입력값으로 사용하고, 매출을 종속변수로 설정함
- 예측 결과는 다음과 같은 평가 지표를 기준으로 검증함
 - MAE: 2,342.20 RMSE: 2,895.44 MAPE: 1.72%

→ 모든 지표에서 오차가 낮고, 평균 오차율도 2% 미만으로 매우 높은 예측 정확도



→ 예측값이 실제값과 거의 일치하는 모습을 보여줌

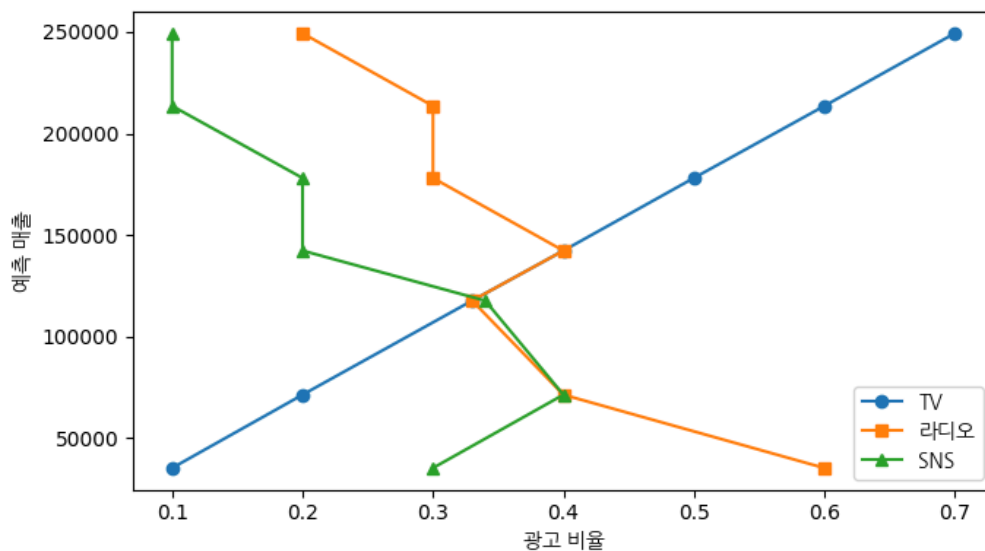
- 오차가 큰 샘플 파악

실제 매출	예측 매출	오차
297,417.98	306,301.25	8,883.27
271,365.13	263,467.93	7,897.20
331,867.17	323,992.00	7,875.17
248,802.38	256,352.51	7,550.13
240,271.68	231,306.50	8,965.18

→ 오차가 상대적으로 큰 20만~33만 매출 구간은, 모델 정밀도 개선의 주요 대상이 될 수 있음.

5. 광고 예산 시나리오 분석

총 광고 예산을 100,000으로 설정한 뒤, TV·라디오·SNS 간 다양한 비율로 예산을 분배하여 예측 매출을 시뮬레이션함.



- TV 광고 비중이 높을수록 예측 매출이 크게 증가
- 반대로 TV 비중이 낮고 라디오·SNS 중심일수록 매출이 급격히 감소
- 동일한 예산 내에서도 채널 배분 방식에 따라 최대 7배의 매출 차이가 발생함.

TV 비율	라디오 비율	SNS 비율	예측 매출
0.70	0.20	0.10	249,341.42
0.10	0.60	0.30	35,263.25

6. 결론

- TV 광고비는 매출과 가장 강한 상관관계(0.99)를 보이며, 다중 회귀 분석에서도 유일하게 통계적으로 유의한 변수로 확인됐다.
- 예측 모델은 평균 오차율(MAPE)이 1.72%로, 매우 높은 예측 정확도를 보였다.
- 광고 예산 시나리오 분석에서는 TV 중심의 예산 배분이 매출 효율을 극대화하는 전략으로 나타났고, TV 비중이 낮을수록 매출이 급격히 감소하는 패턴이 확인되었다.

TV에 예산을 집중하고, 라디오와 SNS는 보조 채널로 활용하는 전략이 효과적이다.