

# 꽃을 든 남자 고객 행동 데이터 분석 보고서

마케터 : 이진

2025.2.17

## 1. 개요

본 보고서는 '꽃을 든 남자'의 고객 행동 데이터를 분석하여 매출에 미치는 영향을 평가하고, 고객 행동의 매출 기여도를 극대화하기 위한 전략을 제안하는 것을 목표로 함.

## 2. 주요 지표 요약

고객 ID	나이	성별	상품 카테고리	구매 금액(원)	구매 빈도	제품 리뷰 평점	반품률	유입된 마케팅 채널	쿠폰 사용 여부	
0	1	56	Female	Skincare	72559.21	18.000000	3.056818	0.43	YouTube	0
1	2	46	Female	Haircare	81902.60	13.000000	2.000000	0.11	YouTube	0
2	3	32	Male	Fragrance	38123.27	7.000000	1.000000	0.48	Facebook	0
3	4	60	Male	Skincare	14136.77	4.000000	2.000000	0.01	Email	0
4	5	25	Male	Skincare	94349.71	9.511111	1.000000	0.48	Instagram	0
...	...	...	...	...	...	...	...	...	...	...
95	96	24	Male	Bodycare	17979.03	12.000000	4.000000	0.07	Email	0
96	97	26	Female	Bodycare	72346.54	3.000000	4.000000	0.01	Instagram	0
97	98	41	Female	Haircare	57517.90	1.000000	3.056818	0.18	Instagram	0
98	99	18	Male	Bodycare	33168.46	1.000000	3.000000	0.29	Website	0
99	100	61	Female	Haircare	44879.18	8.000000	3.000000	0.20	Email	0
100 rows x 10 columns										

총 100 명의 고객 데이터를 포함하며, 각 고객 ID, 나이, 성별, 상품 카테고리, 구매 금액(원), 구매 빈도, 제품 리뷰, 평점, 반품률, 유입된 마케팅 채널, 쿠폰 사용 여부를 주요 지표로 설정하고, 100 명의 고객의 데이터를 포함합니다.

## 3. 주요지표 분석

분석 방법: python 사용, csv 파일 불러온 후 DataFrame 변환 후 분석

```

import pandas as pd
import numpy as np

#1.csv df로 저장
df = pd.read_csv('꽃을 든 남자_고객행동분석.csv')

#2. csv 파일 확인 후 범주형 데이터 변환
#컬럼의 범주 갯수 파악 위해 unique() 사용
df['상품 카테고리'].unique()
df['유입된 마케팅 채널'].unique()

#원핫인코딩
df['상품 카테고리'] = df['상품 카테고리'].map({'Skincare':0, 'Haircare':1, 'Fragrance':2, 'Bodycare':3})
df['유입된 마케팅 채널'] = df['유입된 마케팅 채널'].map({'YouTube':0, 'Facebook':1, 'Email':2, 'Instagram':3, 'Website':4})
df['성별'] = df['성별'].map({'Female':0, 'Male':1})

#결측치 찾기
missing_values_example = df.isnull()

# 결측값 각 컬럼의 평균으로 채우기 # 평균은 항상 float 값을 반환
df.fillna(df.mean(numeric_only = True),inplace =True)

display(df)
print(df.dtypes)

```

```

#3질문 설정을 위한 상관계수 구하기

#종속변수 설정
target_variable = "구매 금액(원)"

#상관계수 계산
correlation_matrix = df.corr()

#종속변수와 상관계수 추출(자기자신제외)
independent_variables = correlation_matrix[target_variable].drop(target_variable).sort_values(ascending=False)

##결과출력
display("구매 금액과의 상관계수:", independent_variables)

```

```

#4시각화
import matplotlib.pyplot as plt
import seaborn as sns
!pip install koreanize-matplotlib
import koreanize_matplotlib

#구매자 성비율
gender_counts = df["성별"].value_counts()
plt.figure(figsize=(6, 6)) # 그래프 크기 조정
set1_colors = list(sns.color_palette("Set1", n_colors=len(gender_counts)))[::-1]
plt.pie(
    gender_counts, labels=gender_counts.index, autopct="%1.1f%%",
    colors=set1_colors, startangle=90
)
plt.title("상품 구매자 남녀 비율")
plt.show()

#카테고리별 성별 수
df["상품 카테고리"]=df["상품 카테고리"].replace([0,1,2,3],['Skincare', 'Haircare', 'Fragrance', 'Bodycare'])
df["성별"] = df["성별"].replace([0,1],["여자","남자"])
plt.figure(figsize=(8, 4))

sns.countplot(x="상품 카테고리", data=df, hue ="성별",palette='Set1')
plt.title("카테고리별 성별 수 비교")
plt.xlabel("상품 카테고리")
plt.ylabel("성별")
plt.show()

#나이대별 구매자 수
bins = [9, 19, 29, 39, 49, 59, 69, 100]
labels = ["10대", "20대", "30대", "40대", "50대", "60대", "70대 이상"]
df["나이대"] = pd.cut(df["나이"], bins=bins, labels=labels, right=True)
plt.figure(figsize=(8, 6))
sns.countplot(x="나이대", data=df, width=0.6, order=labels, palette="Set1")
plt.title("나이대별 구매자수")
plt.xlabel("나이")
plt.ylabel("구매자수")
plt.show()

# 상품카테고리별 나이대 분포
age_group_counts = df["나이대"].value_counts().sort_index()
sns.countplot(x="상품 카테고리", data=df, hue="나이대", palette="Set3")
plt.figure(figsize=(14, 14))
plt.title("카테고리별 나이대 분포")
plt.xlabel("상품 카테고리")
plt.ylabel("구매자수")
plt.legend(title="나이대", bbox_to_anchor=(1.05, 1), loc="upper left")# 범례 추가
plt.show()

```

```

#상품 리뷰 평점과 구매 빈도사이의 상관관계
plt.figure(figsize=(6, 4))
sns.scatterplot(x="제품 리뷰 평점", y="구매 빈도", data=df,palette="Set2")
plt.title("상품 리뷰 평점과 구매빈도 사이의 상관관계")
plt.xlabel("제품 리뷰 평점")
plt.ylabel("반품률")
plt.show()

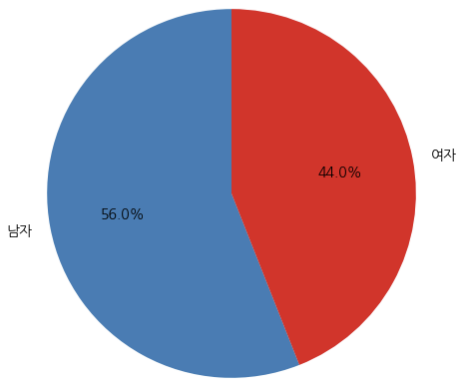
#상품 리뷰 평점과 반품률 사이의 상관관계
plt.figure(figsize=(6, 4))
sns.scatterplot(x="제품 리뷰 평점", y="반품률", data=df, hue="성별", palette="Set2")
plt.title("상품 리뷰 평점과 반품률 사이의 상관관계")
plt.xlabel("제품 리뷰 평점")
plt.ylabel("반품률")
plt.show()

```

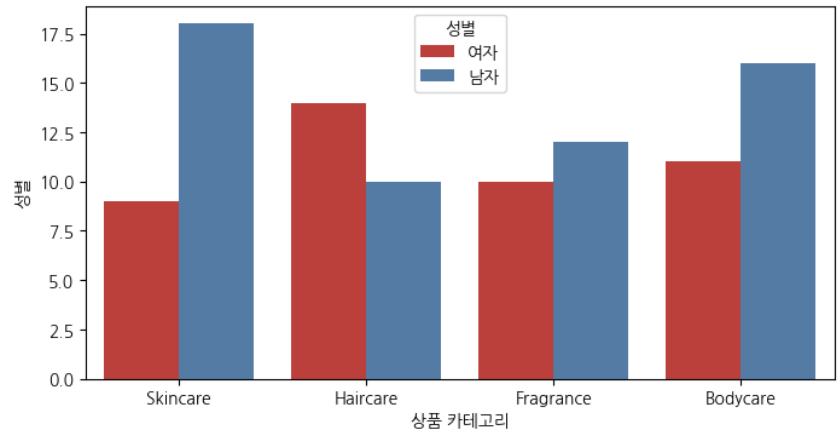
### 3.1 상품 카테고리별 성별 수 비교

#### 분석

상품 구매자 남녀 비율



카테고리별 성별 수 비교



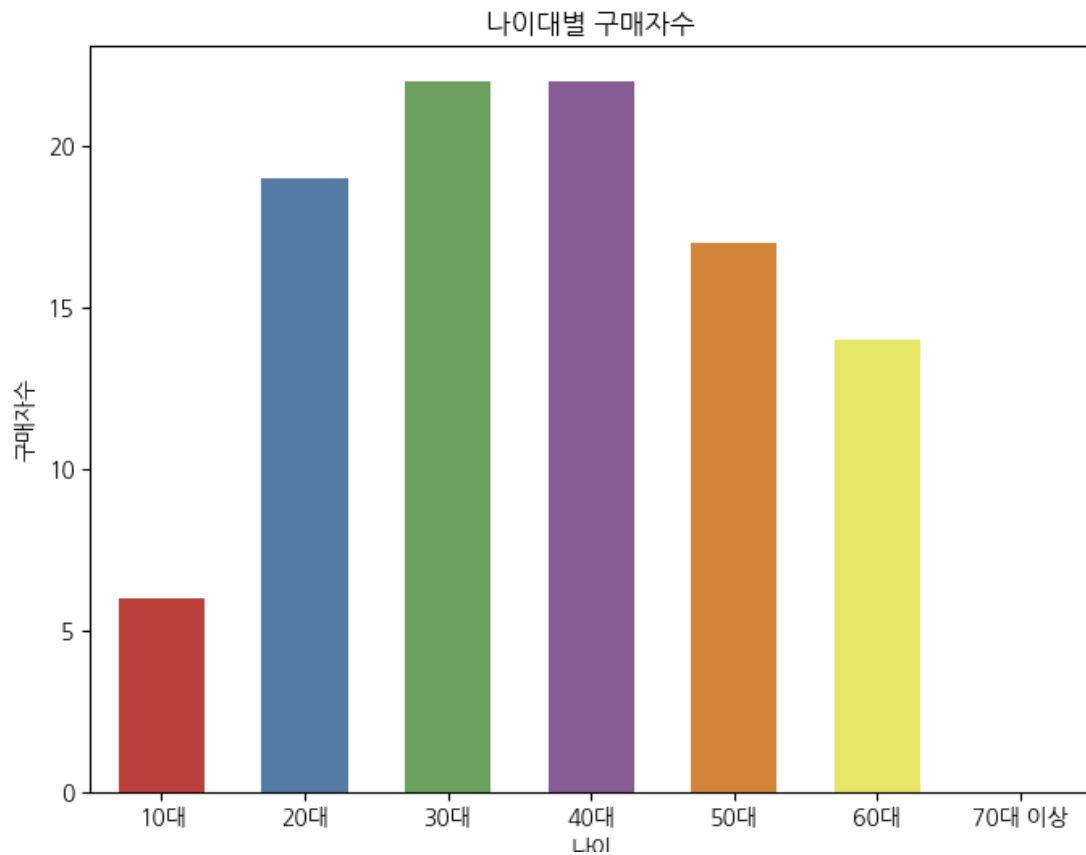
- Skincare 27 명, bodycare 27 명, haircare 24 명, fragrance 22 명의 고객이 구매함.
- 남성 고객을 타겟으로 한 브랜드이기 때문에 전체 고객 중 56%가 남성임.
- Skincare 상품은 남성 구매자(18 명)가 9 명 더 많음.
- Bodycare 상품은 남성 구매자(16 명)가 5 명 더 많음.
- 남성브랜드이지만 헤어상품 구매자(총 24 명)는 여성이 4 명 더 많음.
- 여성 고객을 타겟으로 한 브랜드에서는 헤어 왁스, 젤 상품을 따로 판매하는 경우가 드뭄.
- 헤어 고정제품이 필요한 여성들이 구매한 것으로 예상됨.
- 향수를 구매한 남성 고객과 여성 고객의 수 차이가 2 명임.
- 중성적인 향을 선호하는 여성이거나 선물을 위한 구매로 예측됨.

#### 개선 방향

- 남성 고객 유입 증진을 위한 마케팅 전략 재수립.
- 남성의 구매 심리, 행동을 분석한 제품 최적화.

### 3.2 나이대별 구매자 수

#### 분석



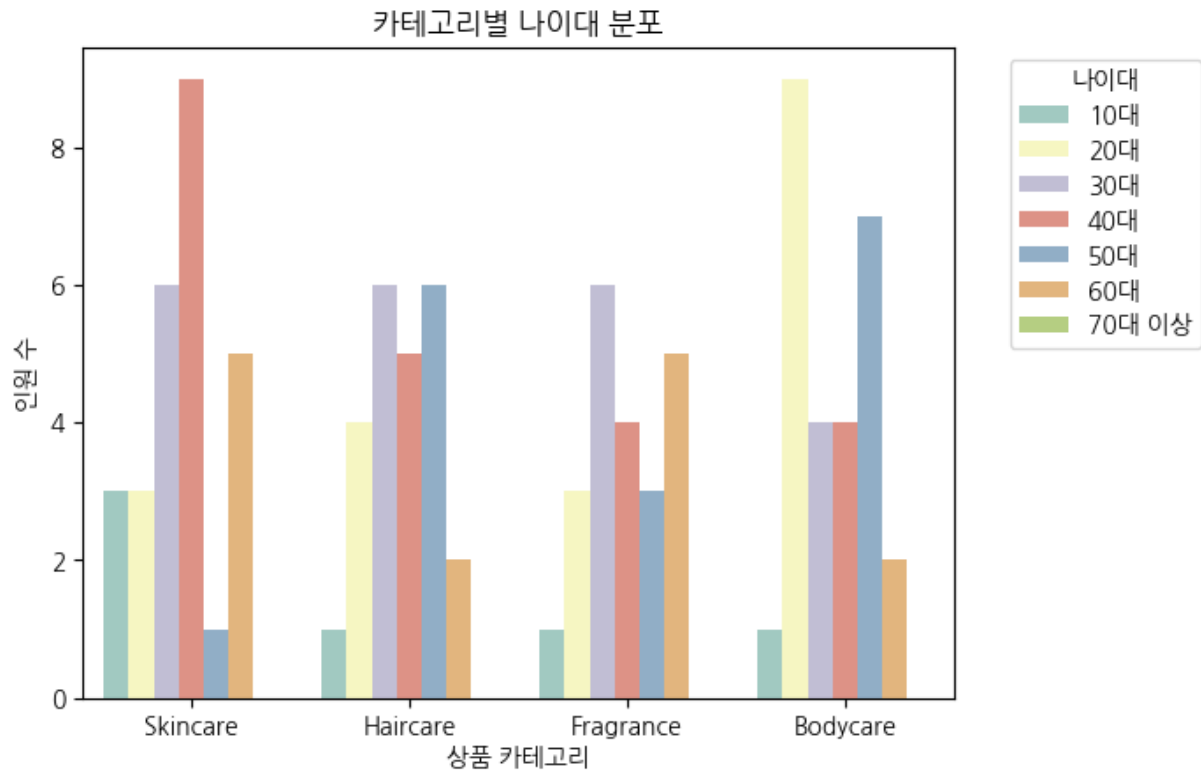
- 30 대와 40 대가 각각 22 명 구매함.
- 20 대가 19 명, 50 대가 17 명, 60 대가 14 명, 10 대가 6 명 구매함.
- 근로를 통한 소득이 왕성한 20 대-40 대가 구매율이 높음.
- 정년 퇴직 후 배움과 소비를 통해 활동적인 삶을 영위하는 5060 세대를 타겟으로 한 마케팅 전략 필요.

#### 개선방향

- 20-40 세대의 고객 충성도 향상을 위한 리텐션 전략 강화
- 멤버십 혜택, 쿠폰 사용 과정 단순화
- 10 대와 5060 세대의 신규 고객 유입을 위한 상품 개발 및 마케팅 전략 수립.
- 비교적 소비 금액에 여유가 있는 5060 을 대상으로 한 프리미엄 라인 출시

### 3.3 상품 카테고리별 나이대 분포

분석



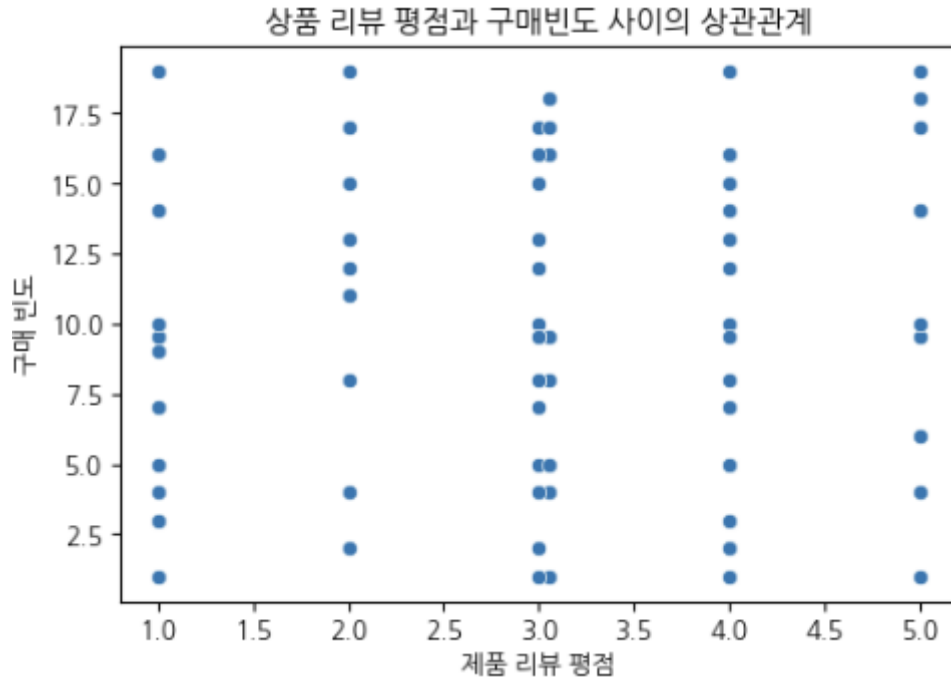
- Skincare 상품은 40 대가 앞도적으로 구매율이 높음.
- Haircare 상품은 30-50 대가 많이 구매함.
- Fragrance 상품은 30 대가 가장 많이 구매함.
- Bodycare 상품은 20 대의 구매 비중이 상당히 큼.
- 전반적으로 10 대 구매자 수가 낮음.
- 10 대는 3040 이 10 대 뭉까지 함께 소비하는 경향이 있음.

개선 방향

- 카테고리 별 전반적인 연령대 구매자 증가 위한 보편적이지만 브랜드 개성이 뚜렷한 상품 개발 필요
- 브랜드 이미지 인식 분석 후 마케팅 전략 개선

### 3.4 상품 리뷰 평점과 구매빈도 사이의 상관관계

#### 분석



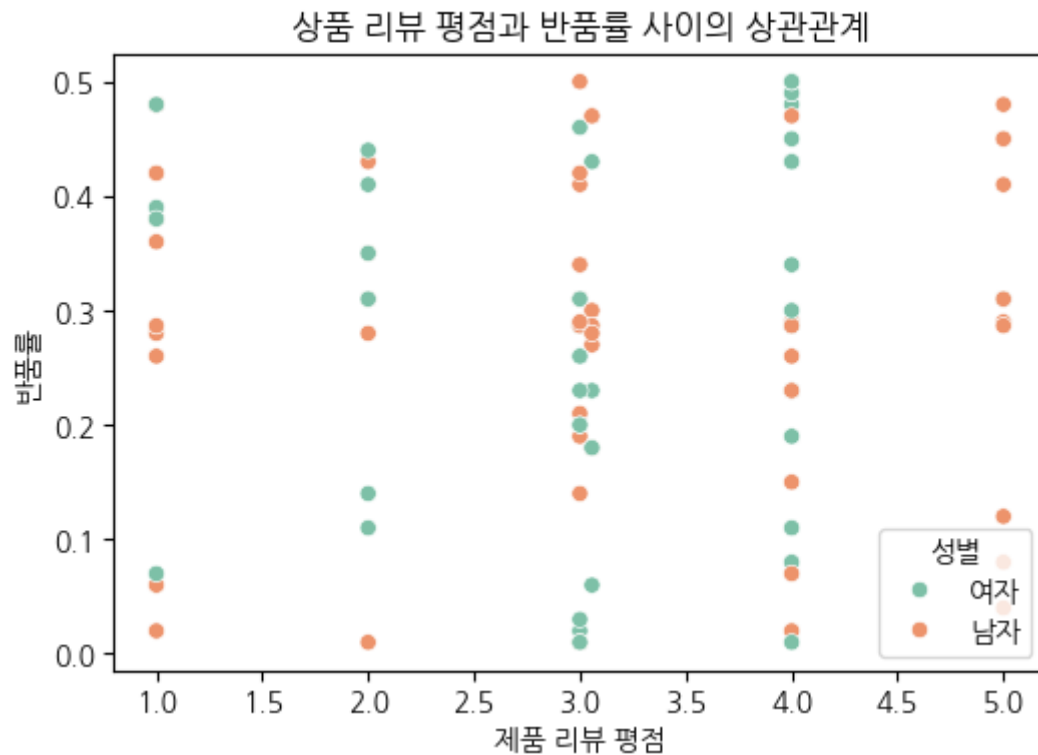
- 리뷰 평점이 1~5 점으로 구분되어 있으며, 각 평점에서 다양한 구매 빈도가 나타남.
- 특정한 패턴 없이 평점별로 고르게 분포된 것으로 보임.
- 리뷰 평점이 낮아도(1~2 점) 일정 수준의 구매 빈도가 존재함.
- 전반적으로 명확한 상관관계가 없어 보이며, 평점이 높다고 해서 반드시 구매 빈도가 높은 것은 아님.

#### 개선 방향

- 평점 시 질문 리스트 세분화하여 고객 만족도 심층 분석
- 구매 빈도와 상관관계가 높은 특징을 찾아 새로운 전략 수립.

### 3.5 상품 리뷰 평점과 반품률 사이의 상관관계

분석



- 특정한 패턴 없이 평점별로 반품률이 고르게 분포된 것으로 보임.
- 리뷰 평점이 높아도 높은 반품률이 존재함.
- 제품 리뷰 평점과 반품률 간의 상관관계가 높은것으로 보이지 않으며 평점이 낮다고해서(1~2 점) 반드시 반품률이 높지 않음.

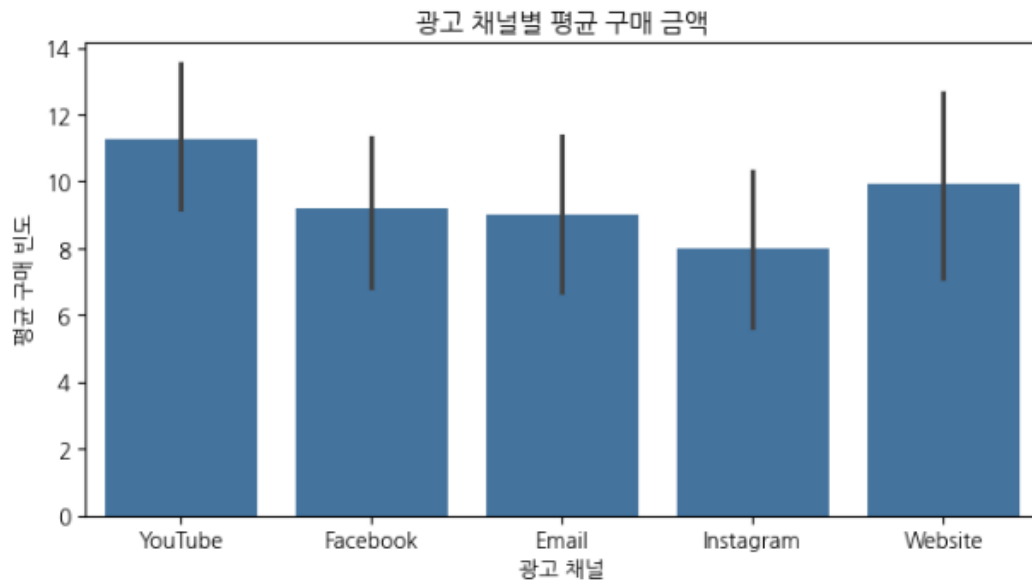
개선방향

- 반품률과 상관관계가 높은 특징을 찾아 반품률을 낮출 전략을 세워야함.
- 평점은 제품 자체의 만족도에 영향을 많기 때문에 전반적인 제품 점검이 필요.



### 3.6 유입된 광고 채널별 성과 분석

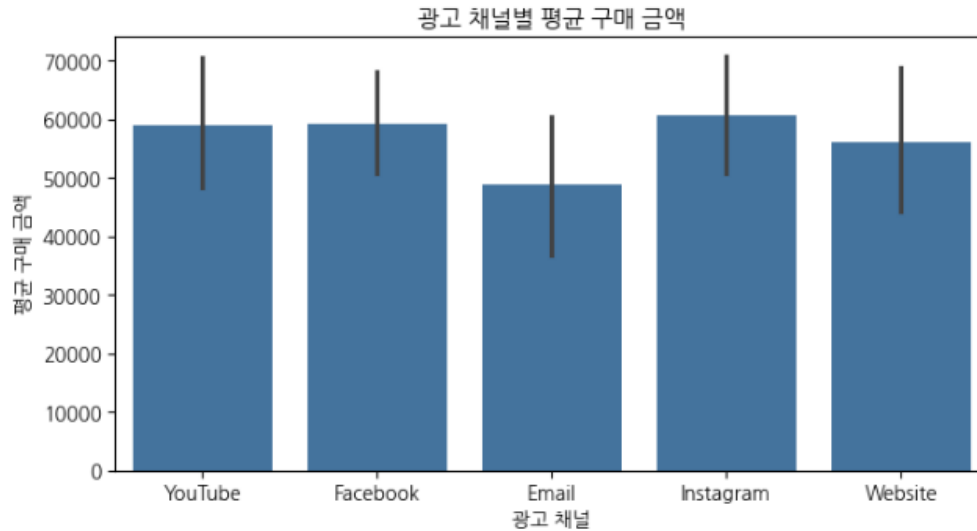
#### 분석



- 광고 채널 별 평균 구매 빈도는 YouTube가 11회로 가장 높고, Website 10회, Facebook 9회, Email 9회, Instagram 8회 순으로 높음.
- YOUTUBE의 구매 전환율이 가장 높음
- 구매 전환율이 낮은 채널의 클릭률 분석 후 낮을 경우 제품 상세 페이지 개선 필요할 것으로 예상.
- 전환율이 높은 채널은 해당 채널의 유사 타겟을 확장.
- 광고 채널의 특성 분석 후 구매율 상승 전략이 필요.
- ROAS 분석 후 성과 높은 마케팅 광고 채널에 예산 집중 분배.

#### 개선 방향

- 광고 채널의 특징을 반영한 콘텐츠 제작 필요
- 전반적인 광고 채널의 내부링크 최적화(광고 채널과 사이트간 긴밀한 연결성 필요.)
- 광고 채널에서 구매로 이어지는 경로 단순화.
- 구매 빈도가 높은 충성 고객 유지 전략 필요



- 유입된 마케팅 채널 별 구매자의 평균 구매금액은 Instagram 60759원으로 가장 높고, Facebook 59240원, YouTube 59072원, Website 56135원, Email : 48763원 순으로 높음.
- Instagram를 통해 유입된 고객의 평균 구매 금액이 가장 높음.
- 그 다음으로 높은 youtube(59072원)은 평균 구매 횟수도 높아 전환율이 가장 높은 채널임.
- 단순히 평균 구매 금액이 높다고 전환율도 높다 판단할 수 없음
- 평균구매금액은 높아도 방문자 중 실제 구매하는 사람은 적으면 전환율이 낮다고 보기 때문임.
- 특정 광고 채널이 구매력이 높은 소수 고객을 타겟 하면 평균구매금액은 높지만 전환율은 낮을 수 있음.
- 고가 제품이라면 방문자 대비 구매자가 적을 수 있음.

#### 개선 방향

- 평균구매 금액을 높이기 위한 업셀링 과 크로셀링 전략 적용
- 타겟별 맞춤 셀링 전략 필요.

#### 4. 종합 결론

지표	높은 경우	낮은 경우	개선 방법
남성 구매자 증가	타겟팅 정확도 높음	타겟팅 정확도 떨어짐	광고 최적화, 고객 맞춤 상품 개발
3040 세대 증가	상품 마케팅 효과 높음	마케팅 효과 낮음	관련 콘텐츠 추천, UX 개선
평균 구매 빈도 높음	콘텐츠 연계성 높음	재방문자 위주	충성 고객 유지 전략 필요
평균 구매 금액 높음	콘텐츠 연계성 높음	한 페이지만 보고 이탈	내부 링크 최적화, UX 개선

#### 5. 최종 제안

1. 신규 사용자 유입을 지속적으로 늘리면서 기존 고객 유지 전략 강화
  - 이메일 마케팅, 맞춤형 추천 기능 활용.

2. **페이지 간 이동을 원활하게 하여 사용자 경험(UX) 개선**
  - 네비게이션 최적화, 관련 콘텐츠 추천 추가.
3. **마케팅 성과 분석 후, 효과적인 광고 채널을 집중적으로 활용**
  - 트래픽 증가 분석 후 가장 성과가 좋은 광고 채널에 예산 집중.