

# E-commerce 전환율 최적화: 트래픽 소스와 고객 행동 분석을 통한 마케팅 최적화 전략

이인  
데이터분석 프로젝트  
2024.08.28

# 목차

- 
- 1 프로젝트 개요
  - 2 데이터 분석
  - 3 결론

# 01. 프로젝트 개요



### 주제 선정 배경

쇼핑몰 고객의 웹사이트  
방문 경로가 다양함



특정 트래픽 소스에서  
구매율이 높거나 낮음을 보임



어떤 요인들이 고객의 구매 전환에  
영향을 미치는 걸까?



어떻게 하면 최적의 전환을  
유도할 수 있을까?



### 분석 목표

웹사이트 트래픽과 고객 행동 패턴을  
다각도로 분석하여  
**채널 최적화와 전환율 향상을 위한  
전략을 도출해보자!**



### 분석 도구



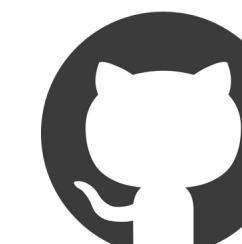
언어 및 DB:  
**python/mysql**



Tool:  
**Jupyter notebook/  
mysql workbench**



시각화:  
**matplotlib, seaborn**



작업 공유:  
**GitHub**



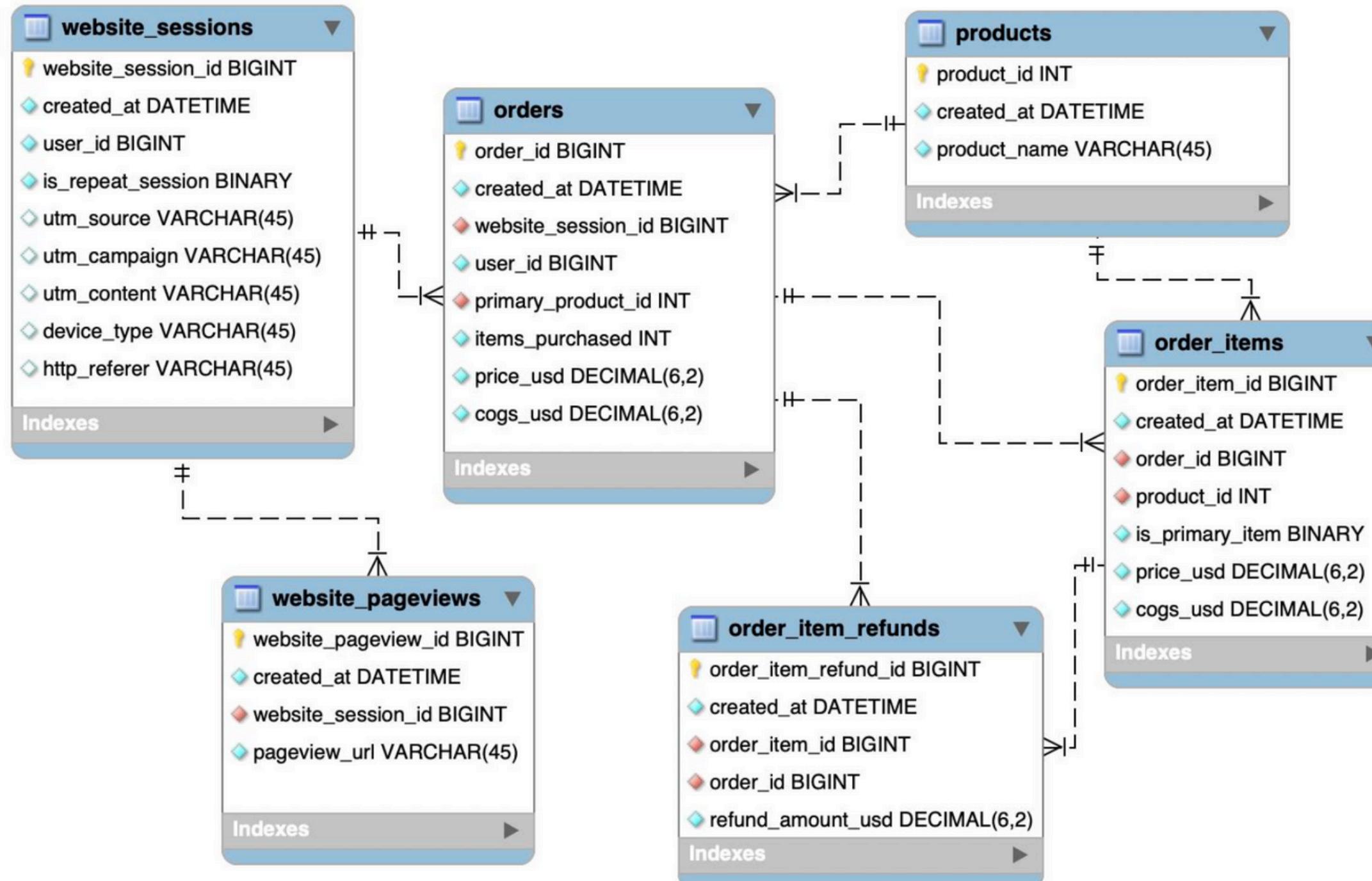
## 수행 절차



수행 기간	활동
08.07 – 08.09 (3일)	<ul style="list-style-type: none"><li>• 프로젝트 주제 및 목표 설정</li><li>• 데이터 범위 및 요구사항 정의</li></ul>
08.10 – 08.12 (3일)	<ul style="list-style-type: none"><li>• 데이터 수집</li><li>• 데이터 정제 및 통합</li></ul>
08.13 – 08.17 (5일)	<ul style="list-style-type: none"><li>• 기초 통계 분석(mysql)</li><li>• 데이터 시각화(python)</li><li>• 이상 값 및 패턴 탐색</li></ul>
08.18 – 08.22 (5일)	<ul style="list-style-type: none"><li>• 트래픽 세분화</li><li>• 전환 퍼널 및 A/B 테스트 분석</li></ul>
08.23 – 08.28 (6일)	<ul style="list-style-type: none"><li>• 주요 인사이트 정리</li><li>• 비즈니스 전략 제안</li><li>• 시각화 및 ppt 보고서 작성</li></ul>



## 데이터 설명



**데이터셋:** e-commerce 웹사이트의 트래픽 및 고객 행동 데이터

**데이터 출처:** udemy 온라인 교육 사이트의 [Advanced SQL: MySQL for Ecommerce Data Analysis](#) 교육자료

**활용 목표:** MySQL 사용하여 고객 접근 및 상호작용 분석, 랜딩 페이지 성능 및 전환율 분석, 고객 행동 분석



### 활용 데이터 테이블 및 분석 주제

#### 트래픽 소스 분석

- **데이터:** website\_sessions의 UTM 정보 및 orders
- **분석:** 트래픽 소스별 주문 건수 및 전환율 분석에 활용.

#### 랜딩 페이지 성과 분석

- **데이터:** website\_sessions, website\_pageviews 및 orders
- **분석:** 특정 랜딩 페이지의 방문 수와 전환율 분석에 활용.

#### 퍼널 테스트 분석

- **데이터:** website\_sessions, website\_pageviews의 pageview\_url 및 orders
- **분석:** 전환 퍼널의 각 단계에서 이탈률 분석을 통한 주요 이탈 지점 식별에 활용.

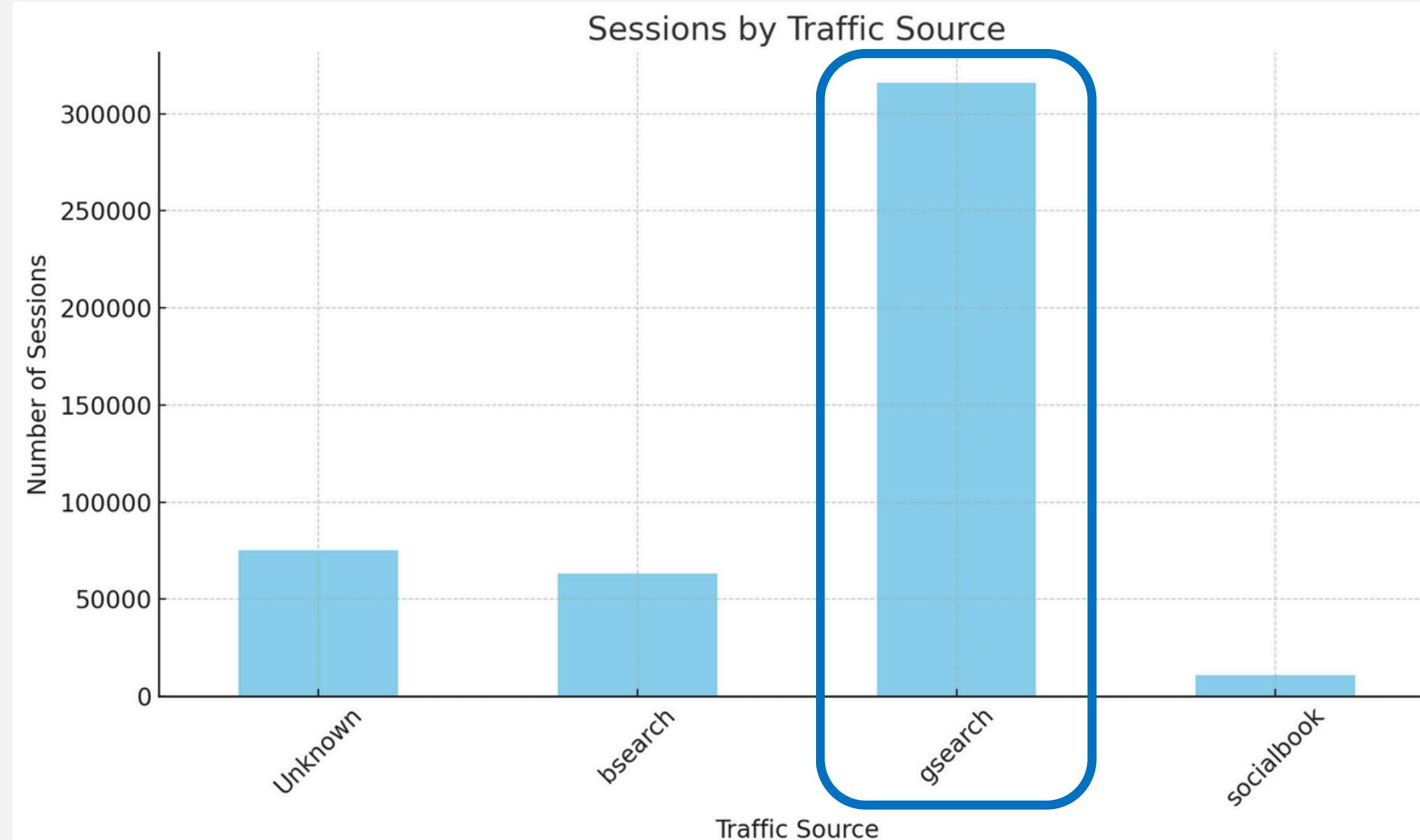
#### 고객 행동 분석

- **데이터:** website\_sessions의 is\_repeat\_session과 user\_id
- **분석:** 반복 방문 고객의 행동을 파악해 구매율 분석에 활용.

## 02. 데이터 분석



## 트래픽 별 세션수 분석



## 결과 테이블

Result Grid Filter Rows: Search Export:

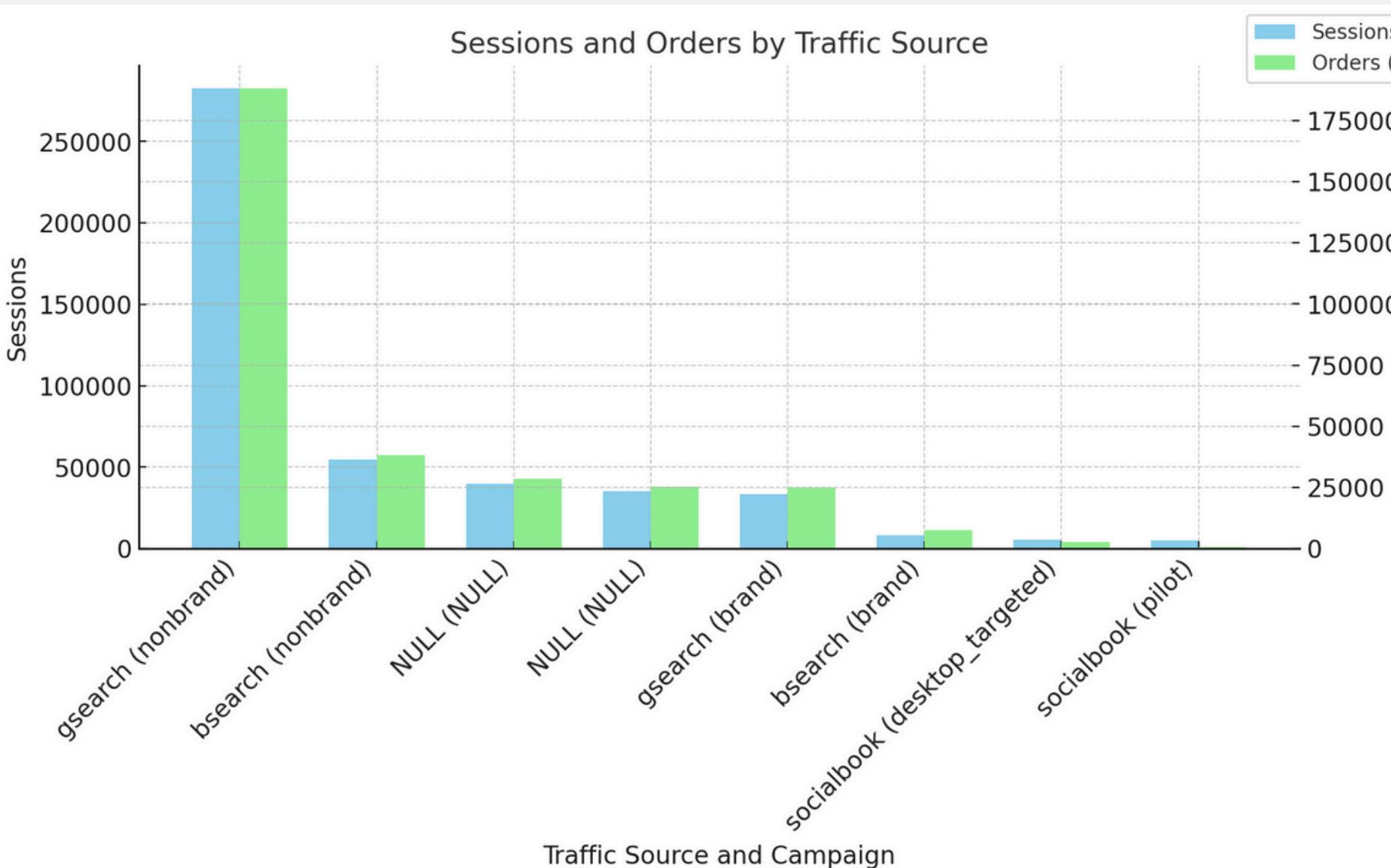
utm_source	utm_campaign	http_referer	sessions
gsearch	nonbrand	https://www.gsearch.com	282706
bsearch	nonbrand	https://www.bsearch.com	54909
NULL	NULL	NULL	39917
NULL	NULL	https://www.gsearch.com	35202
gsearch	brand	https://www.gsearch.com	33329
NULL	NULL	https://www.bsearch.com	8209
bsearch	brand	https://www.bsearch.com	7914
socialbook	desktop_targeted	https://www.socialbook.com	5590
socialbook	pilot	https://www.socialbook.com	5095

## 결과 요약

- **gsearch**는 **nonbrand** 캠페인을 통해 가장 많은 세션을 유도하고 있음.
- **bsearch**는 또한 의미있는 세션을 유도하고 있으나 **gsearch**에 비해 성과가 낮음.
- **socialbook**은 **desktop\_targeted** 캠페인이 좋은 성과를 보이고 있음.
- **NULL(unknown)** 값이 많이 포함되어 있어 직접 방문이나 트래킹이 안된 트래픽에 대한 추가 분석이 필요.



## 트래픽 별 세션수와 주문수 비교



## 결과 테이블

utm_source	utm_campaign	http_referer	sessions	orders
gsearch	nonbrand	https://www.gsearch.com	282706	18822
bsearch	nonbrand	https://www.bsearch.com	54909	3818
NULL	NULL	NULL	39917	2856
NULL	NULL	https://www.gsearch.com	35202	2519
gsearch	brand	https://www.gsearch.com	33329	2511
NULL	NULL	https://www.bsearch.com	8209	743
bsearch	brand	https://www.bsearch.com	7914	701
socialbook	desktop_targeted	https://www.socialbook.com	5590	288
socialbook	pilot	https://www.socialbook.com	5095	55

## 결과 요약

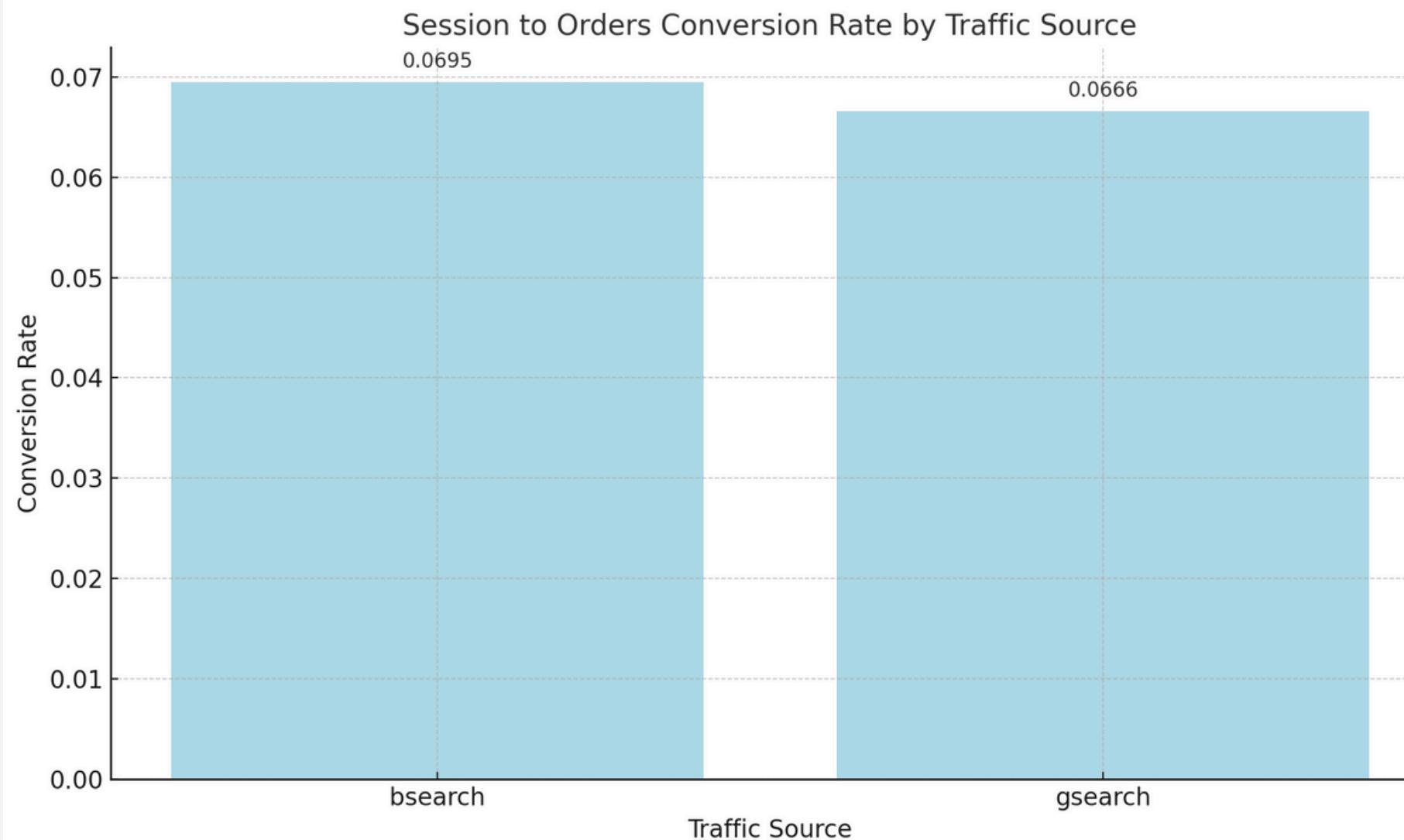
- **gsearch의 nonbrand 캠페인은 가장 높은 성과를 보임.**
- **bsearch는 gsearch보다 낮은 전환율을 보임.**
- **socialbook의 pilot 캠페인은 성과가 가장 낮음.**



gsearch와 bsearch의 nonbrand 캠페인 트래픽을 더 상세히 분석 비교해보자



## Gsearch vs. Bsearch 전환율 비교



## 결과 테이블

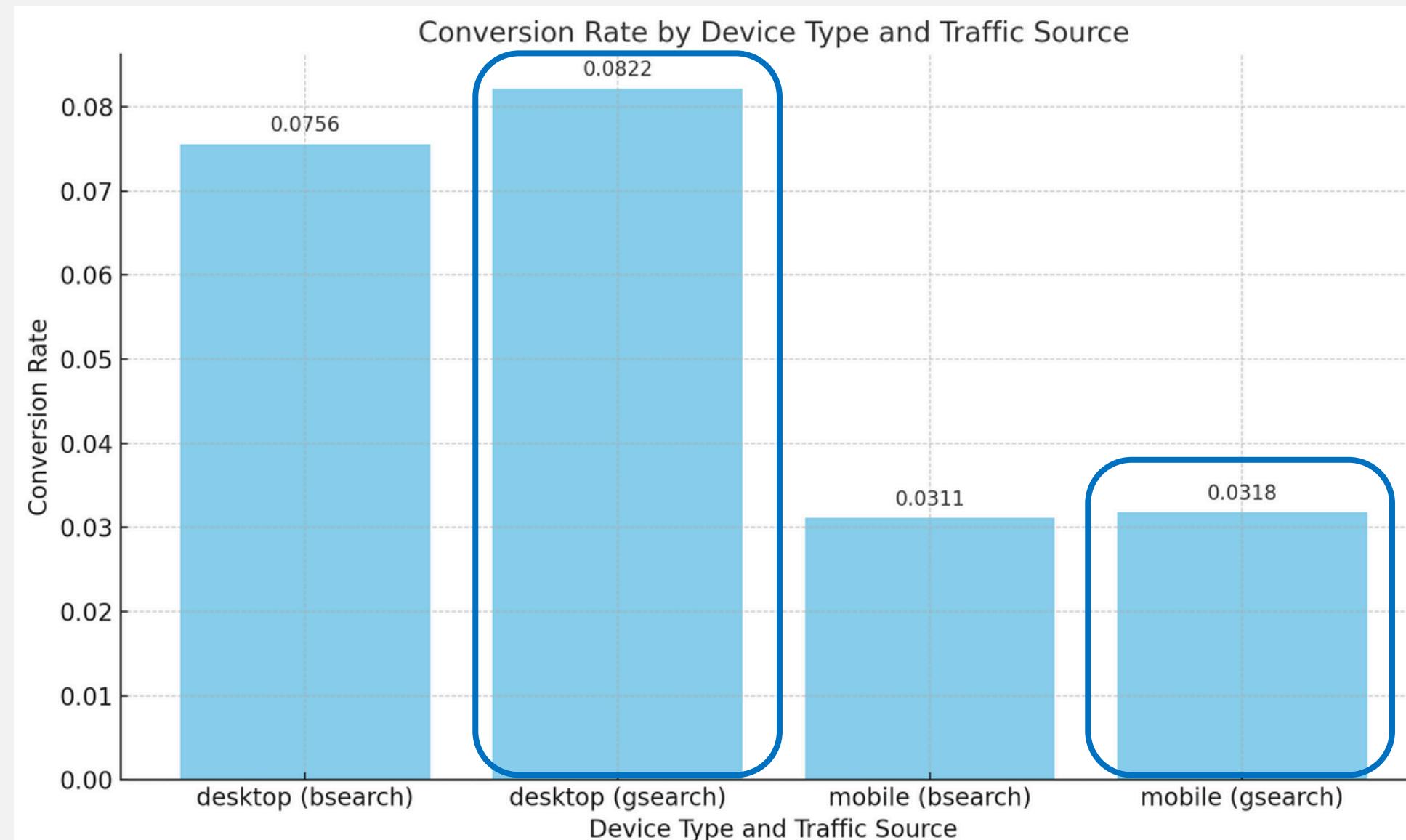
utm_source	utm_campaign	sessions			orders	session_to_orders_conv.
		nonbrand	54909	3818		
bsearch	nonbrand				282706	18822
gsearch	nonbrand					0.0666

## 결과 요약

- **bsearch**는 상대적으로 적은 세션(54,909)을 기록 하지만, 전환율(6.95%)이 더 높음.
- **gsearch**는 훨씬 많은 트래픽(282,706)을 기록하지만, 전환율(6.66%)이 상대적으로 약간 더 낮음.



## 7|7| 유형별(Device Type) 전환율 비교



## 결과 테이블

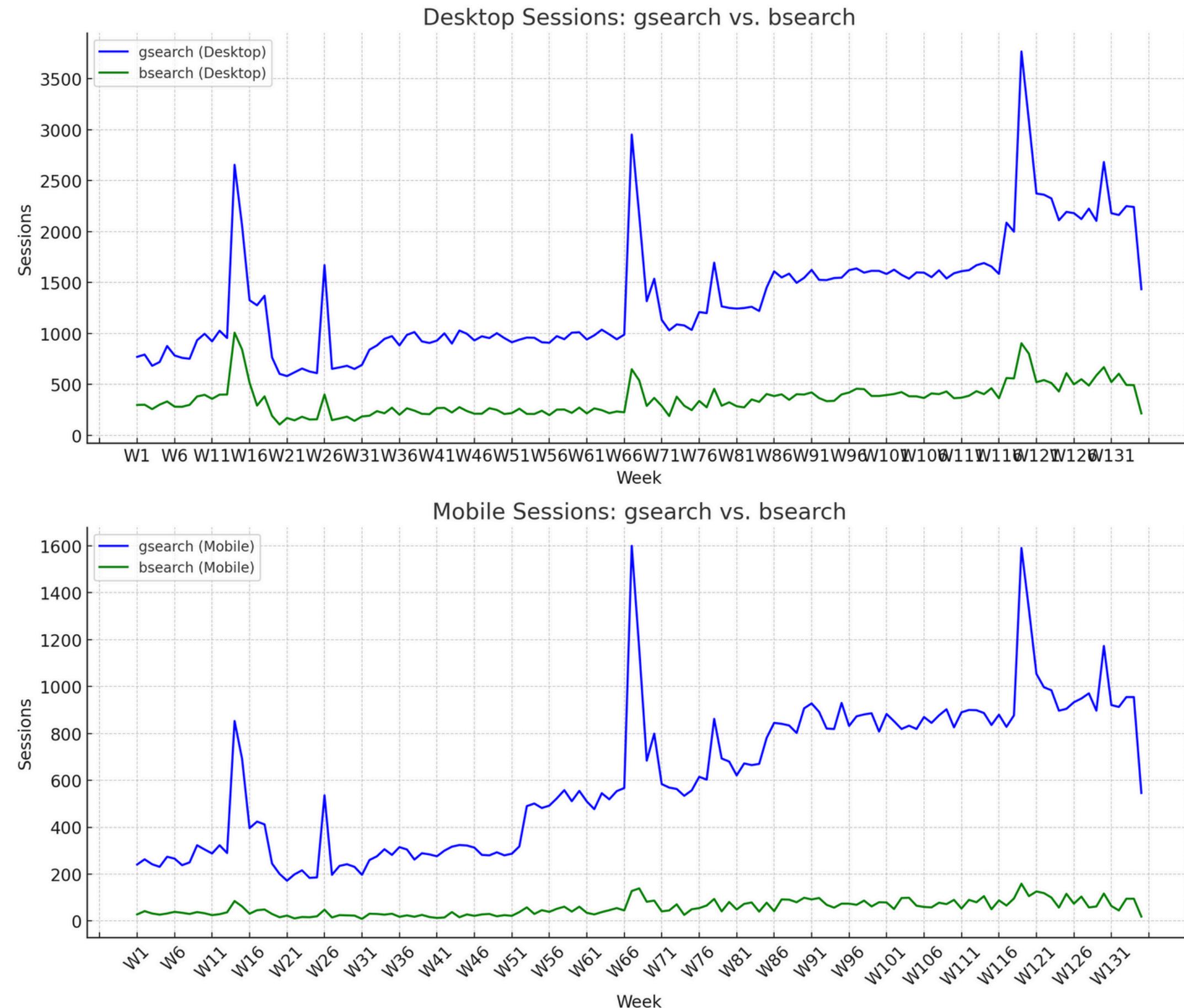
device_type	utm_source	sessions	orders	conv_rate
desktop	bsearch	47395	3584	0.0756
desktop	gsearch	195155	16037	0.0822
mobile	bsearch	7514	234	0.0311
mobile	gsearch	87551	2785	0.0318

## 결과 요약

- **데스크탑**에서 gsearch(8.22%)와 bsearch(7.56%)는 높은 전환율을 보임.
- **모바일**에서는 gsearch(3.18%)와 bsearch(3.11%) 모두 전환율이 낮은 편.
- **gsearch**는 모두 기기에서 **bsearch**보다 전반적으로 더 나은 성과를 보이고 있음.

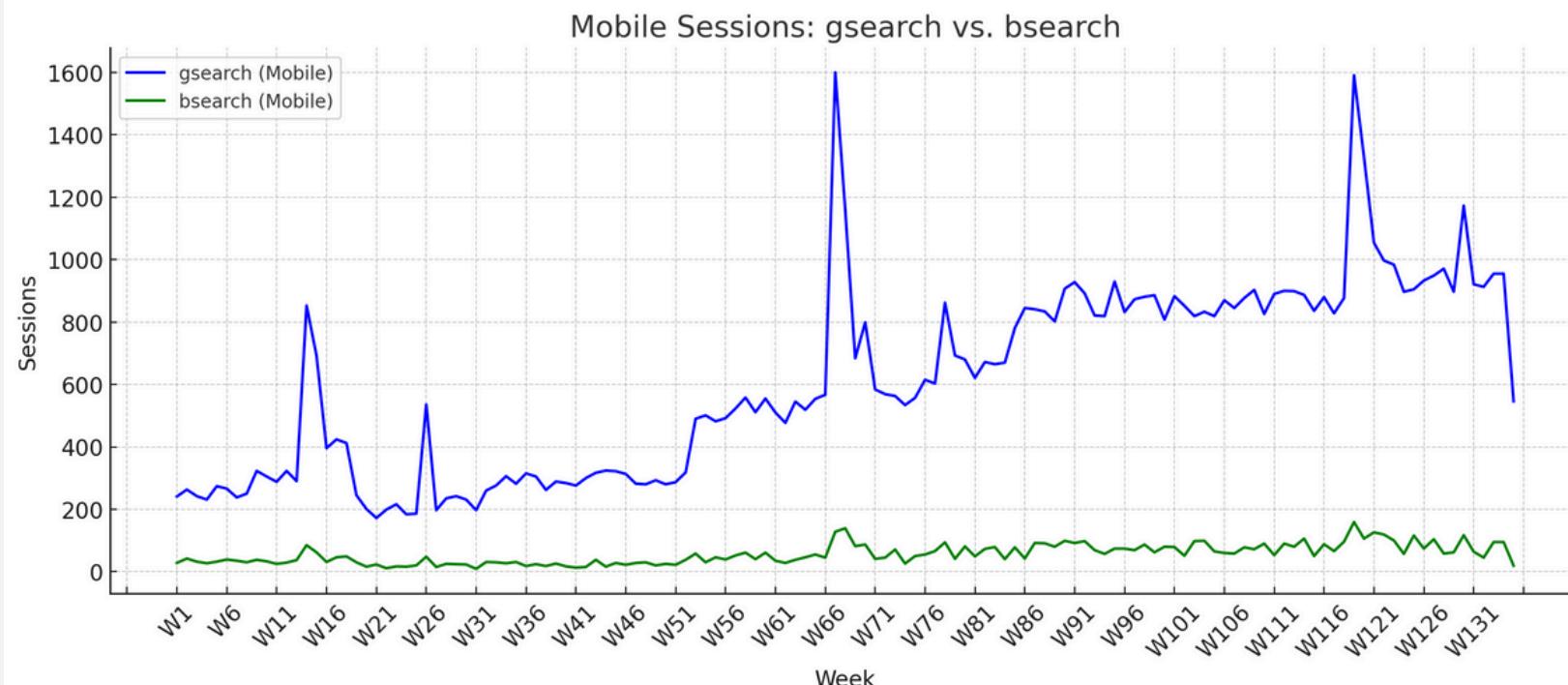
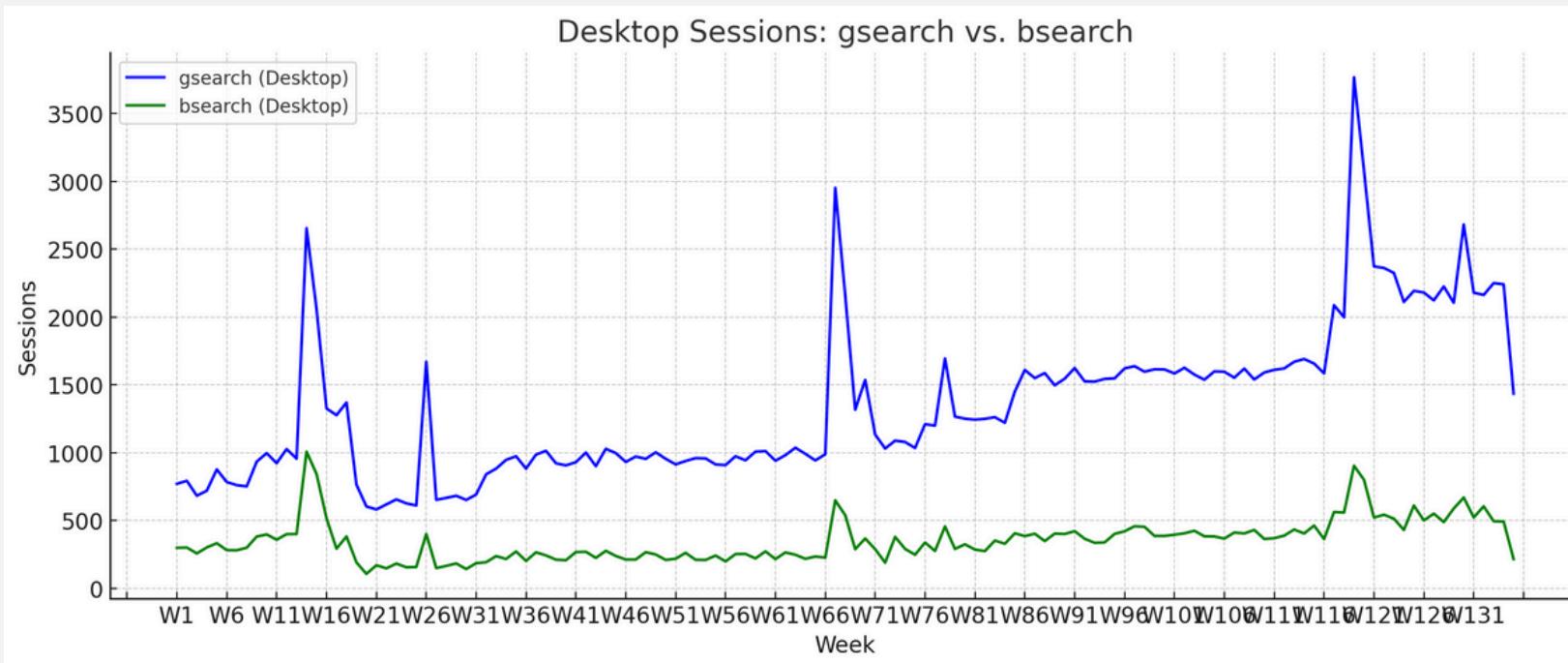


## 7|7| 유형별 세션수(주간 단위) 비교





## 7/17 기 유형별 세션수(주간 단위) 비교

**Gsearch****Desktop**

세션 수

&gt;

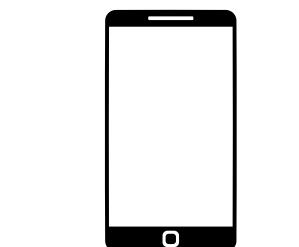
변동성

&gt;

**Bsearch**

세션 수

변동성

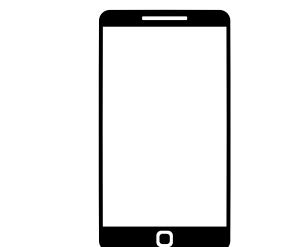
**Mobile**

세션 수

&gt;

변동성

&gt;



세션 수

변동성

#### 트래픽 유형별 정의

특정 브랜드 광고 캠페인  
을 통해 유입된 방문자

Brand

직접적으로 웹사이트 주소  
를 입력해 유입된 방문자

Direct

브랜드와 관계없는 일반  
유료 광고 캠페인을  
통해 유입된 방문자

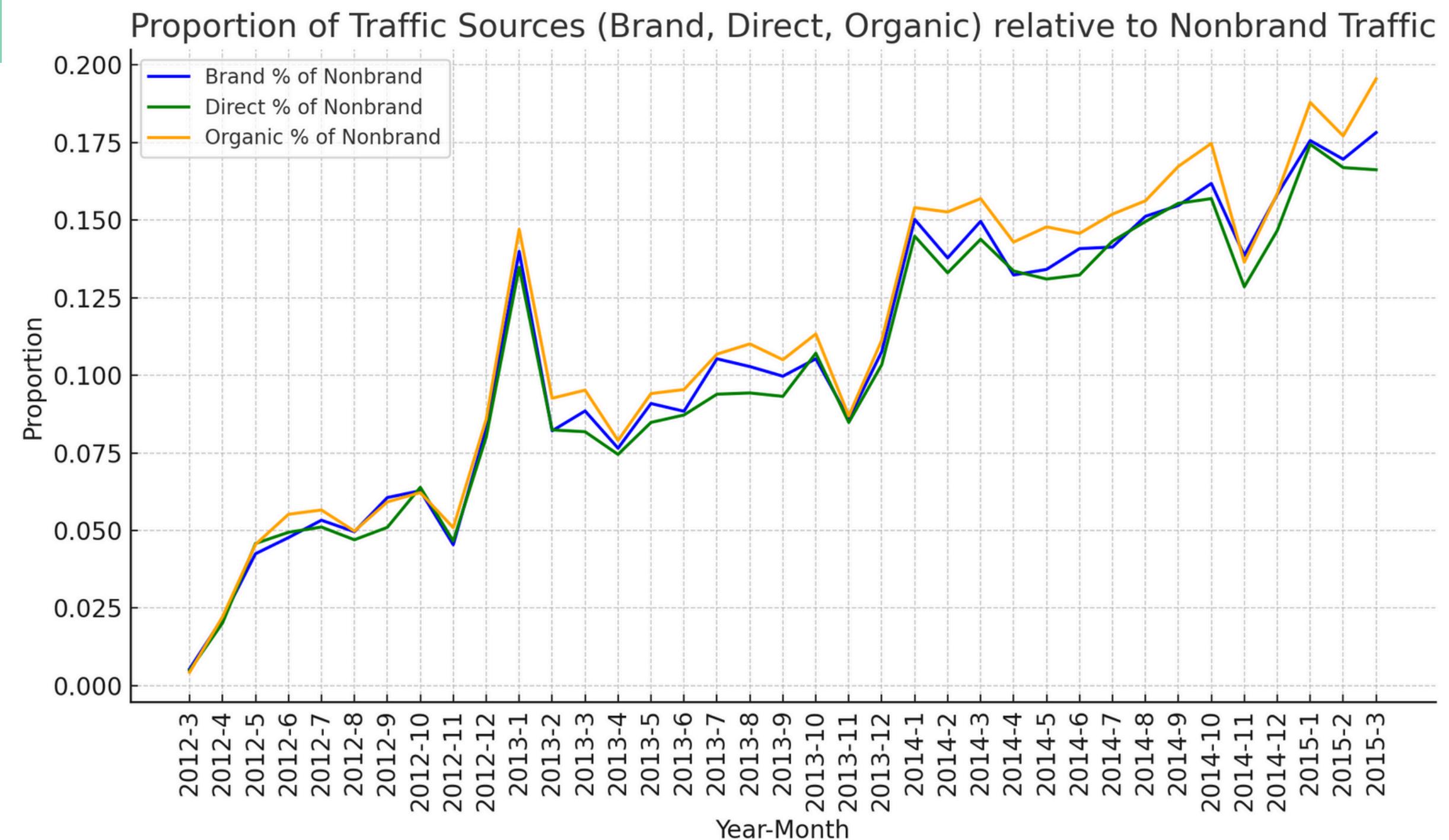
Nonbrand

유료 광고 없이 자연 검색  
결과를 통해 유입된 방문자

Organic

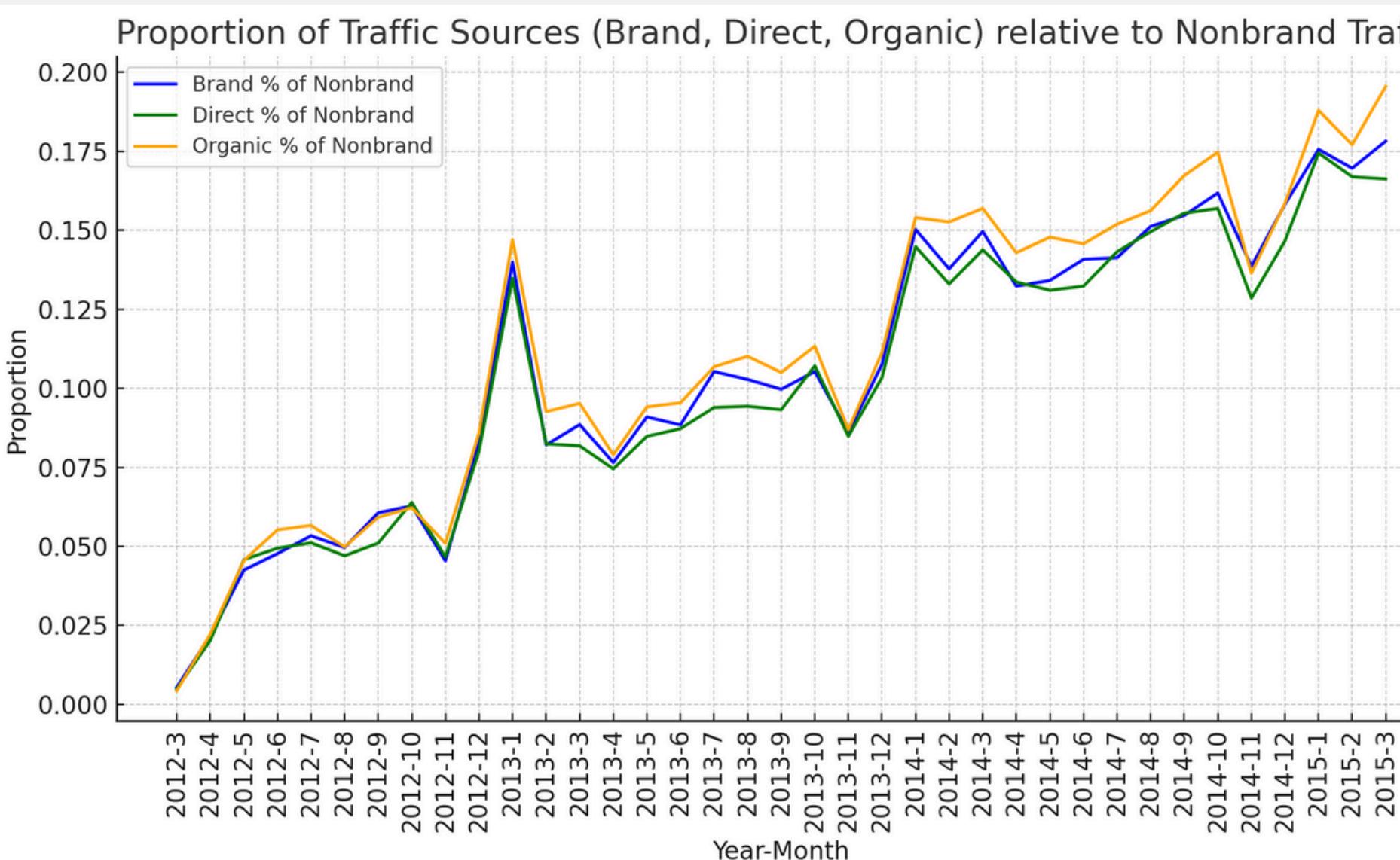


## Nonbrand 세션 수 대비 월별 (weekly) 변화량





## Nonbrand 세션 수 대비 월별 (weekly) 변화량



## 결과 요약

### 브랜드 트래픽 (paid\_brand):

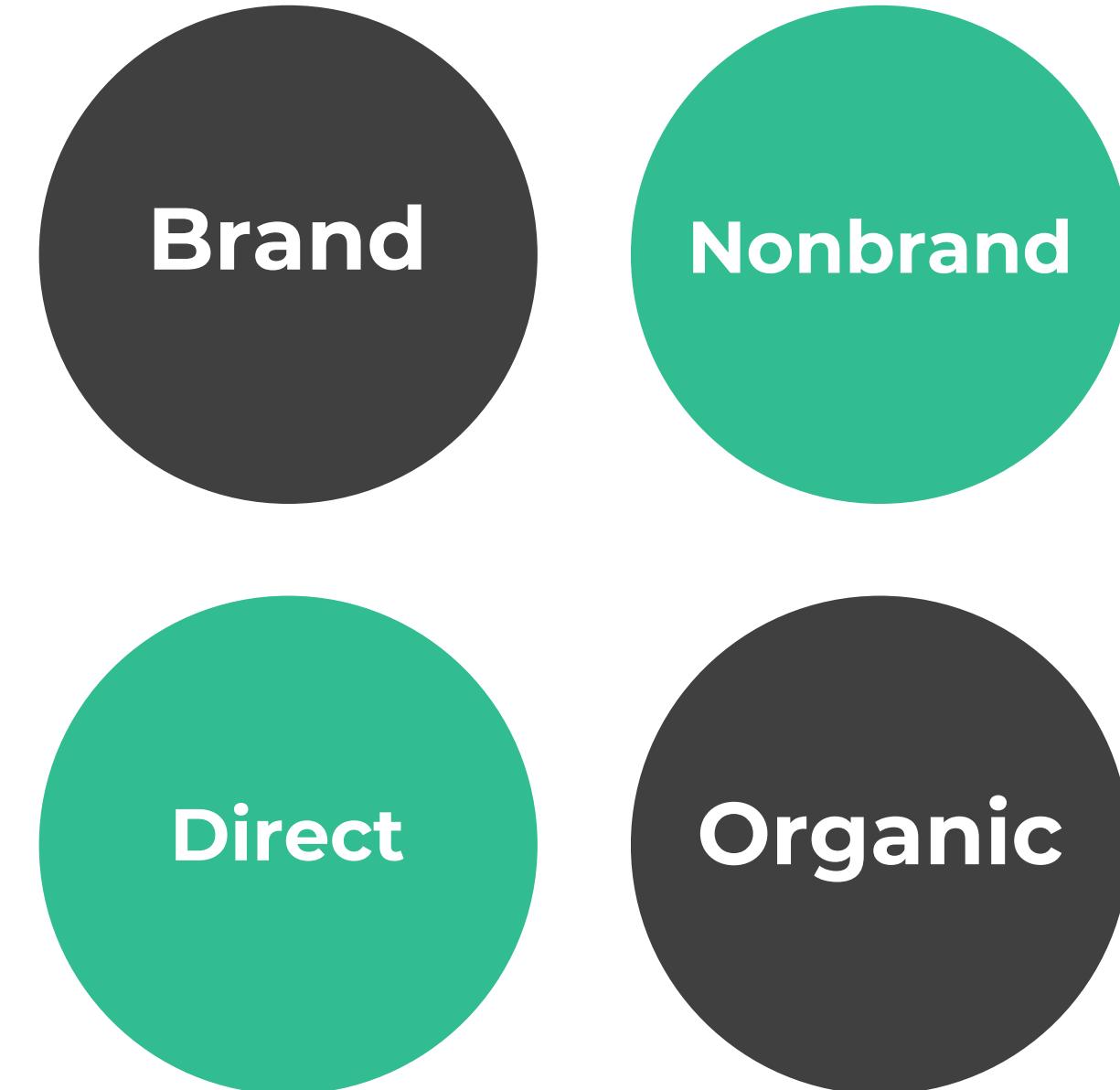
- 시간이 지남에 따라 점진적으로 증가하고 있음.
- 2014년 말에서 2015년 초 사이에 급격히 증가 추세
- nonbrand 트래픽 비율도 꾸준히 상승함.

### 다이렉트 트래픽 (direct type in):

- 변동성 적고 꾸준한 성장세를 보임.
- 비율은 다른 트래픽 소스에 비해 상대적으로 낮음.

### 오가닉 트래픽 (organic search):

- 변동성이 크며, 2013년 초에서 2014년 말까지 급격히 성장함.



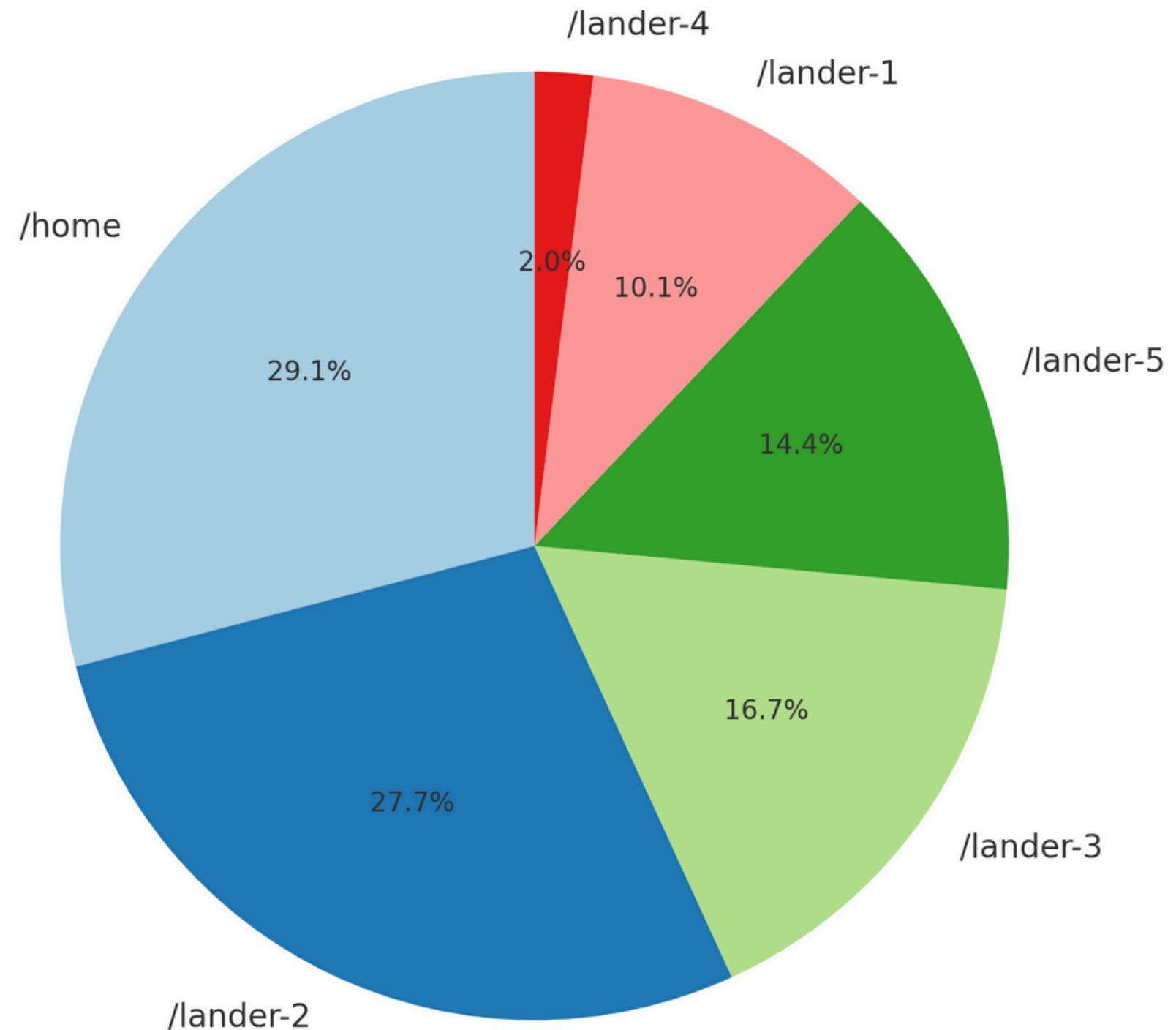
## 결론

- 브랜드 캠페인은 효과적으로 성장 중이며, 브랜드 인식이 확립되고 있음.
- 다이렉트 트래픽은 비교적 낮지만, 꾸준히 성장하며 브랜드 충성 고객을 유지 중임.
- 오가닉 트래픽은 변동이 있지만 성과가 향상 중이며, SEO 전략 강화를 통해 지속 성장이 가능함.



## 랜딩 페이지별 세션 비율

Landing Page Sessions Distribution





## 랜딩 페이지별 세션 비율

## 결과 테이블

landing_page	sessions_hitting_this_page
/home	137576
/lander-2	131170
/lander-3	79000
/lander-5	68166
/lander-1	47574
/lander-4	9385

## 결과 요약

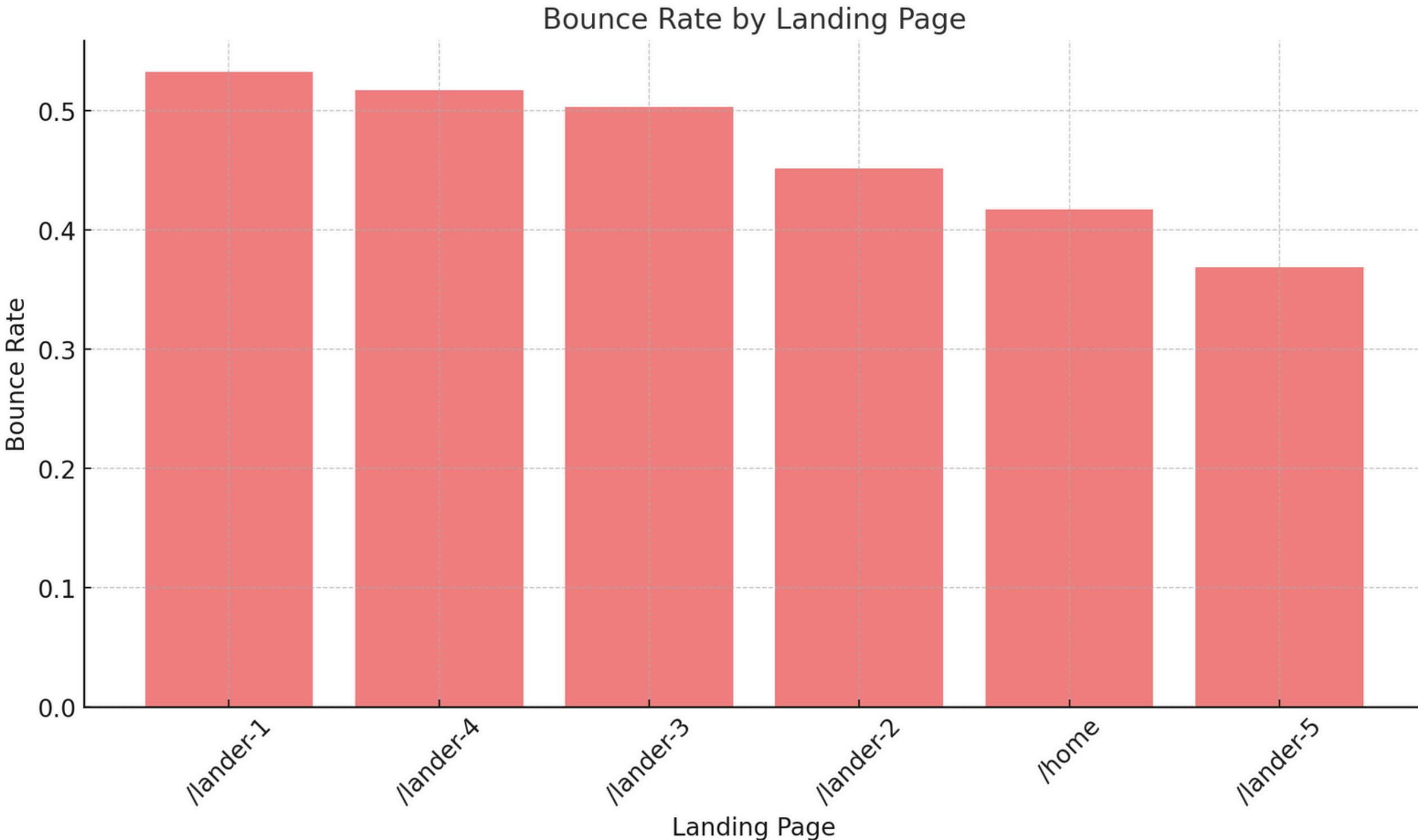
- **/home** 페이지는 전체 세션의 **29.1%**로, 가장 많은 사용자가 첫 방문 시 도달한 페이지임.
- **/lander-2** 페이지가 그 다음으로 많은 세션을 차지하며, **27.7%**의 비율을 보임.
- **/lander-3** 페이지는 **16.7%**, **/lander-5**는 **14.4%**의 비율로 상대적으로 적은 세션을 유도함.
- **/lander-1**과 **/lander-4**는 각각 **10.1%**와 **2.0%**의 세션 비율로, 가장 적은 방문자 수를 기록함.



랜딩 페이지별 이탈률(Bounce rate)을 분석하여 전환에 효과적인 랜딩페이지와 개선이 필요한 페이지 탐색



## 이탈률(Bounce Rate) 분석



- 이탈률이 높다는 것은 콘텐츠나 첫 방문 시 제공되는 정보가 부족하거나 탐색이 어려울 수 있음.
- 분석을 통해 이탈률이 높은 페이지의 개선 전략을 수립할 수 있고 이탈률이 비교적 낮은 페이지를 메인으로 했을 때 사용자가 실제 주문까지 이어질 확률이 높음.



## 이탈률(Bounce Rate) 분석

## 결과 테이블

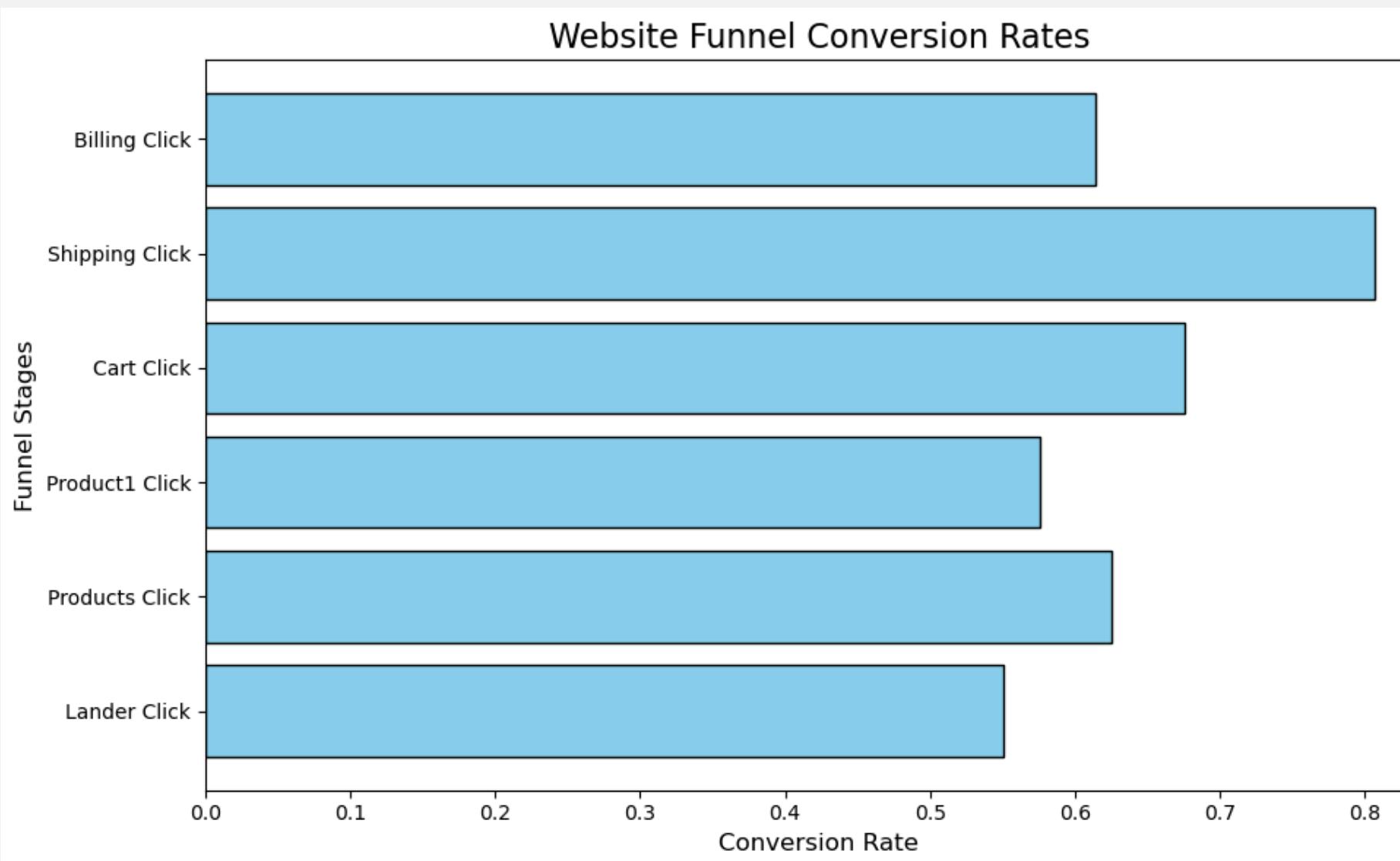
landing_page	sessions	bounced_sessions	bounce_rate
/lander-1	47574	25330	0.5324
/lander-4	9385	4851	0.5169
/lander-3	79000	39733	0.5029
/lander-2	131170	59249	0.4517
/home	137576	57346	0.4168
/lander-5	68166	25131	0.3687

## 결과 요약

- **/lander-1** 는 가장 높은 이탈률을 기록하고 있음.  
/lander-1에 도착한 후 사이트를 바로 떠나는 경향이 큼.
- **/lander-4** 와 **lander-3**도 높은 이탈률(50% 이상)을 보이고 있음. 방문자가 사이트에서 머무르지 않고 이탈하는 경우가 많으므로 구조나 사용자 경험(UX) 개선이 필요할 수 있음.
- **/lander-2** 는 세번째로 낮은 이탈률을 보이고 있는 반면, 두번째로 많은 유입량을 보이고 있음. 많은 유입량을 전환으로 돌릴 기회가 있음.
- **/home** 은 두 번째로 낮은 이탈률을 보이고 있음. 이는 홈페이지의 구조나 콘텐츠가 사용자의 탐색에 효과적일 가능성이 있음.
- **/lander-5**는 가장 낮은 이탈률을 보이고 있음. 사용자 경험이 긍정적일 수 있음.



## 페이지별 전환율 분석



## 결과 테이블

lander_click_rt	products_click_rt	product1_click_rt
-----------------	-------------------	-------------------

0.5512	0.6252	0.5764
--------	--------	--------

cart_click_rt	shipping_click_rt	billing_click_rt
---------------	-------------------	------------------

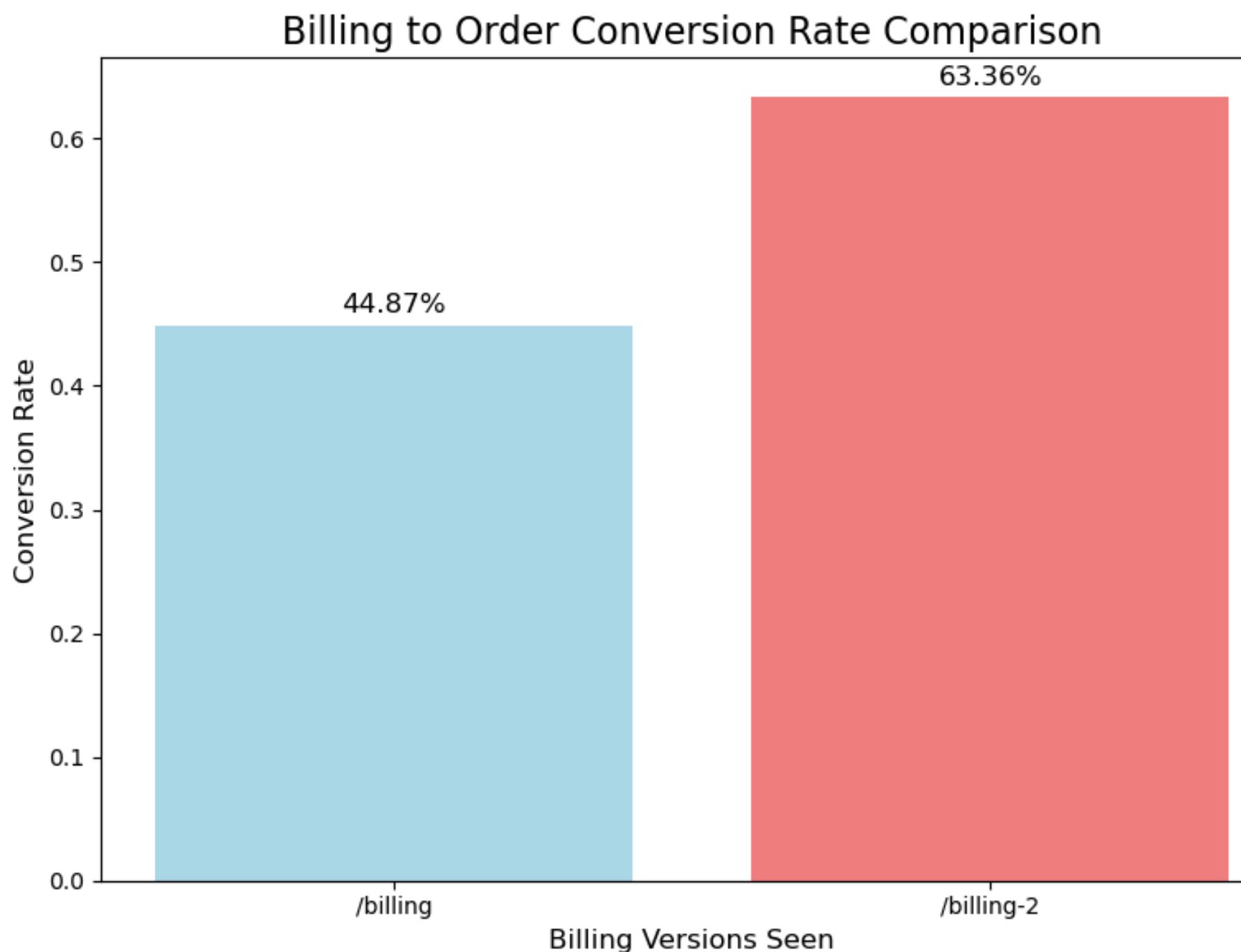
0.6761	0.8068	0.6146
--------	--------	--------

## 결과 요약

- 랜딩 페이지에서 제품 페이지로의 전환율이 약 **55%**로 상대적으로 낮음. 이는 첫 페이지에서 더 많은 제품 정보나 유인 요소가 필요할 수 있음을 시사함.
- 배송 페이지에서 결제 페이지로의 **전환율(80.68%)**이 가장 높음, 이는 사용자들이 이 과정에 대한 심리적 장벽이 크지 않다는 것을 보여줌.
- 결제 페이지에서 구매완료 페이지로의 **전환율(61.46%)**은 상대적으로 낮아 결제 과정에서 문제가 발생할 가능성을 시사함. 결제 과정의 간소화 또는 신뢰성 강화가 필요할 수 있음.



## 결제 전환율 A/B TEST



## 결과 테이블

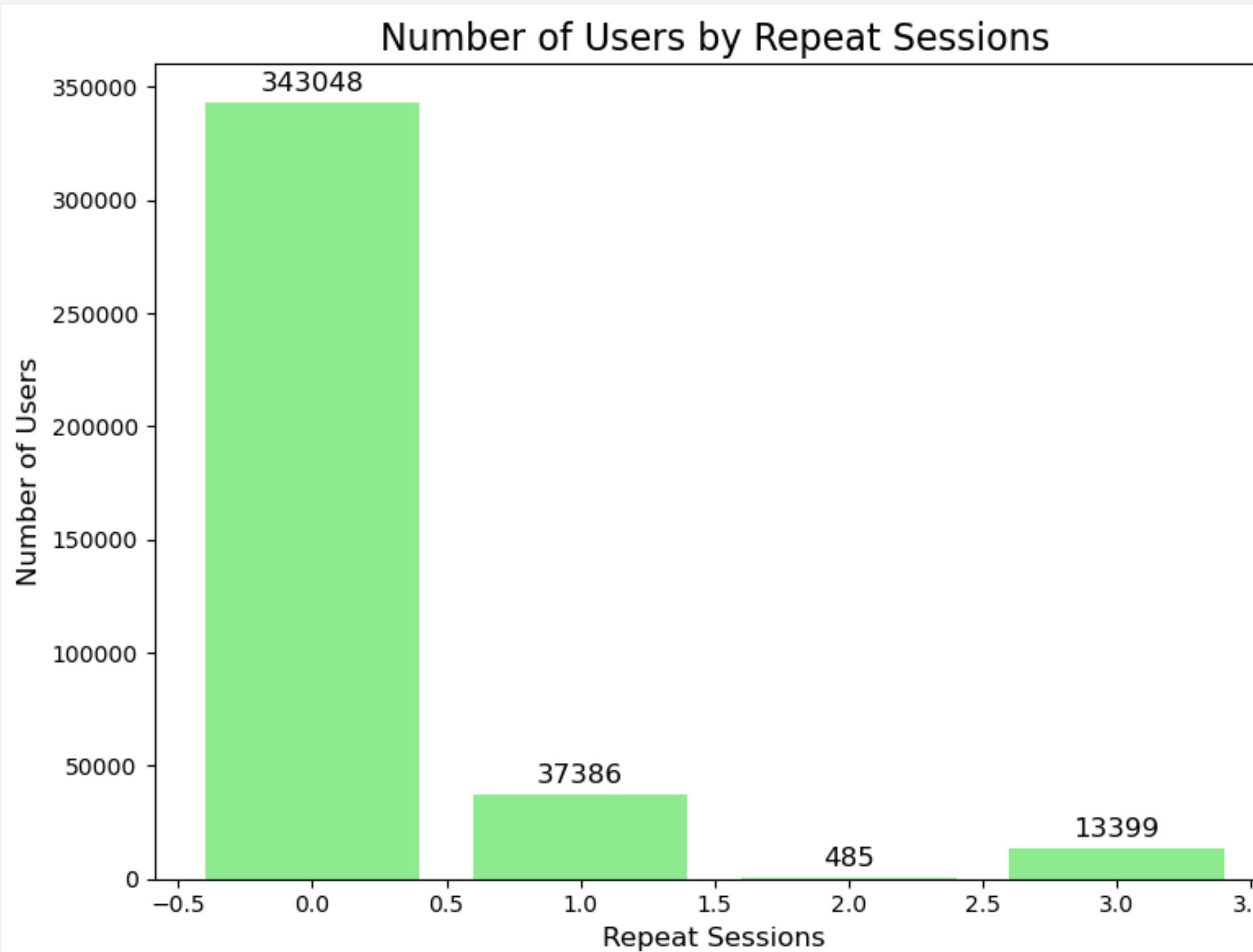
billing_versions_seen	sessions	orders	billing_to_order_rt
/billing	1627	730	0.4487
/billing-2	48411	30674	0.6336

## 결과 요약

- **/billing-2** 페이지는 **/billing** 페이지에 비해 더 높은 결제 전환율을 보이고 있음 (**63.36% vs 44.87%**).
- **/billing-2**는 더 많은 사용자가 결제를 완료하는 경향을 보임. 이는 **/billing-2**의 구조나 사용성이 더 나은 결제 경험을 제공할 수 있음을 시사함.
- **/billing** 페이지는 세션당 전환율이 낮음. **/billing** 페이지에서 결제 전환율이 낮은 이유를 파악하고, 페이지 구조나 결제 프로세스를 개선할 필요가 있음.



## 반복 방문 고객 세션 수



## 결과 테이블

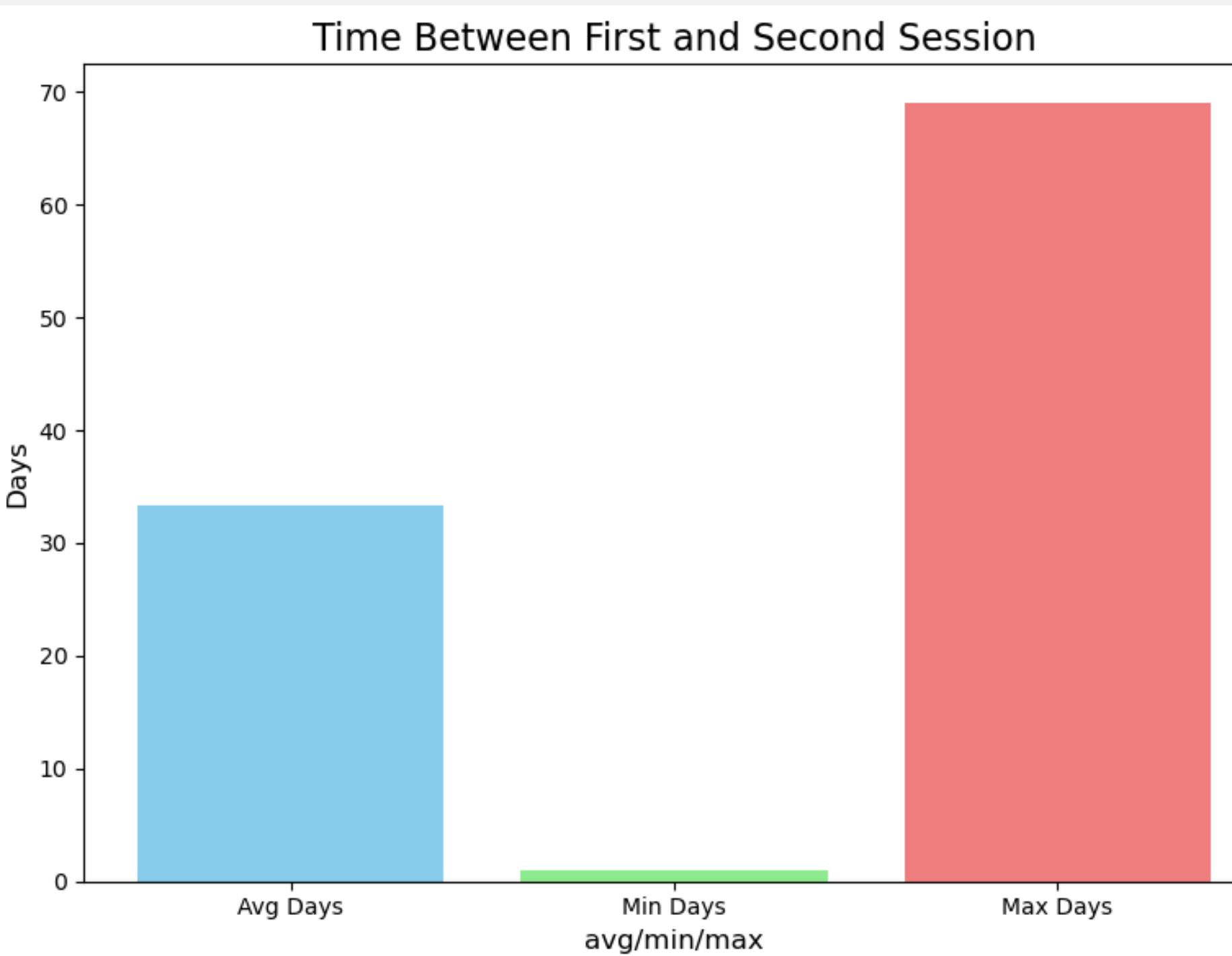
repeat_sessions	users
0	343,048
1	37,386
2	485
3	13,399

## 결과 요약

- 대다수의 사용자(**343,048명**)는 한 번만 사이트를 방문하고 다시 방문하지 않음.
- **1회 반복 방문한** 사용자의 수는 37,386명으로 상대적으로 많은 비율을 차지함.
- **2번 반복 방문한** 사용자는 485명으로 **매우 적음**.
- **3번 반복 세션을 가진** 사용자는 13,399명으로 **2번 반복 방문자보다 훨씬 많음**.



## 두번째 방문까지 소요일



## 결과 테이블

avg_days_first_to_second	min_days_first_to_second	max_days_first_to_second
33.2622	1	69

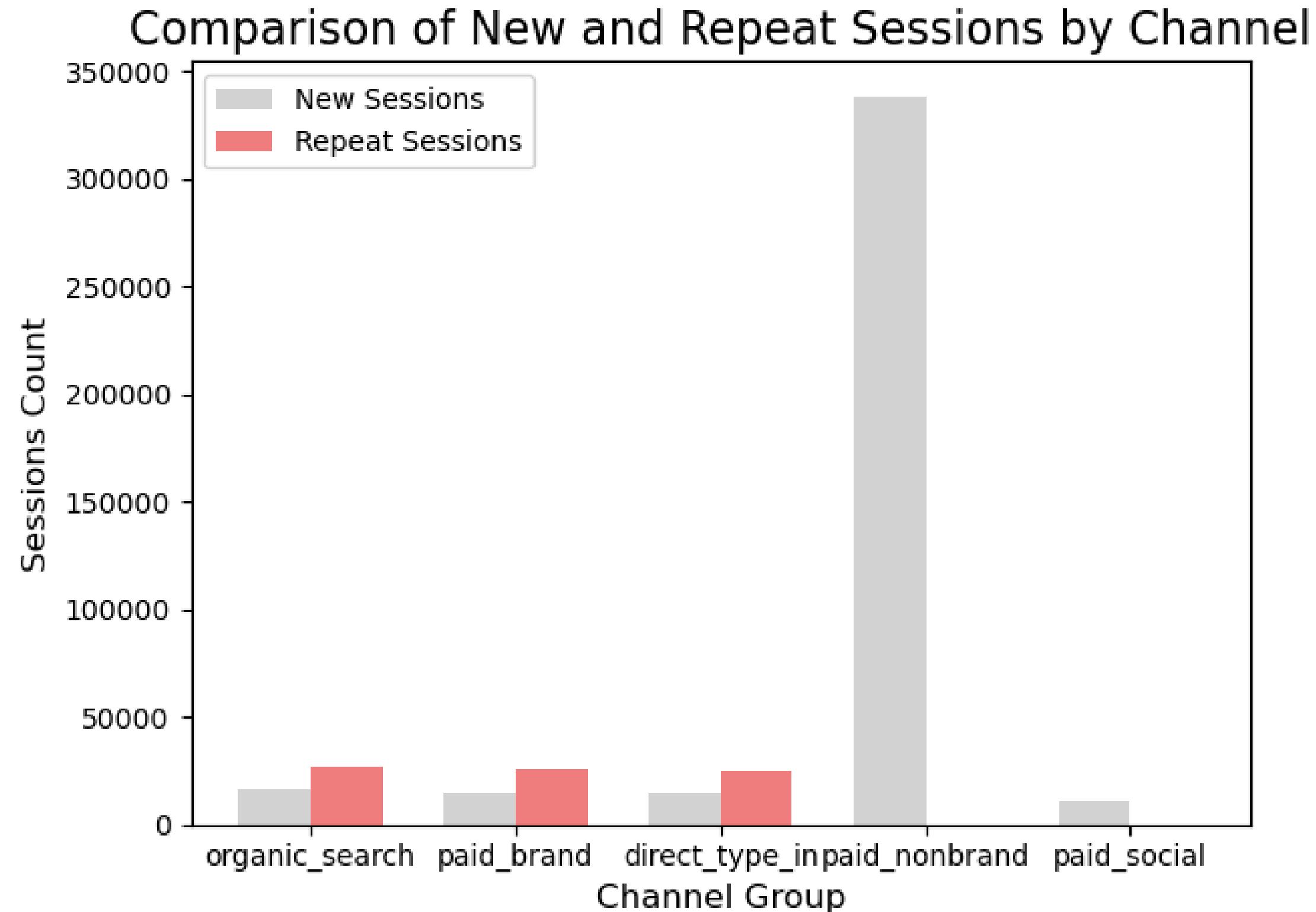
## 결과 요약

**대다수의 사용자는 평균 약 33일 후 재방문:**  
평균적으로 첫 방문 후 한 달 내에 사용자가 다시  
방문하는 것으로 나타남.

**최소 1일 이내에서 최대 69일 이후에 재방문한**  
사용자가 있음.



## 채널별 재방문 세션 수





## 채널별 재방문 세션 수

## 결과 테이블

channel_group	new_sessions	repeat_sessions
organic_search	16299	27112
paid_brand	15019	26224
direct_type_in	14700	25217
paid_nonbrand	337615	0
paid_social	10685	0

## 결과 요약

- **Organic Search:** 반복 방문율이 가장 높음.
- **Paid Brand:** 반복 방문율이 두번째로 높게 나타남.
- **Direct Type In:** 북마크나 주소 직접 입력을 통해 반복 방문이 세번째로 많음.
- **Paid Nonbrand:** 새로운 유입은 가장 많지만 반복 방문 없음.
- **Paid Social:** 소셜 광고로 유입된 사용자는 반복 방문하지 않음.

## 03. 결론



## 트래픽 소스 분석 요약

**gsearch**

- 가장 많은 세션을 유도
- 트래픽은 많지만 전환율(6.66%)은 약간 낮음.
- 데스크탑(8.22%)과 모바일(3.18%)에서 모두 높은 성과를 보여줌.

**bsearch**

- 트래픽 유입은 적지만 전환율(6.95%)이 **gsearch**보다 높음.
- 데스크탑(7.56%)에서 특히 성과가 좋음.

**Social book**

- 소셜 미디어 기반 트래픽으로, desktop\_targeted 캠페인이 성과를 보였음.
- pilot 캠페인의 성과가 낮아, 개선 전략이 요구됨.

**Direct & Null**

- **Direct Type In:** 브랜드 충성도가 높은 사용자들이 방문하며 꾸준한 성장세를 보임. 세션 수는 낮지만, 추가 성장 가능.
- **NULL 값이 있는 트래픽:** 트래킹 되지 않은 트래픽이 많아, 추가 분석 통해 마케팅 기회 탐색 필요.



## 트래픽 소스 별 최적화 전략



## 강점

가장 많은 세션을 유도하고 있어 트래픽 확보에 탁월한 성과를 보임.



세션 수는 적지만 전환율이 gsearch보다 높아 고효율적인 트래픽을 유도함.

## 기회

전환율이 상대적으로 낮으므로, 재방문을 유도하는 리마케팅 캠페인이나 랜딩 페이지 개선의 기회가 있음

## 전략

두 접속 기기에서 모두 성과가 좋은 만큼, 추가 마케팅 예산을 집중하고, 각 기기별 맞춤형 최적화 전략을 시행함.

트래픽 자체를 증가시켜, 현재 높은 전환율을 바탕으로 더욱 많은 전환을 유도할 수 있음.

데스크탑에서 성과가 좋은 만큼 데스크탑 사용자 대상 캠페인을 강화하고 모바일 사용자 대상의 최적화 전략도 병행함.

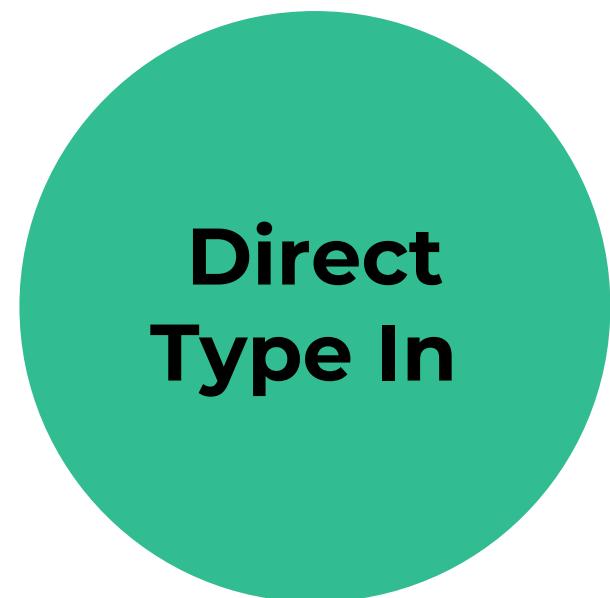


## 트래픽 소스 별 친절화 전략



## 강점

desktop\_targeted  
캠페인이 상대적으로 나은 성과를 보임.

Direct  
Type In

브랜드 충성도가 높은 고객층  
이 직접 웹사이트를 방문하며,  
지속적인 성장을 보여줌.

## 기회

pilot 캠페인의 전환율이 낮아, 프로모션 및 타겟팅을 통한 사용자의 관심을 끌어올릴 기회가 있음.

## 전략

소셜 미디어 기반의 트래픽 유도 전략을 재검토하고,  
참여율 상승을 위한 콘텐츠와 타겟팅 전략을 강화함.

세션 수 자체는 낮으므로,  
리텐션 마케팅을 강화해 충성 고객의 재방문을 유도할 수 있는 전략이 필요함.

이메일 마케팅, VIP 고객 프로그램 시행 등을 통해 이탈률을 줄이고, 충성 고객의 재방문을 장려할 수 있음.



## 랜딩 페이지 분석 요약

**Sessions %****29.1%****/home 페이지:**

- 전체 랜딩 세션의 29.1%를 차지하며 가장 많은 유입을 보임.
- 두 번째로 낮은 이탈률을 보이며, 방문자 유지에 효과적.

**Sessions %****27.7%****/lander-2 페이지:**

- 27.7%의 유입 비율로 두 번째로 높지만, 높은 이탈률을 보임.
- 이탈률 개선을 위한 UX 및 구조 개선 필요.

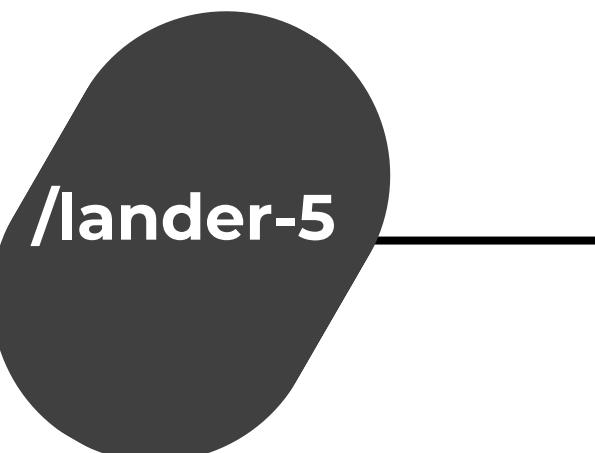
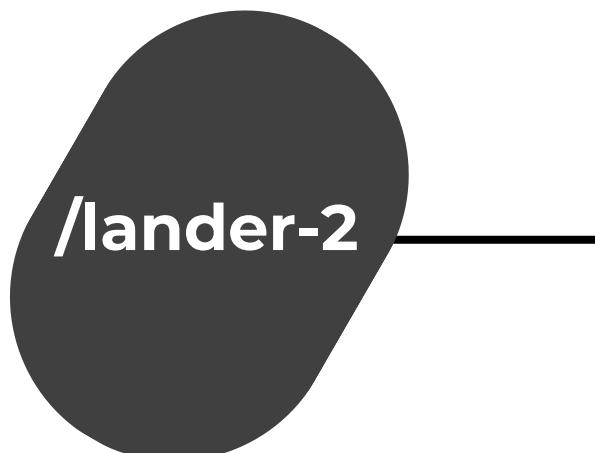
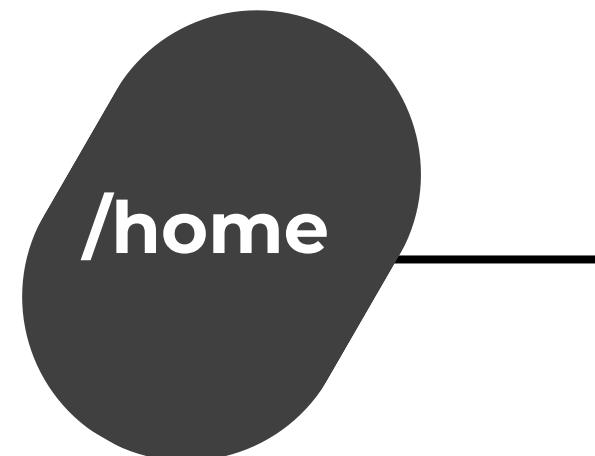
**bounce rate****36.9%****/lander-5 페이지:**

- 가장 낮은 이탈률을 기록, 긍정적인 사용자 경험을 제공할 가능성이 있음.



## 랜딩 페이지 최적화 전략

## 강점



## 기회

여전히 높은 이탈률을 낮추기 위한 방도 필요.

페이지 내에서 구매를 유도할 수 있는 콘텐츠와 구조 개선이 필요함.

페이지의 장점을 분석해 활용할 기회가 있음.

## 전략

사용자 피드백을 반영해 페이지 내 정보 구성 UI 최적화 및 신뢰감을 줄 수 있는 콘텐츠 추가를 통한 이탈률 감소 전략 시행

추천 제품 섹션을 추가하거나, 이벤트 및 프로모션 배너를 통해 구매로 이어질 수 있는 구매 유도 요소 강화

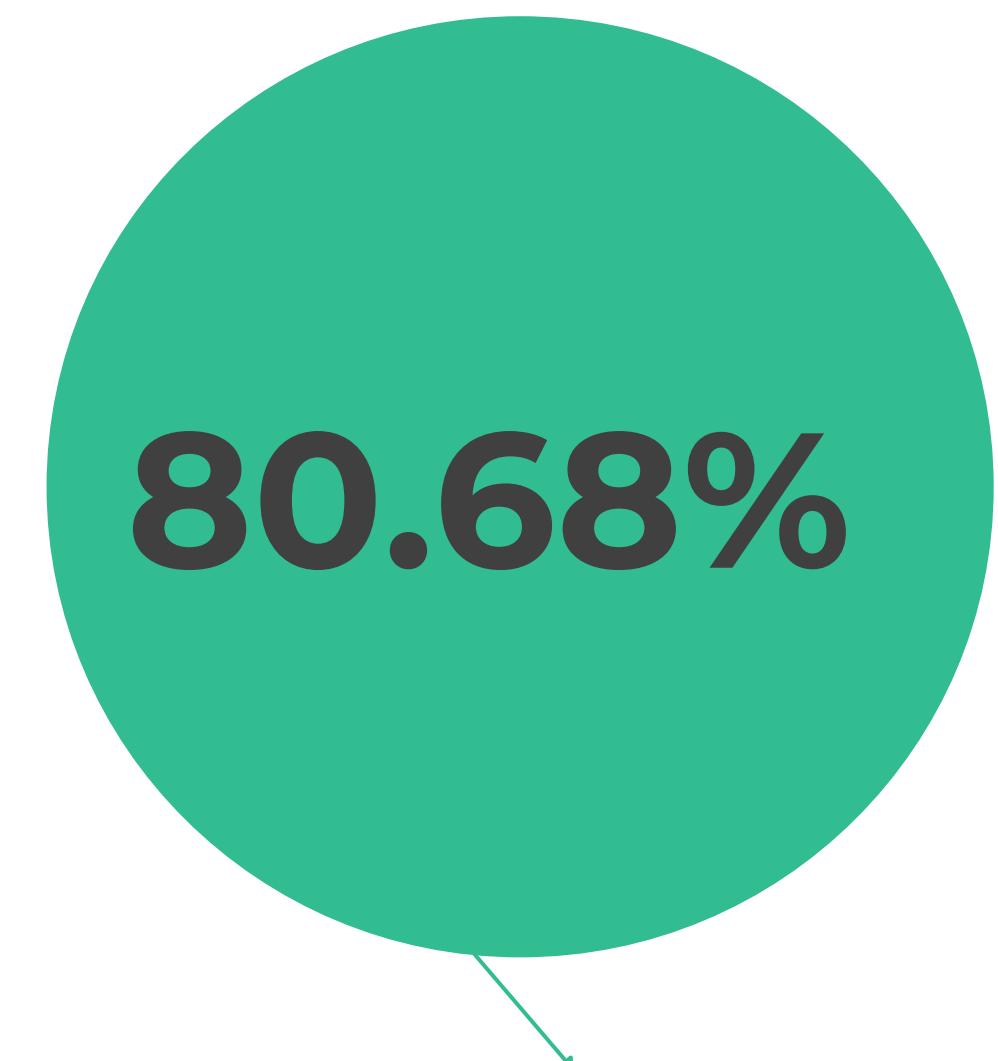
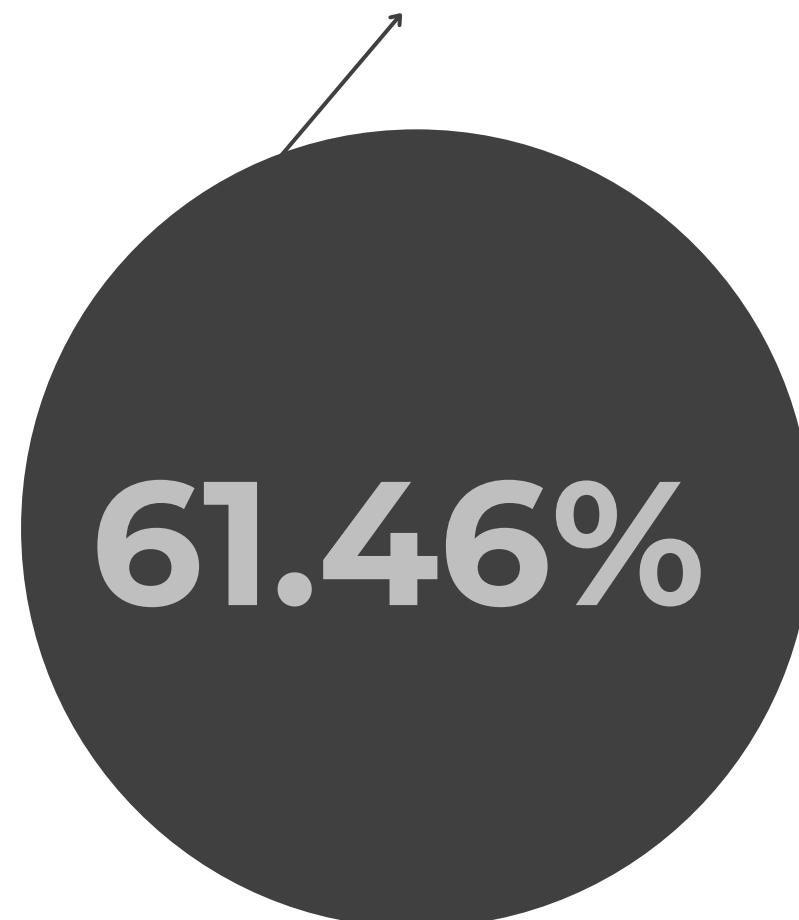
페이지의 장점을 적용해 다른 랜딩 페이지에도 유사한 사용자 경험을 제공할 수 있도록 구조 및 콘텐츠 개선



## 퍼널 테스트 분석 요약

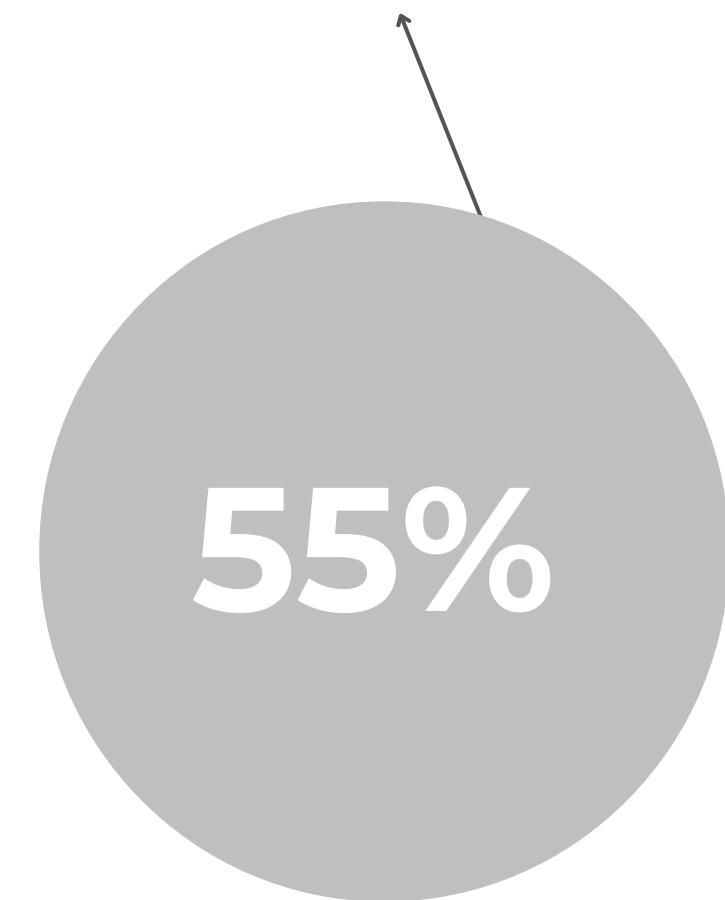
## • 결제 페이지 전환율:

구매 완료페이지로의 전환율은 상대적으로 낮아 결제 과정에서 문제가 발생할 가능성을 시사.



## • 제품 페이지 전환율:

랜딩 페이지에서 제품 페이지로의 전환율이 낮아, 유인 요소 추가 필요.

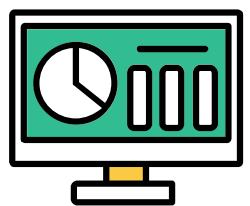


## • 배송 페이지 전환율:

배송페이지에서 결제 페이지로 전환율이 제일 높음.



## 전환 퍼널 최적화 전략



### 제품 페이지 전환율 개선

랜딩 페이지에서 제품  
페이지로의 전환율이  
55%로 낮음.



### 결제 전환율 개선

결제에서 구매 완료로  
의 전환율이 상대적으  
로 낮음 (61.46%).

## 문제점

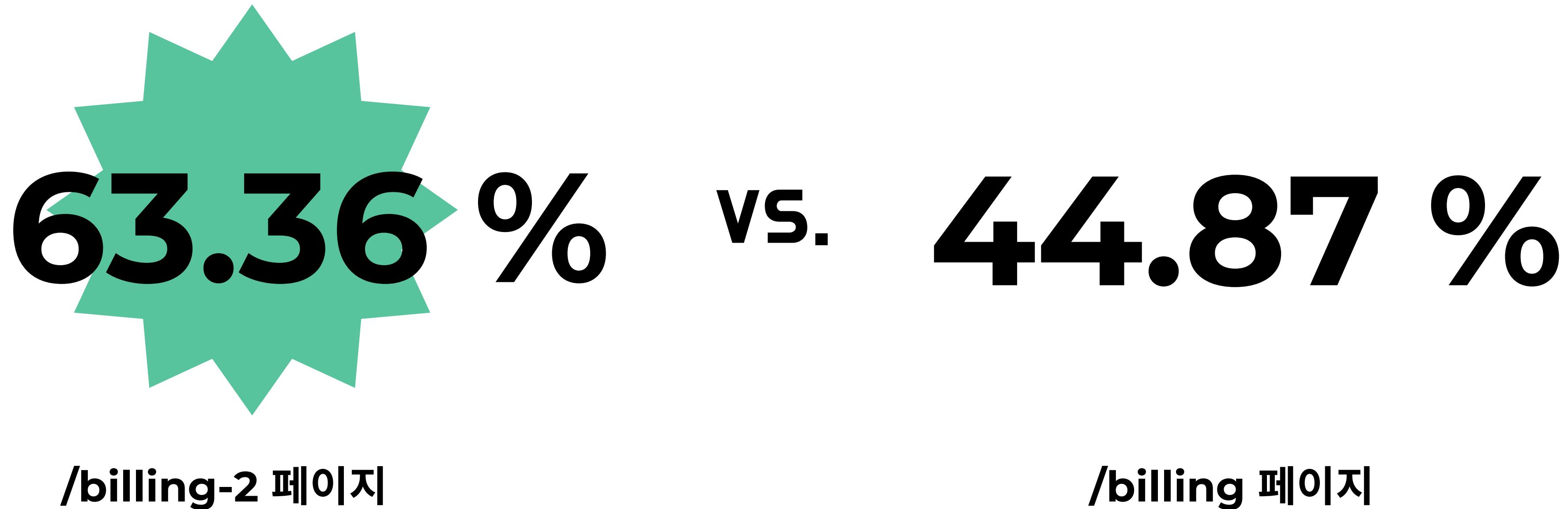
## 개선 전략

제품 페이지로 연결되는 링크  
강화, 초기 방문자에게 제품 특징  
을 강조하는 배너 및 팝업 등을  
추가해 전환율을 높일 수 있음.

결제 과정 간소화, 신뢰할 수  
있는 결제 보안 메시지 추가 등  
을 통해 결제 과정에서 발생할  
수 있는 이탈을 줄임.



### A/B 테스트 분석 요약



### 결제 페이지 최적화 전략



#### 문제점

/billing 페이지의 전환율이 낮아, 결제 과정에서의 이탈이 높음.

A/B 테스트 기반  
결제 페이지 최적화

#### 개선 전략

/billing-2 페이지의 성공적인 요소를 /billing 페이지에 반영해 구조를 개선하고, 사용자 피드백을 수집해 지속적인 최적화를 진행함.



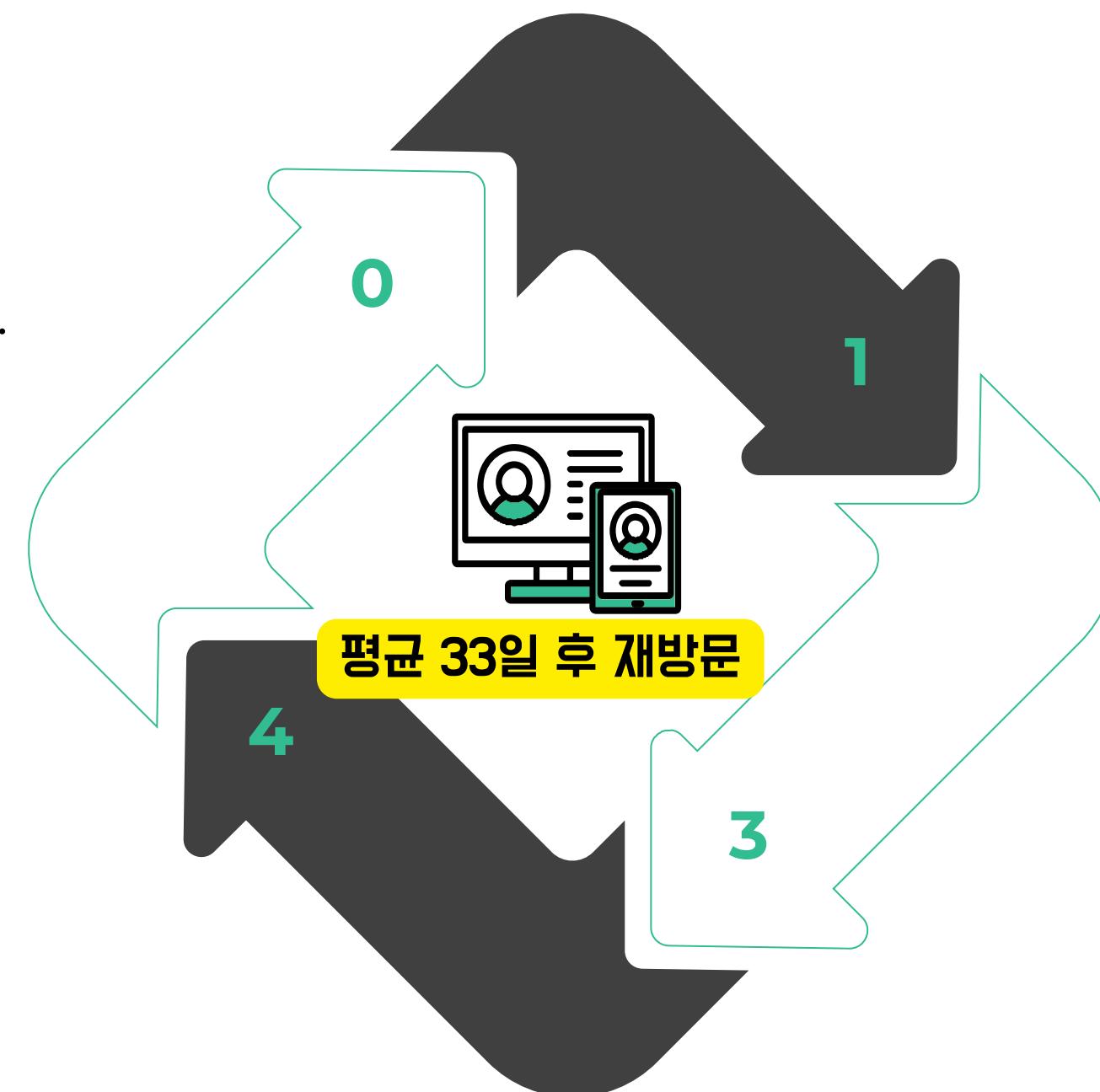
## 고객 행동 분석 요약

### 첫 방문이 마지막인 사용자

대다수 사용자(343,048명)는 한 번만 사이트를 방문하고 다시 방문하지 않음.

### 3회 이상 반복 방문 사용자

13,399명으로, 충성도 높은 사용자가 많음을 시사.



### 1회 반복 방문 사용자

37,386명으로, 비교적 많은 비율을 차지함.

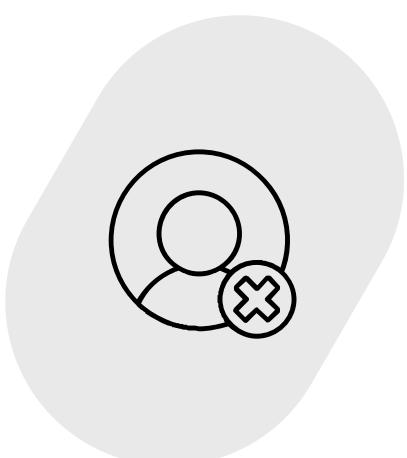
### 2회 반복 방문 사용자

485명으로 매우 적음. 1회 방문 후 다시 방문하지 않는 이유를 분석할 필요가 있음.



## 고객 리텐션 전략

## 문제점



대다수의 사용자는 한 번  
방문 후 재방문하지 않음.



## 개선 전략

## 리마인드 알림

재방문 사용자가 약 33일 후 다시 사이트를 방문하는 경향이 있으므로, 리마인드 이메일 및 푸시 알림을 이 시점에 맞춰 발송해 방문을 유도함.

## 재방문 유도 프로모션

첫 방문 후 즉각적인 할인 코드나 특별 혜택을 제공해 빠른 재방문을 유도할 수 있음.

## 장기 리마케팅 전략

최대 69일 후에 재방문하는 사용자를 대상으로 장기적인 리마케팅 전략을 수립해, 이탈한 고객을 다시 유입시킴.



### 한계점 및 아쉬운 점

- 1. 적절한 데이터 확보의 어려움:** 실제 이커머스 데이터 접근이 어려워 강의 데이터셋 활용. 비즈니스 복잡성을 충분히 반영하지 못함.
- 2. 정제된 데이터셋 사용:** 이미 정제된 데이터셋 사용으로 실무 데이터 준비 단계의 도전성 부족.
- 3. 다양한 분석 기법의 적용:** 전환율 최적화를 위한 여러 분석 기법 적용과 결과 해석에 시간 많이 소요됨. 분석 기법을 한 두개만 선정해서 더 깊게 분석 해보는 방법이 더 나은 결과를 보여줄 수도 있었음.

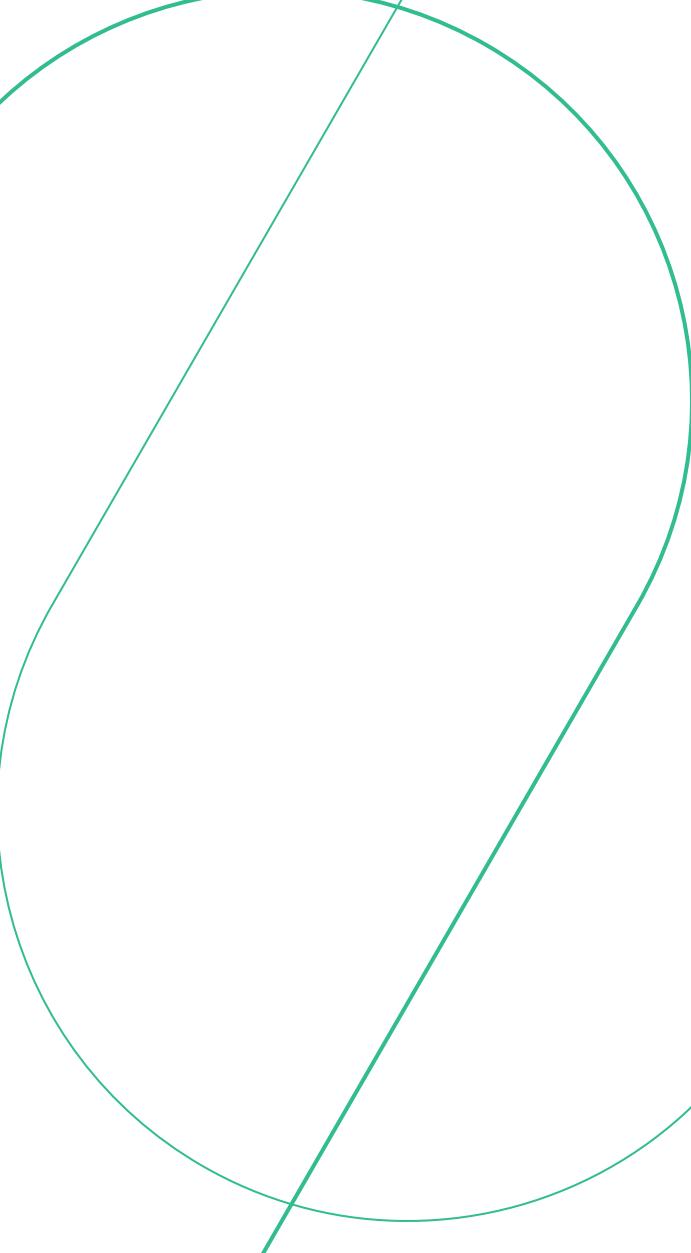
### 향후 도전해볼만 분석 계획

**고객 생애 가치(Lifetime Value) 예측:** 고객 세그먼트별 장기적인 가치를 예측하여 마케팅 자원을 효율적으로 배분하고 최적화된 고객 관리 전략 수립.

**추천 시스템 최적화:** 데이터 기반 상품 추천 알고리즘을 개선하여 개인화된 고객 경험 제공.

**동적 가격 전략 수립:** 실시간 수요와 경쟁사 가격 데이터를 활용해 가격을 최적화하여 매출 극대화.

**반품 패턴 분석 및 개선:** 반품 사유를 파악하여 제품 개선을 통한 반품율 감소 목표.



**감사합니다.**