

# **STARBUCKS**

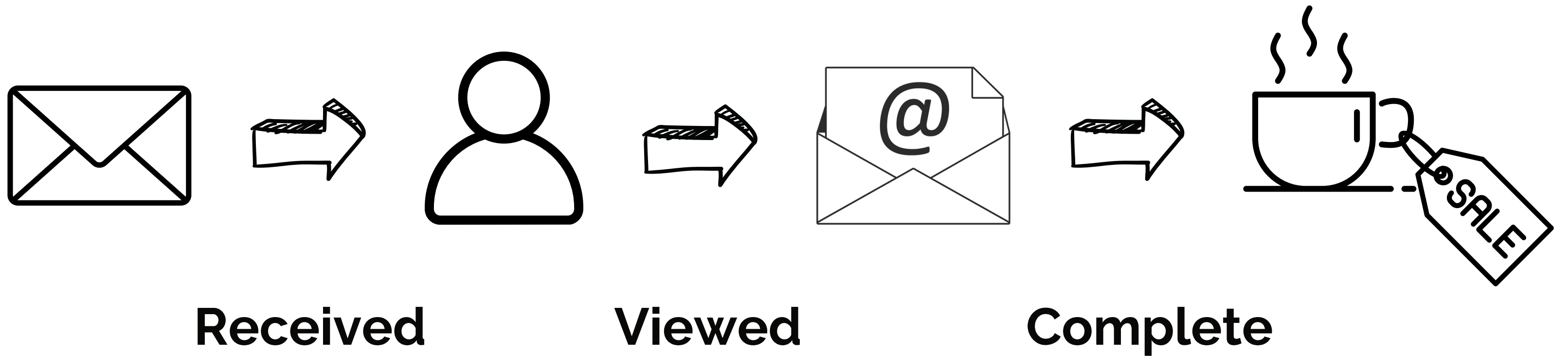
## **MARKETING**



# Contents

<b>01</b>	_____	프로젝트 개요
<b>02</b>	_____	데이터 전처리
<b>03</b>	_____	데이터 EDA
<b>04</b>	_____	통계 과정 검토
<b>05</b>	_____	인사이트 및 비즈니스 전략

# Question



# 프로젝트 개요

---

# 프로젝트 목표

---

## 프로모션 효율 극대화

- 어떤 오퍼 유형이 가장 높은 전환율을 보이는가?
- 조회여부가 전환율에 미치는 영향은?
- 채널 다양성이 전환율 향상에 기여하는가?

## 고객 세그먼트별 타겟 전략 수립

- 성별에 따른 평균 거래금액 차이는?
- 나이대별 거래금액의 관계는?
- 소득 수준과 거래금액의 관계는?

# 프로젝트 목표

---

- 전환율과 평균 거래금액 **10%** 초과 달성
- 프로모션 ROI **2.0** 이상 개선

# 데이터 전처리

---

# 전처리

Starbucks Customer Data.csv

데이터	컬럼(Column)	행(Row)	데이터 설명
portfolio	7	10	프로모션 정보 및 배포 방식
profile	6	17,000	고객 정보
transcript	5	306,534	프로모션 혜택 및 거래내역



# 전처리

## [1] 필수 컬럼 선별

데이터 카테고리 내 컬럼 묶음을 표시한 것

대표컬럼	한글설명	유형구분
person	고객 고유 ID	범주형
event	프로모션 행위 (받음, 조회, 완료)	
offer_type	프로모션 유형	
id_x	고객ID	
gender	성별	
merged_offer_id	프로모션 고유 ID	
id_y	고객 프로모션 닉네임ID	
channels	프로모션 받은 경로	

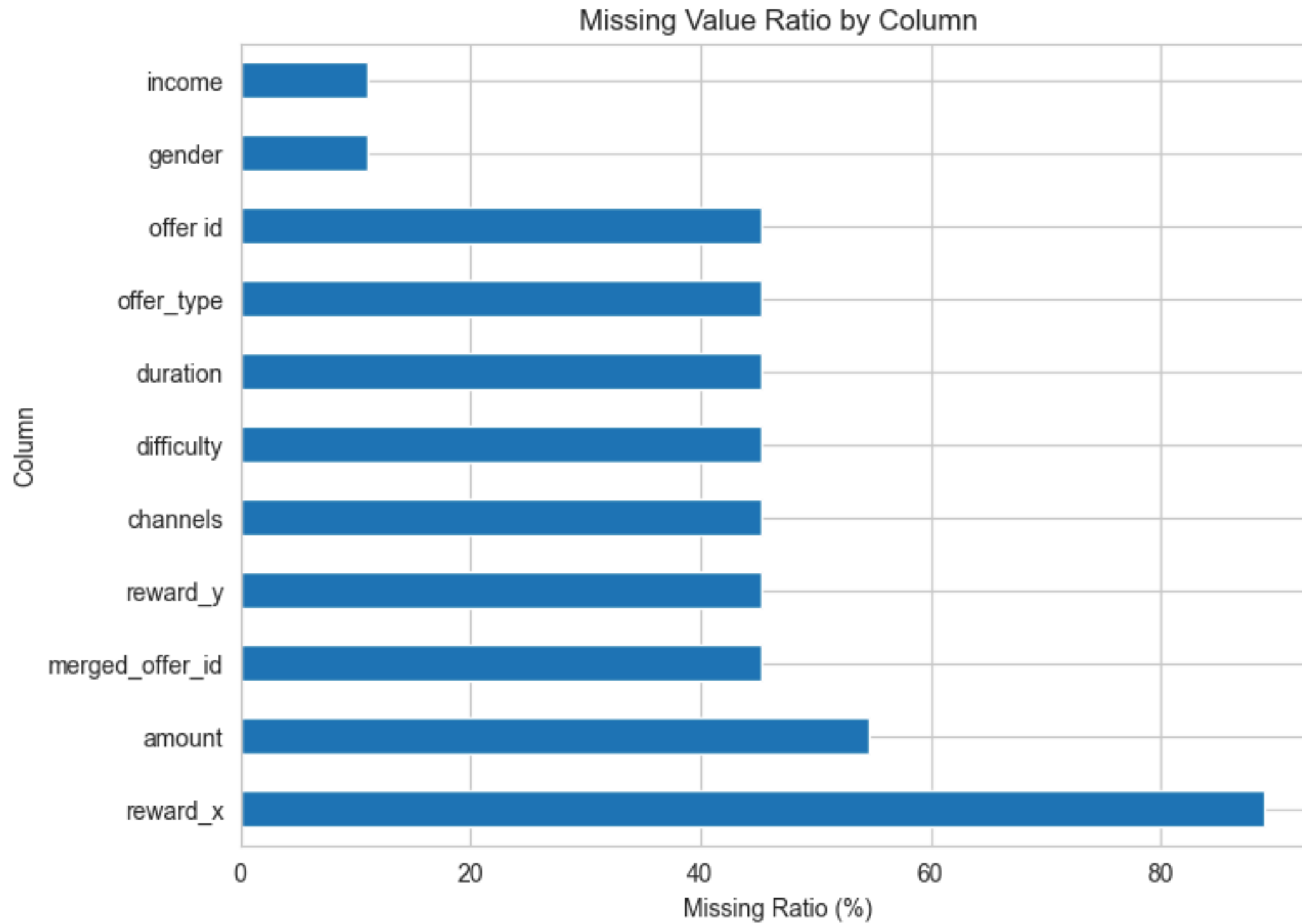
# 전처리

## [1] 필수 컬럼 선별

데이터 카테고리 내 컬럼 묶음을 표시한 것

대표컬럼	한글설명	유형구분
difficulty	프로모션 완료까지 남은 금액	수치형
duration	프로모션 기간	
time	프로모션 시간	
amount	결제 금액	
reward_x	리워드 금액	
age	나이	
reward_y	프로모션 적립 금액	
become_member_on	멤버 가입 날짜	
income	소득	

# 결측치 비율



- 수치형: 결측치는 0으로 통일
- 문자형: 결측치는 'Unknown'으로 대체  
→ 데이터 손실 최소화 및 통계 왜곡 방지

# 데이터 EDA

---

# 사용된 데이터

## 주요 컬럼

컬럼	컬럼 설명	사용
value	프로모션 정보 집합	파싱(offer id, amount)
person	고객 고유 ID	데이터 병합
event	프로모션 행위(받음, 조회, 완료)	분석
channels	프로모션 받은 경로	분석
offer_type	프로모션 유형	분석

# 사용된 데이터

---

## 주요 컬럼 - value

- 이 컬럼은 딕셔너리 형태로 되어 있기에 **파싱 작업**이 필요
- **키값은 컬럼**으로 **값은 행**으로

value 컬럼의 키값 들['amount', 'offer id', 'offer\_id', 'reward']

```
{'offer id': '3f207df678b143eea3cee63160fa8bed'}
```

```
{'offer id': '0b1e1539f2cc45b7b9fa7c272da2e1d7'}
```

```
{'offer id': '3f207df678b143eea3cee63160fa8bed'}
```

```
{'amount': 34.56}
```

```
{'amount': 13.23}
```

```
{'amount': 19.51}
```

\* 오퍼아이디는 2개 컬럼 있음 → 두 컬럼 겹치는 값이 없기때문에 병합 진행

# 통계검정 ①



# 퍼널 분석

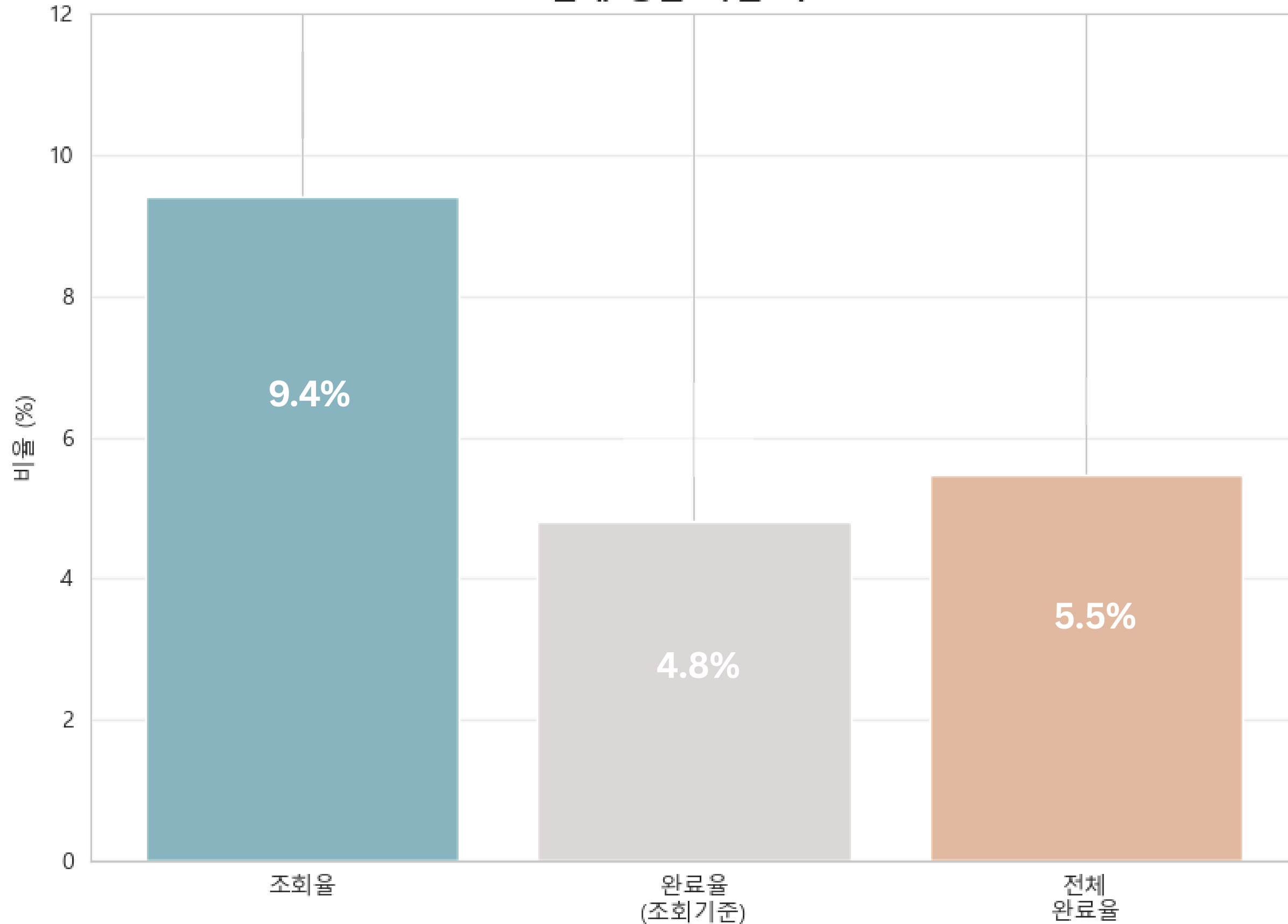
단계별 전환율을 통해 고객 이탈 구간과 프로모션 성과를 진단

단계	의미	전환율 계산식	해석
① 오퍼 수신 (offer_received)	고객이 오퍼 받음	—	—
② 오퍼 조회 (offer_viewed)	받은 오퍼 열람함	<ul style="list-style-type: none"><li>조회 전환율 = 조회/수신 (View Rate)</li></ul>	오퍼 받은 고객 중 조회한 비율
③ 오퍼 완료 (offer_completed)	프로모션 조건 충족	<ul style="list-style-type: none"><li>완료 전환율 = 완료/조회 (Completion Rate)</li><li>전체 전환율 = 완료/수신 (Overall Completion Rate)</li></ul>	<ul style="list-style-type: none"><li>조회 고객 중 완료한 비율</li><li>오퍼 받은 고객 중 최종 완료한 비율</li></ul>



# 고객 ID x 오퍼 ID x 오퍼 유형 → 전환율

전체 평균 퍼널 지표

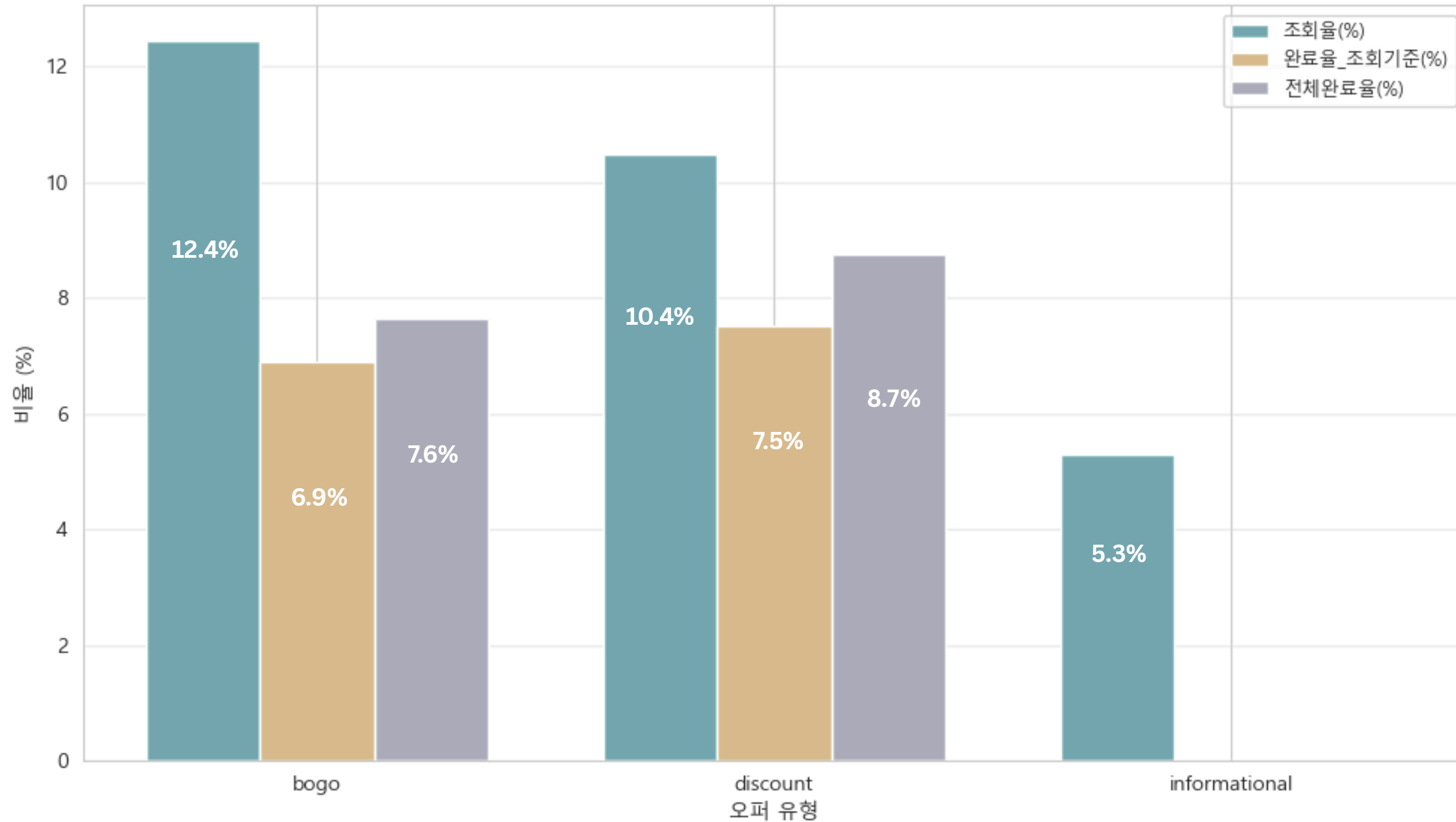


## 병목 구간 분석

- 평균 조회율 : 9.40%
- 평균 완료율 : 4.81%
- 평균 전체 완료율 : 5.47%  
→ 조회율이 낮음 (< 50%)

# 고객 ID x 오퍼 ID x 오퍼 유형 → 퍼널지표

오퍼 유형별 퍼널 지표 비교



## BOGO

- 조회률 : 12.4%
- 완료율 : 6.9%
- 전체완료율 : 7.6%

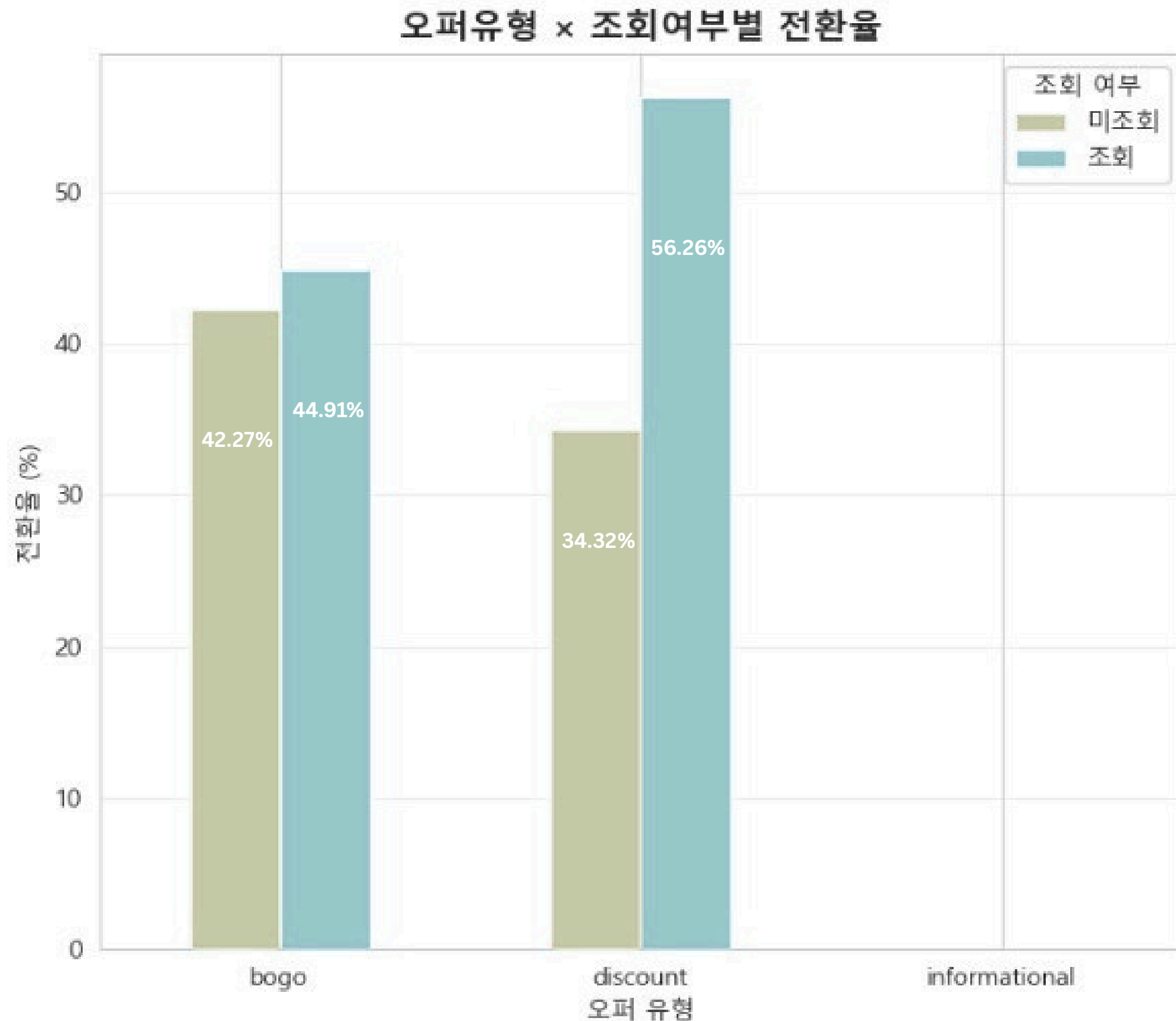
## Discount

- 조회률 : 10.4%
- 완료율 : 7.5%
- 전체완료율 : 8.7%

## Informational

- 조회률: 5.3%

# 오퍼유형 × 조회여부 → 전환율



## Discount

- 가장 효과적인 오퍼
- 조회 시 전환율: 56.26%
- 미조회 시 전환율: 34.32%
- 전환율 개선 핵심 → 오퍼 조회율 향상 필요

# 카이제곱 독립성 검정

검정	Cramér's V
<p>[검정 A: 오퍼 유형 × 완료 여부]</p> <ul style="list-style-type: none"><li>• 카이제곱 통계량: 15554.3869</li><li>• 자유도: 2</li><li>• p-value: 0.0000000</li><li>• ✓ 결과: <math>p &lt; 0.001 \rightarrow</math> 귀무가설 기각 (매우 강한 연관성)</li></ul>	<p>[검정 A] Cramér's V = 0.1747 해석: <b>중간 연관성</b></p>
<p>[검정 B: 조회 여부 × 완료 여부]</p> <ul style="list-style-type: none"><li>• 카이제곱 통계량: 197081.8593</li><li>• 자유도: 1</li><li>• p-value: 0.0000000</li><li>• ✓ 결과: <math>p &lt; 0.001 \rightarrow</math> 귀무가설 기각 (매우 강한 연관성)</li></ul>	<p>[검정 B] Cramér's V = 0.6217 해석: <b>강한 연관성</b></p>
<p>[검정 C: (오퍼유형 × 조회여부) × 완료여부]</p> <ul style="list-style-type: none"><li>• 카이제곱 통계량: 254381.0422</li><li>• 자유도: 5</li><li>• p-value: 0.0000000</li><li>• ✓ 결과: <math>p &lt; 0.001 \rightarrow</math> 귀무가설 기각 (매우 강한 연관성)</li></ul>	<p>[검정 C] Cramér's V = 0.7064 해석: <b>강한 연관성</b></p>

# 카이제곱 독립성 검정

## [WEB 채널]

- $\chi^2 = 4588.12$ , p-value = 0.00
- 유의함 ( $p < 0.05$ )

### 전환율:

- WEB O: 54.5%
- WEB X: 29.2%
- 차이: +25.3%p (1.86배)

## [EMAIL 채널]

- $\chi^2 = 0.00$ , p-value = 1.00
- 비유의 ( $p \geq 0.05$ )

### 전환율:

- EMAIL O: 49.6%
- EMAIL X: 49.6%

## [MOBILE 채널]

- $\chi^2 = 1.89$ , p-value = 0.17
- 비유의 ( $p \geq 0.05$ )

### 전환율:

- MOBILE O: 49.5%
- MOBILE X: 50.3%
- 차이: -0.7%p (0.99배)

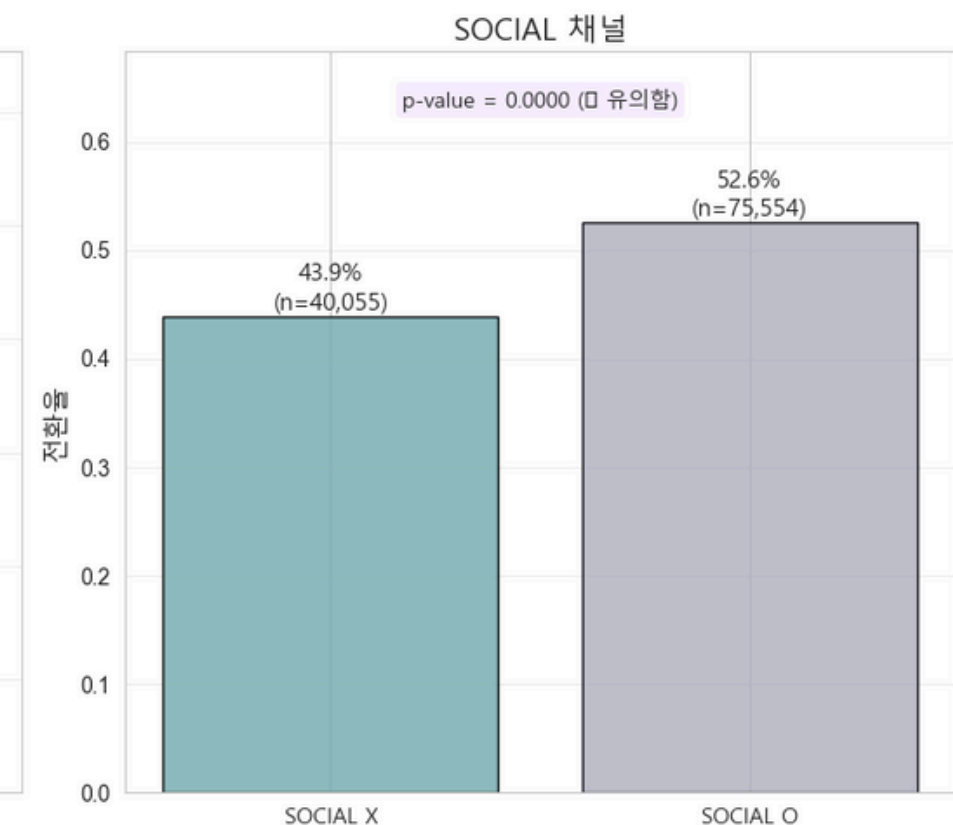
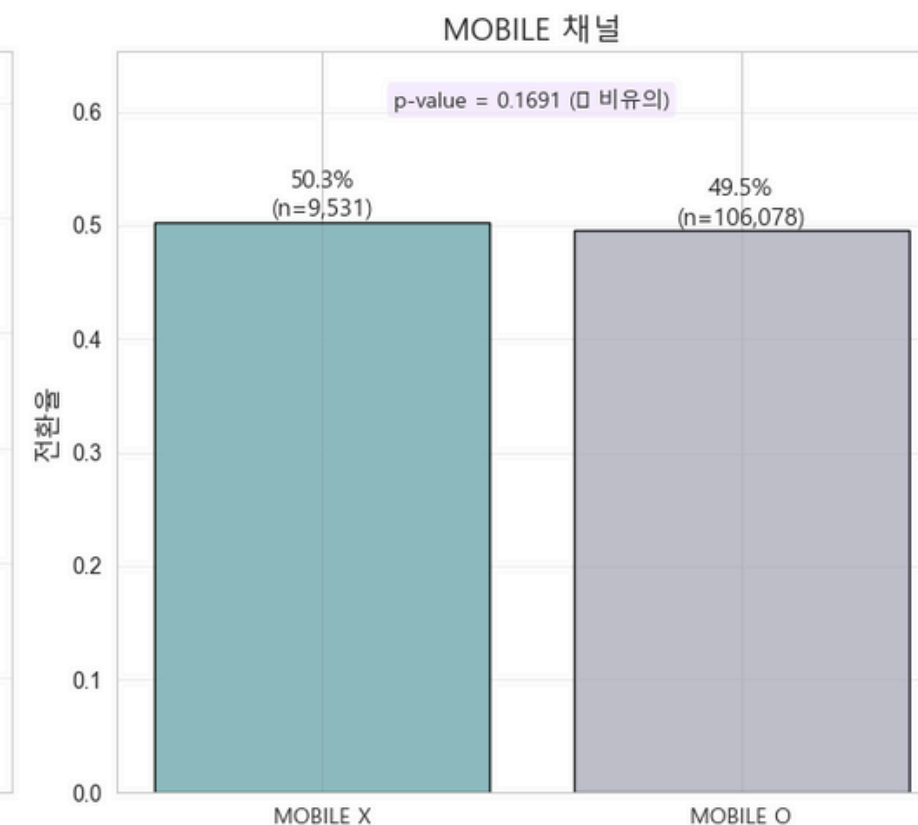
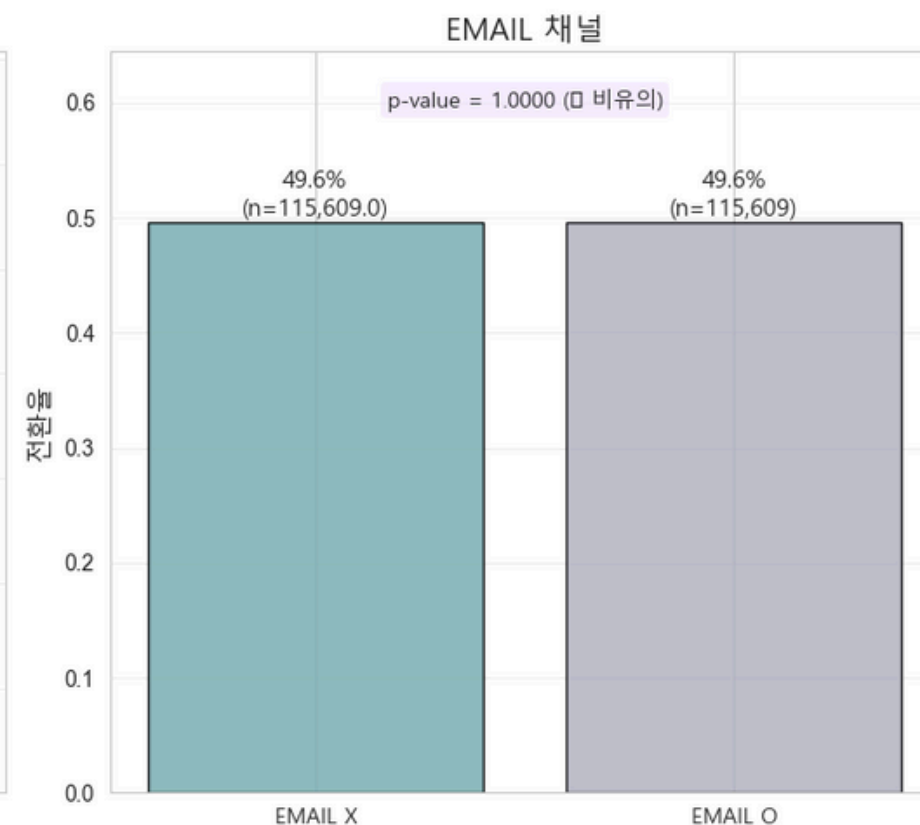
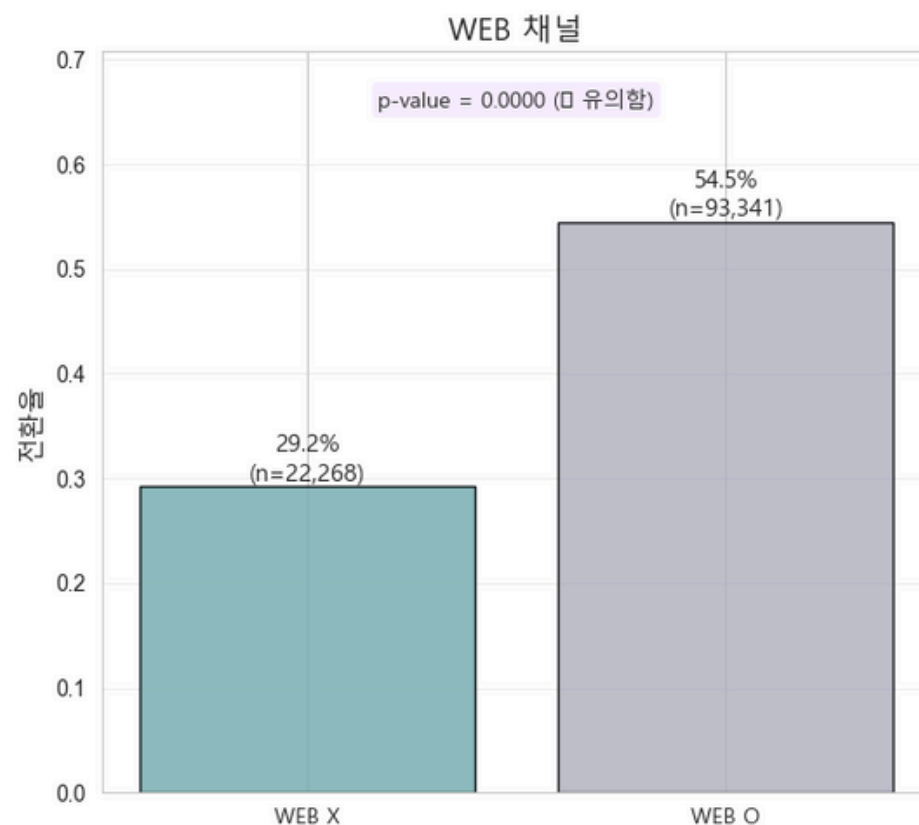
## [SOCIAL 채널]

- $\chi^2 = 781.94$ , p-value = 0.00
- 유의함 ( $p < 0.05$ )

### 전환율:

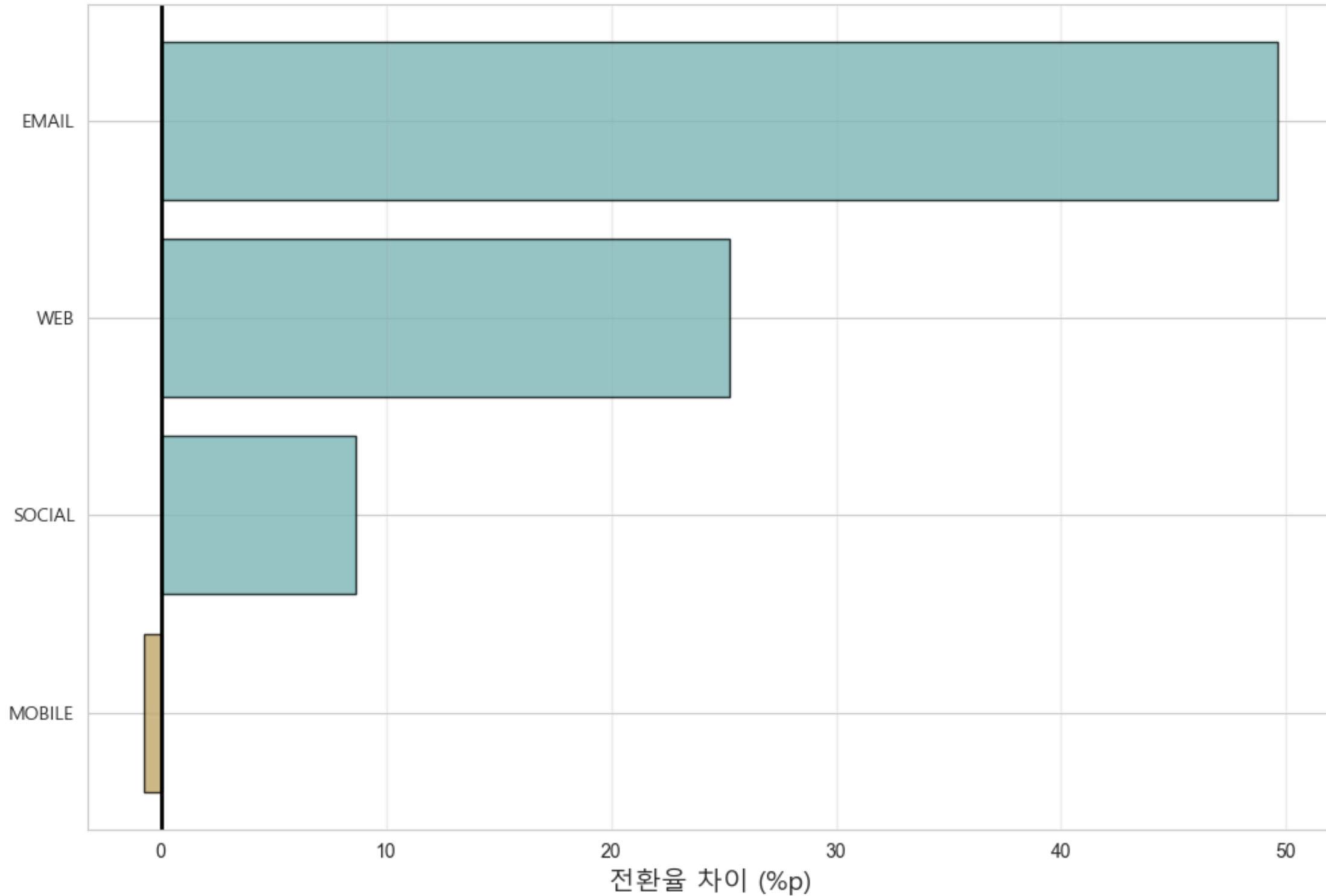
- SOCIAL O: 52.6%
- SOCIAL X: 43.9%
- 차이: +8.6%p (1.20배)

개별 채널별 전환율 비교



# 채널별 효과 순위

채널 사용 여부에 따른 전환율 차이



## WEB

- 전환율 +25.3%p ( $p=0.0000$ )
- 채널 사용시: 54.5%
- 미사용시: 29.2%

## SOCIAL

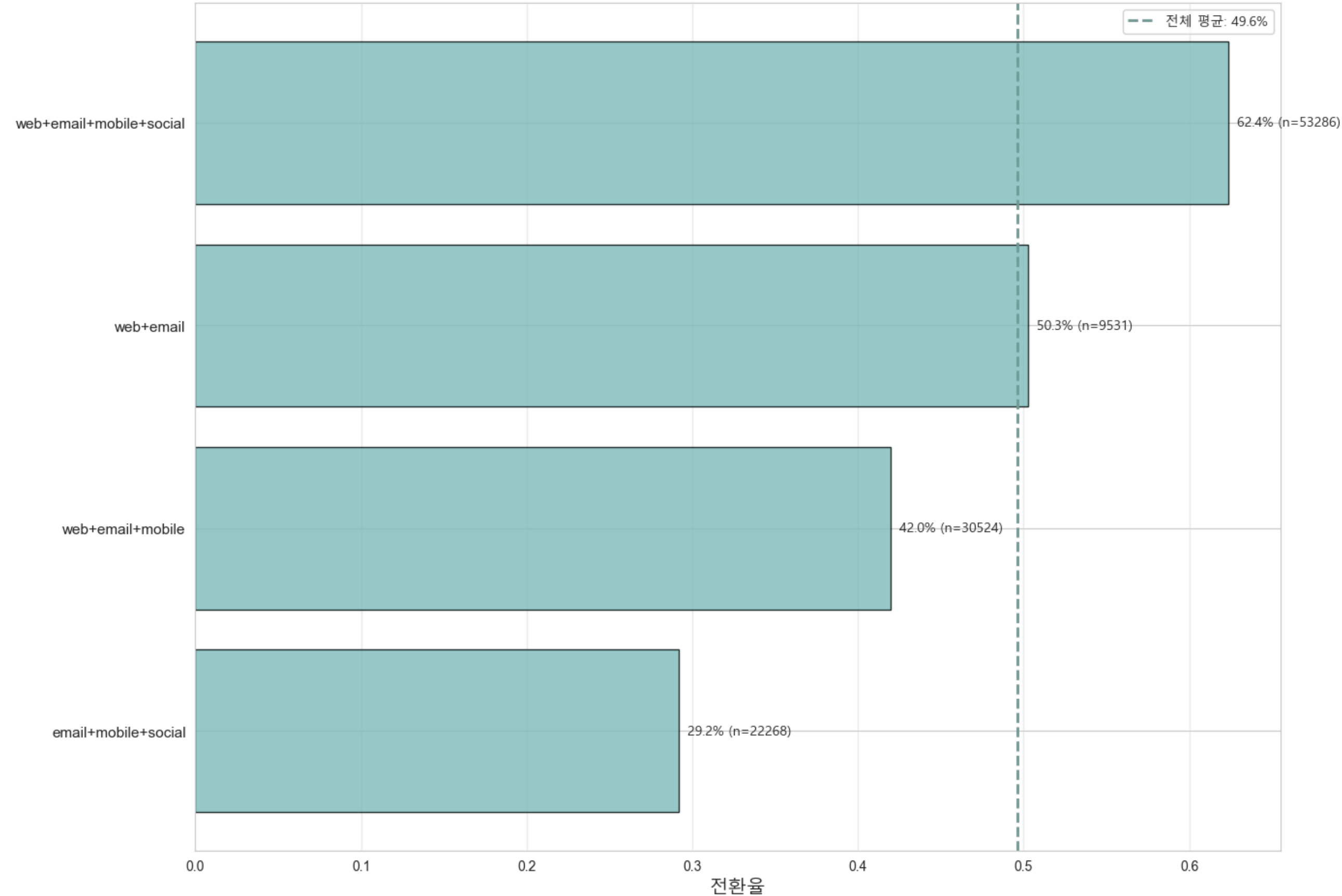
- 전환율 +8.6%p ( $p=0.0000$ )
- 채널 사용시: 52.6%
- 미사용시: 43.9%

## EMAIL & MOBILE

- $p=1.000$  (비유의)
- $p=0.169$  (비유의)
- 효과 미미한 채널 (유지 여부 재검토)

# 채널 조합별 전환율 분석

채널 조합별 전환율 Top 15 (샘플 100개 이상)

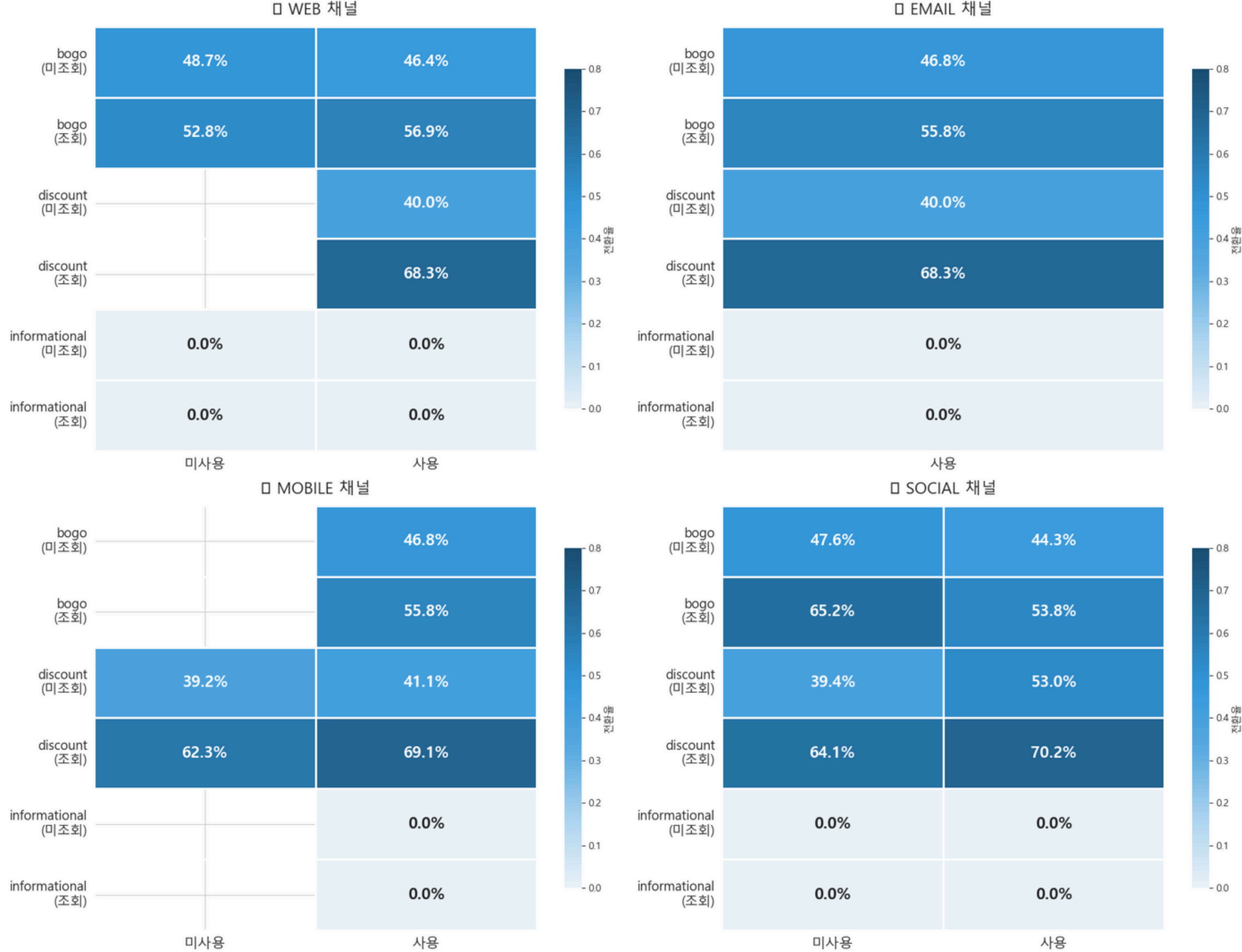


## WEB+SOCIAL+EMAIL+MOBILE

- 전환율: 62.4% (n=53286)
- 전체 평균 대비: +25.7% 상승
- ✓ 최고 채널 조합

# [오퍼 유형 × 조회 여부 × 개별 채널 → 전환율]

채널별 전환율 분석: 오퍼 유형 × 조회 여부



## 최고 성과 조합

- **Discount** : 전환율 65~69%
- **Bogo** : 전환율 50~55%

## 채널별 특징

- **Web: Discount** : 전환율 68.3%
- **Email / Social**: 전반적 균형 잡힌 성과
- **Mobile**: 전환율 변동폭 큼 (채널불안정 )



# 인사이트 및 비즈니스 전략

---

# 인사이트 및 비즈니스 전략

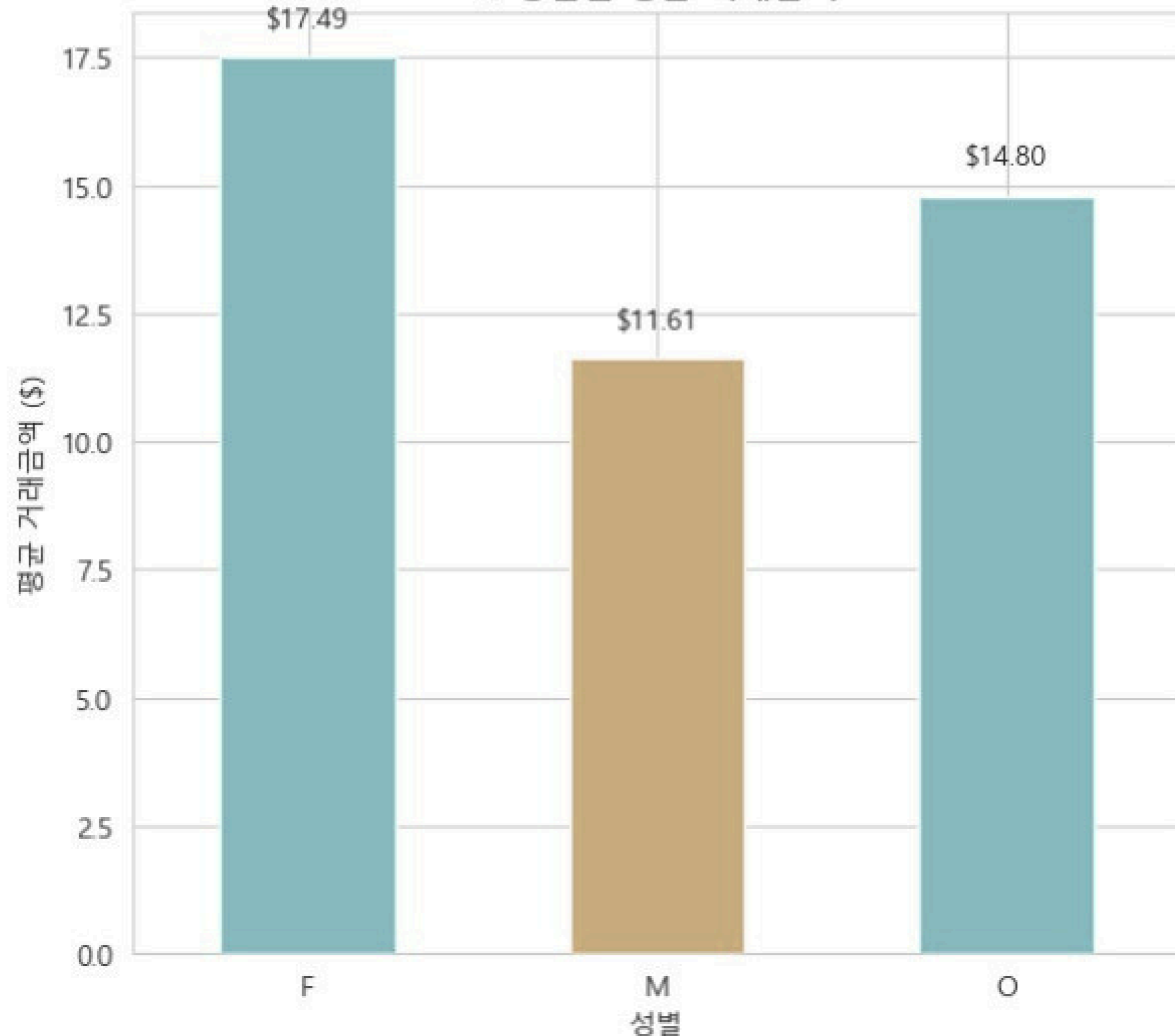
인사이트	비즈니스 전략
<b>WEB</b> <ul style="list-style-type: none"><li>채널 사용시: 54.5%</li><li>미사용시: 29.2%</li></ul>	<b>WEB 집중 투자</b> <ul style="list-style-type: none"><li>전환율 +25.3%p 증가 효과</li><li>마케팅 예산의 50% 이상 배정 권장</li></ul>
<b>EMAIL &amp; MOBILE</b> <ul style="list-style-type: none"><li>p=1.000 (비유의)</li><li>p=0.169 (비유의)</li><li>효과 미미한 채널 (유지 여부 재검토)</li></ul>	<b>EMAIL, MOBILE</b> <ul style="list-style-type: none"><li>비효율 채널 예산 축소 대상</li><li>절감 예산 → 효과적 채널로 이동</li></ul>
<b>WEB+SOCIAL+EMAIL+MOBILE</b> <ul style="list-style-type: none"><li>전환율: 62.4% (n=53286)</li><li>전체 평균 대비: +25.7% 상승</li></ul>	<b>EMAIL, MOBILE</b> <ul style="list-style-type: none"><li>비효율 채널 예산 축소 대상</li><li>절감 예산 → 효과적 채널로 이동</li></ul>

# 통계 검정 ②



# [남성/여성 고객 → 평균 거래금액]

1. 성별별 평균 거래금액



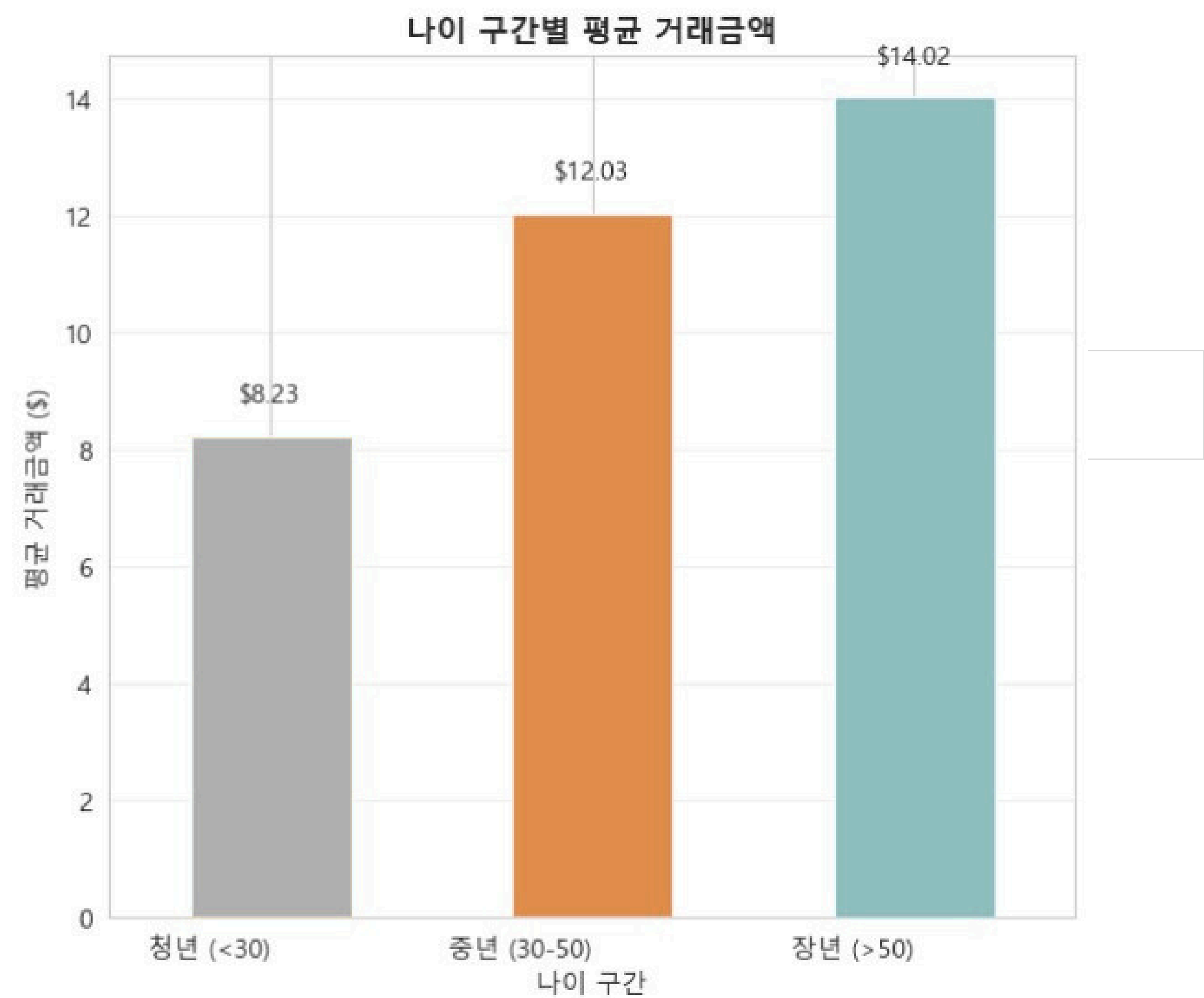
## 기술통계:

- 남성 평균 거래금액: \$11.61 (n=72794)
- 여성 평균 거래금액: \$17.49 (n=49382)

## 검정 결과:

- t-statistic: -30.6616
- p-value: 0.00000 (유의미한 차이 있음)
- 유의수준  $\alpha=0.05$  기준: 귀무가설 기각
- 효과크기 (Cohen's d): -0.1821

# [나이 → 평균 거래금액]



## Pearson 상관분석 결과 (선형 관계):

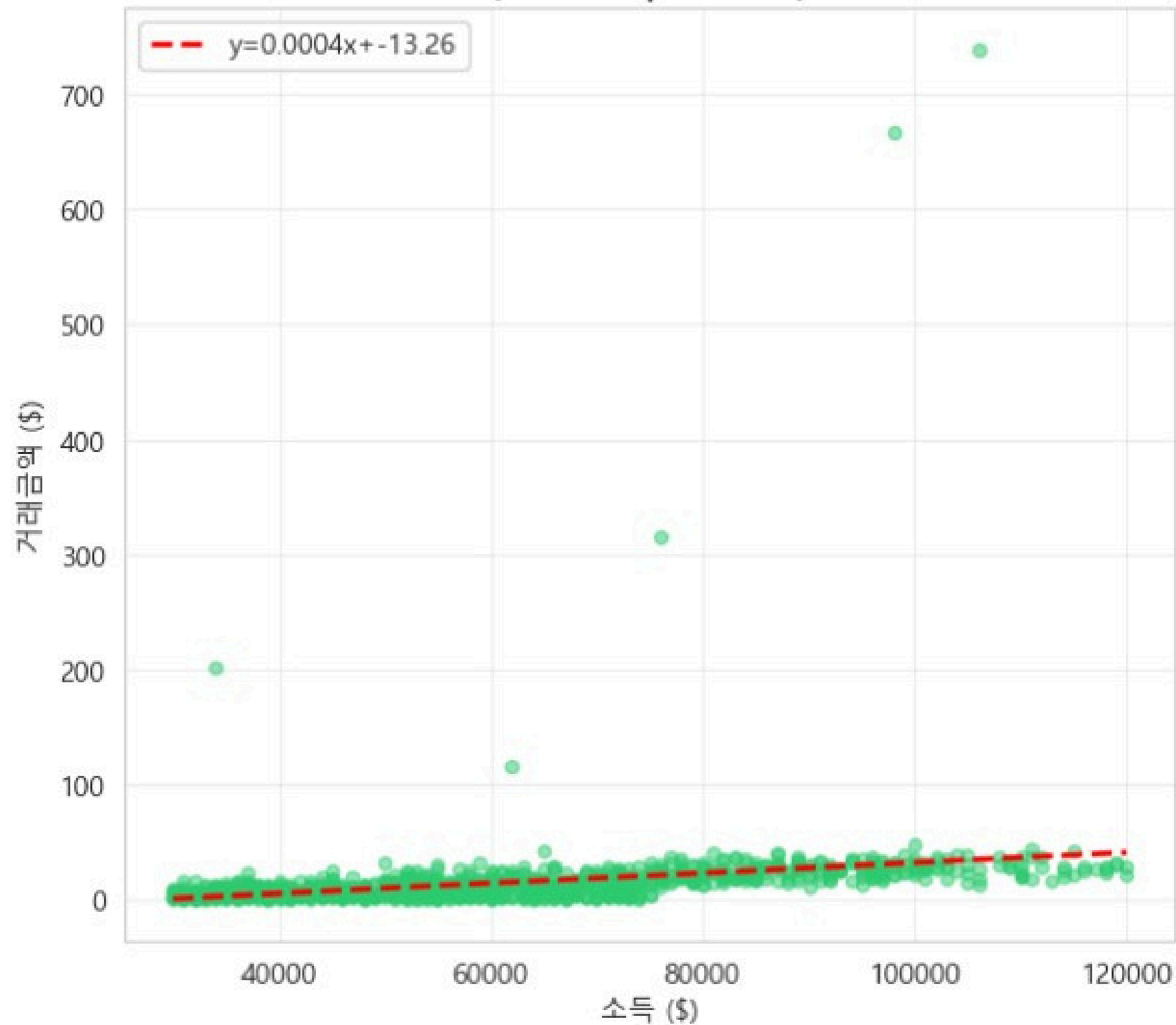
- 상관계수 (r): 0.0915
- p-value: 0.0000

## Spearman 상관분석 결과:

- 표본 크기: 123802
- 상관계수 (ρ): 0.2614
- p-value: 0.0000
- 유의수준  $\alpha=0.05$  기준: 통계적으로 유의미함

# [소득 → 거래금액]

5. 소득 vs 거래금액  
( $r=0.237$ ,  $p=0.0000$ )



## Pearson 상관분석 결과:

- 표본 크기: 138953
- 상관계수 ( $r$ ): 0.237
- p-value: 0.0000
- 유의수준  $\alpha=0.05$  기준: 통계적으로 유의미함
- 약한 상관관계

# 인사이트 및 비즈니스 전략

---

# 인사이트 및 비즈니스 전략

인사이트	비즈니스 전략
<p><b>성별 관계:</b></p> <ul style="list-style-type: none"><li>여성 고객 평균 거래금액 \$5.88 더 높음</li><li>(p=0.0000, 통계적 유의미)</li></ul>	<p><b>성별 차이 활용:</b></p> <ul style="list-style-type: none"><li>디저트·콜드브루 세트 중심 오퍼</li><li>SNS·앱 내 감성형 캠페인(포토 리뷰드, 한정 굿즈) 확산</li></ul>
<p><b>나이 관계:</b></p> <ul style="list-style-type: none"><li>나이와 거래금액 간 유의미한 관련성 발견</li><li>연령대별 소비 성향 차이</li></ul>	<p><b>연령대별 세분화:</b></p> <ul style="list-style-type: none"><li>20~30대: 포인트 적립·한정 메뉴 퀘스트형 오퍼</li><li>40대 이상: 잦은 방문 리뷰드·모바일 오더 할인</li></ul>
<p><b>소득 관계:</b></p> <ul style="list-style-type: none"><li>소득 수준 높을수록 거래금액 증가</li><li>(r=0.237, p=0.0000)</li></ul>	<p><b>프리미엄 세그먼트 집중:</b></p> <ul style="list-style-type: none"><li>시그니처 음료·리저브 라인업 프로모션 강화</li><li>리저브 회원 전용 리뷰드</li></ul>



**감사합니다**

