

인과 효과 추정 모형으로 온사이트 마케팅 캠페인 효과 분석하기

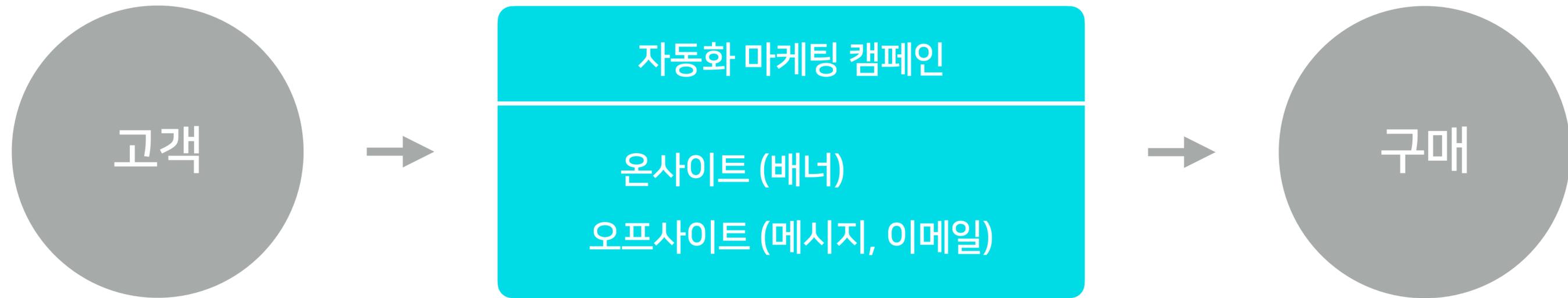
이민호 데이터라이즈

CONTENTS

1. 배경
2. 인과 효과를 추정하는 방법
3. 신규 기능 배포가 지표에 미치는 인과 효과 확인하기
4. 온사이트 마케팅 캠페인이 어떤 고객에 효과적이었는지 확인하기
5. 마무리

1. 배경

고민의 시작



우리의 마케팅 캠페인은 잘 동작하고 있을까? 🤔

마케팅 캠페인의 효과를 확인하는 방법



인과 효과를 확인하는 가장 좋은 방법은 A/B 테스트

그런데 A/B 테스트를 수행하기 어렵다면 어떻게 해야 할까?

A/B 테스트가 어려운 환경

고객사에서 원하지 않는 경우

- 경험적으로 이벤트 효과를 알고 있고, 최대한의 성과를 달성하고자 트래픽을 분산시키지 않을 때
- 특정 고객군에만 제공된 혜택으로 인해 고객이 클레임을 제기할 우려가 있을 때
- 이런 경우에는 실험을 진행하는 것이 어렵다.

여러 가지 온사이트 마케팅 캠페인이 서로 영향을 미치는 환경

- 고객이 사이트에 방문해서 구매를 완료할 때까지 여러 캠페인이 서로 영향을 주고 받는다.
- 특정 캠페인의 효과를 확인하기 어렵다.

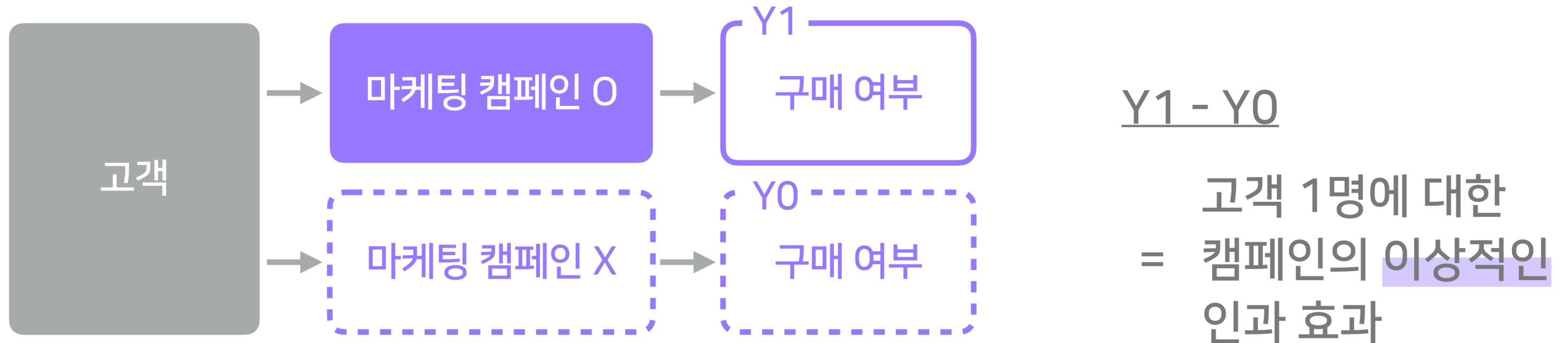
A/B 테스트를 하지 못하더라도, 마케팅 효과는 확인하고 싶다면 어떻게 할까?



우리는 데이터로 먹고 사는 회사.
데이터를 통해 인과 효과를 추정해 볼 수는 없을까?

2. 인과 효과를 추정하는 방법

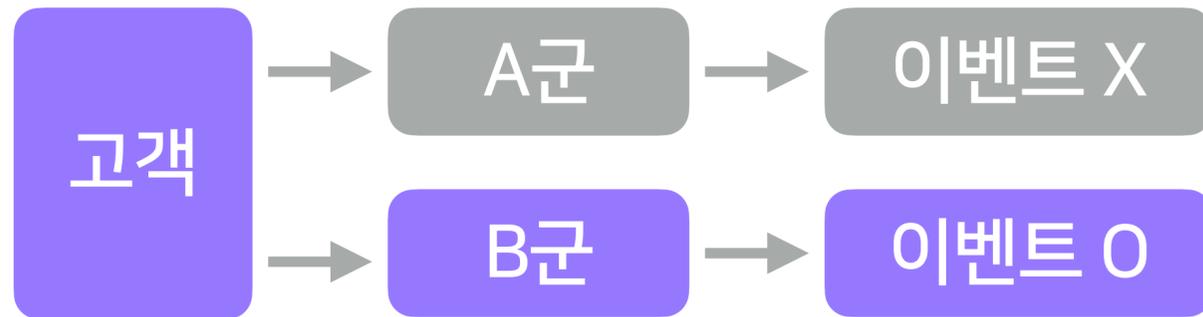
인과 효과를 어떻게 추정할 수 있을까?



하지만 현실에서는 두 가지 상황을 **동시에 경험할 수 없다.**

현실적인 대안

두 그룹으로 나누어
평균적인 효과를 확인한다.



인과 모형을 통해
효과를 추정한다.



- A/B 테스트
- 두 그룹이 구분하기 어려울 정도로 비슷하게 나누어져 있다면 평균적인 인과 효과를 확인할 수 있다.

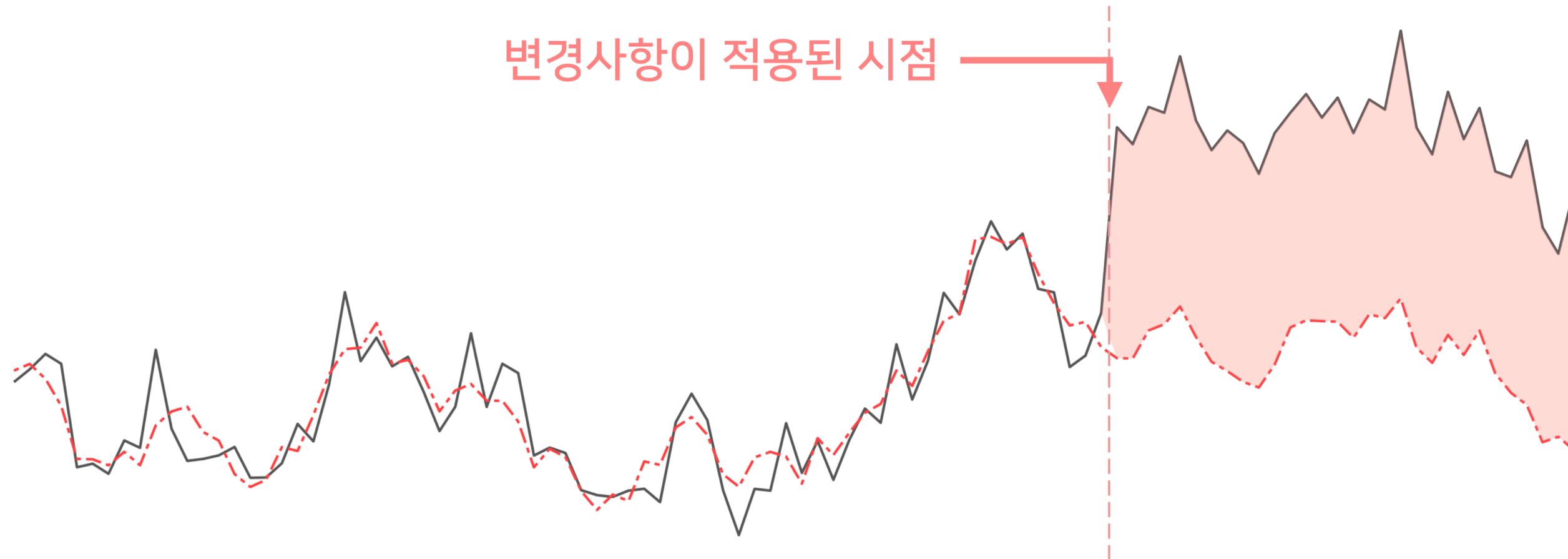
오늘 다루게 될 주제!

인과 효과를 추정해야 하는 비즈니스 사례

1. 특정 날짜에 반영된 변경 사항이 지표에 얼마나 영향을 줄까?
2. 외부에서 발생한 변화가 지표에 미치는 영향은 어떻게 확인할 수 있을까?
3. 전체적인 인과 효과가 아니라 고객 개개인에 대한 영향은 어떻게 알 수 있을까?
4. 변경 사항이 긍정/부정적인 영향을 미치는 고객군은 어떤 특성을 가지고 있을까?

3. 신규 기능 배포가 지표에 미치는 인과 효과 확인하기

일별 지표 추이를 보고 인과 효과를 확인하는 방법



- 어떤 변화가 발생한 시점을 알고 있을 때, 일별 지표에 미친 영향을 확인하고 싶다면 어떻게 해야 할까?
- 배포, 이벤트, 정책 변경, 외부 영향 등등

일별 지표 추이를 보고 인과 효과를 확인하는 방법

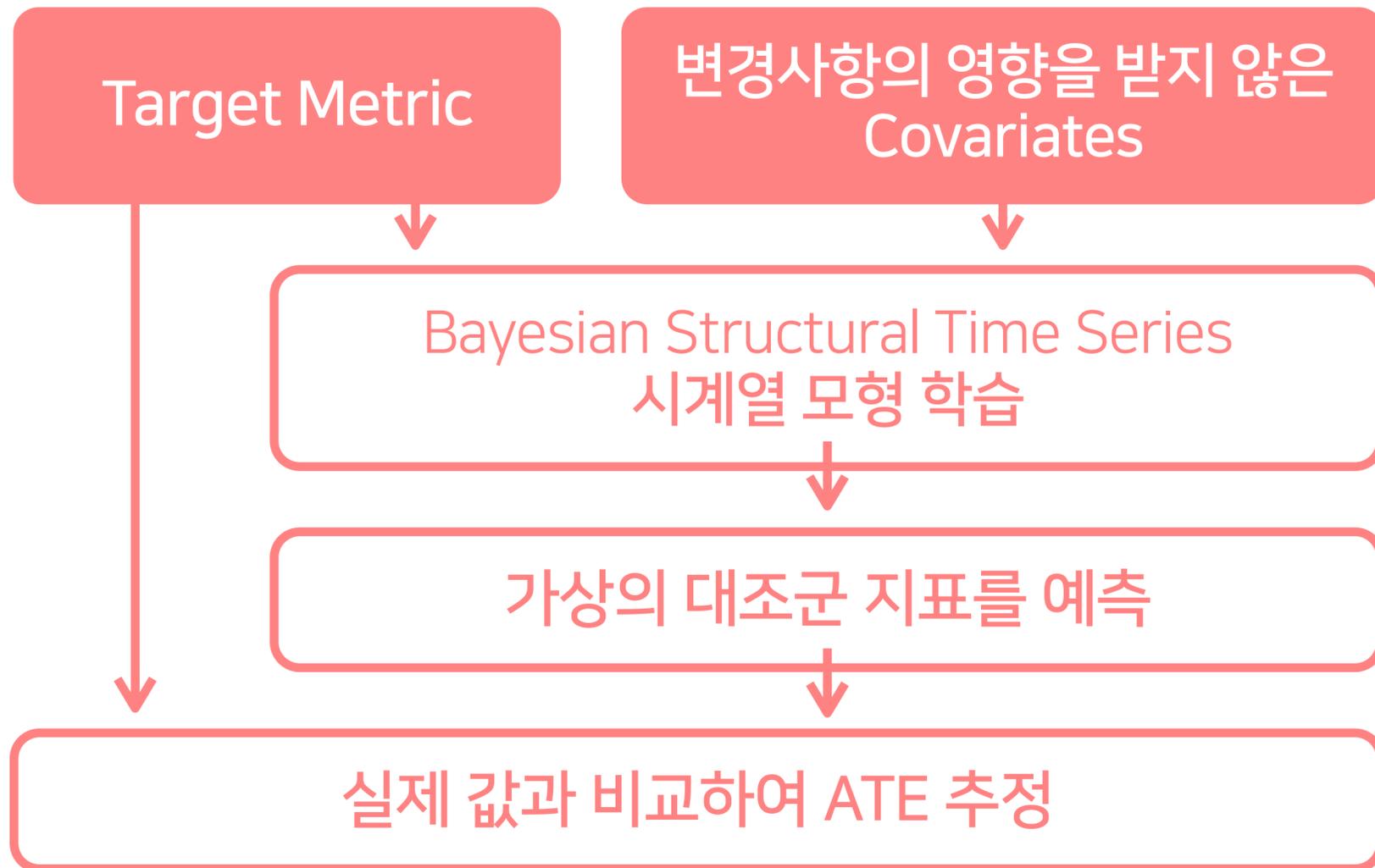
목표 : 변경 사항으로 인한 인과 효과를 확인한다.

- 전체적인 영향을 ATE Average Treatment Effect 를 계산하여 확인한다.



가상의 대조군 지표를 추정하여 비교하기

CausalImpact



대조군 지표를 추정하기 위해 사용하는 변수에는 변경사항으로 인한 영향이 없어야 한다.

목표 변수와 독립 변수들 사이의 관계가 변경사항 적용 전/후에 동일하게 유지되어야 한다.

참고.

Synthetic Control 방법론은 시계열 모형 대신 독립 변수들의 Convex Combination으로 예측한다.

[CASE 1] 신규 기능 배포로 인한 영향

Background

- 온사이트 배너에 신규 기능을 배포했다.
- 고객사 요청으로 A/B 테스트 없이 전체 반영했다.
- 비슷한 시기에 고객사에서 대규모 이벤트가 진행되었다.

목표

- 신규 기능 배포가 CTR에 미친 영향을 확인한다.

[CASE 1] 신규 기능 배포로 인한 영향

데이터

목표 변수

배포 전/후 CTR 추이

분석 대상 사이트

주요 지표
(세션 수 등)

배너 관련 지표
(변경사항 반영되지
않은 페이지)

독립 변수

유사 사이트
(업종, 규모)

주요 지표

배너 관련 지표

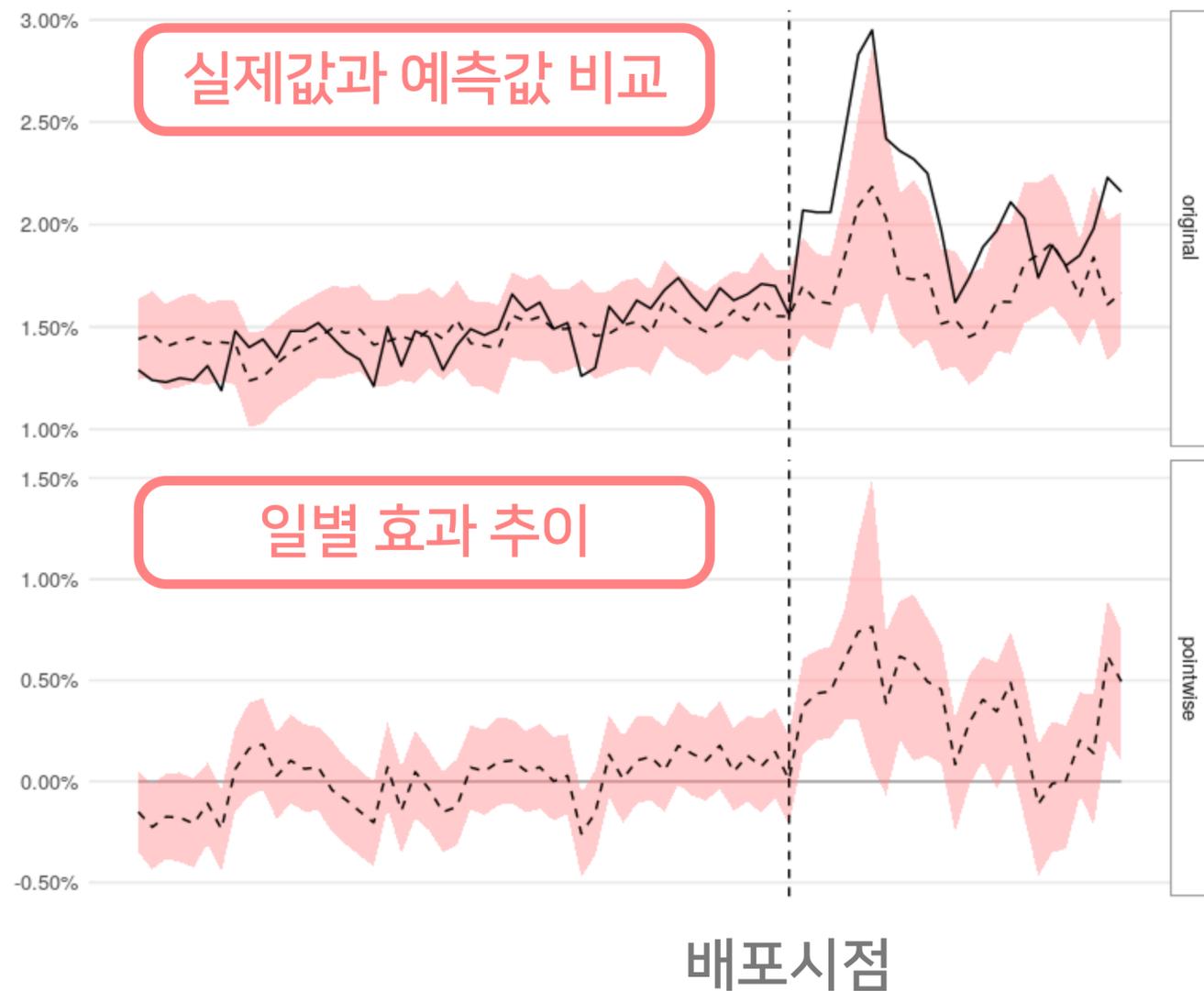
검색어 트렌드
(네이버/구글)

사이트 검색어

주요
유입 검색어

[CASE 1] 신규 기능 배포로 인한 영향

모델링 결과 해석 (1) 배포 효과 확인



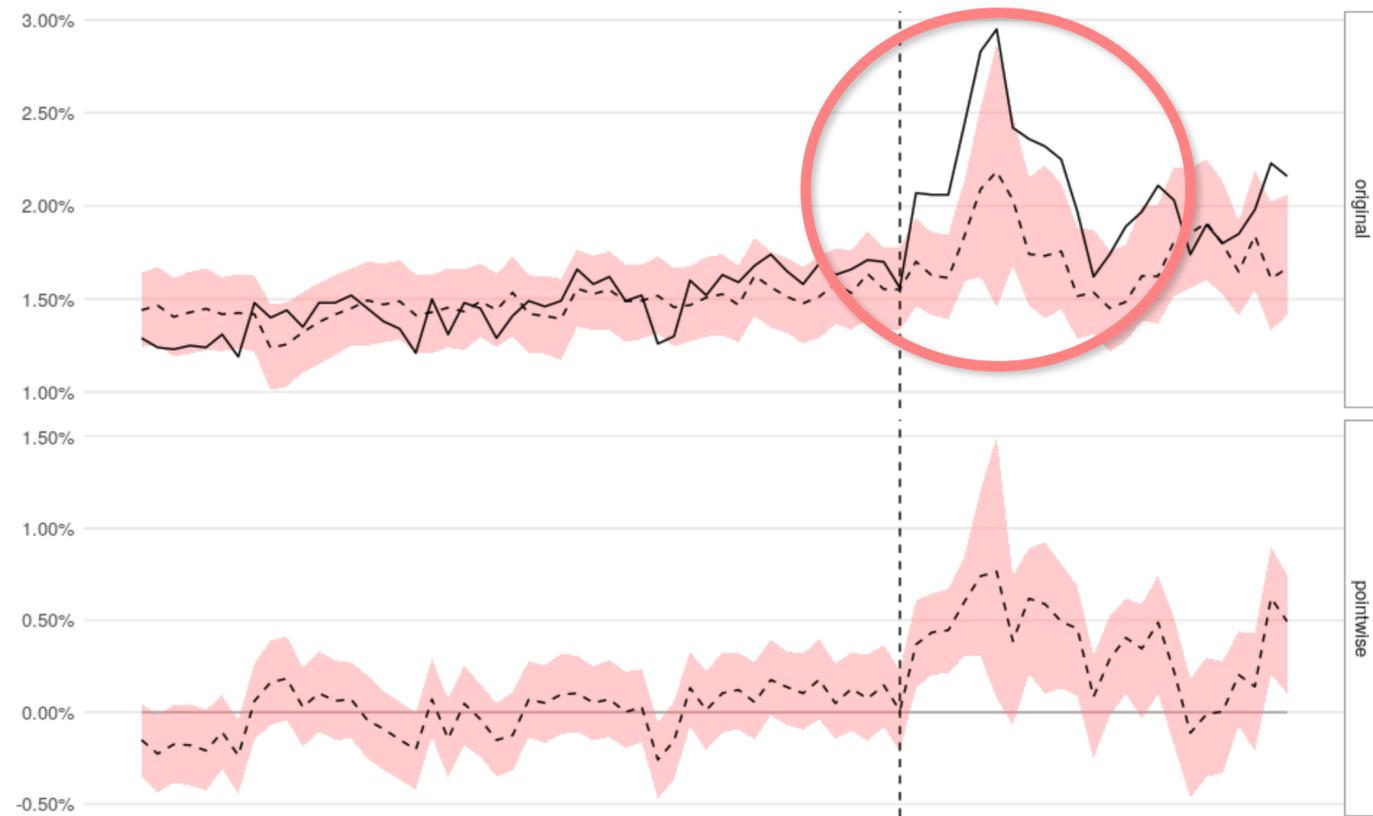
변경사항 배포 이후 0.38%p (22%) CTR 증가 효과

다만, 이벤트 기간이 지난 이후에는
CTR 증가폭이 명확하게 보이지 않고있다.

[CASE 1] 신규 기능 배포로 인한 영향

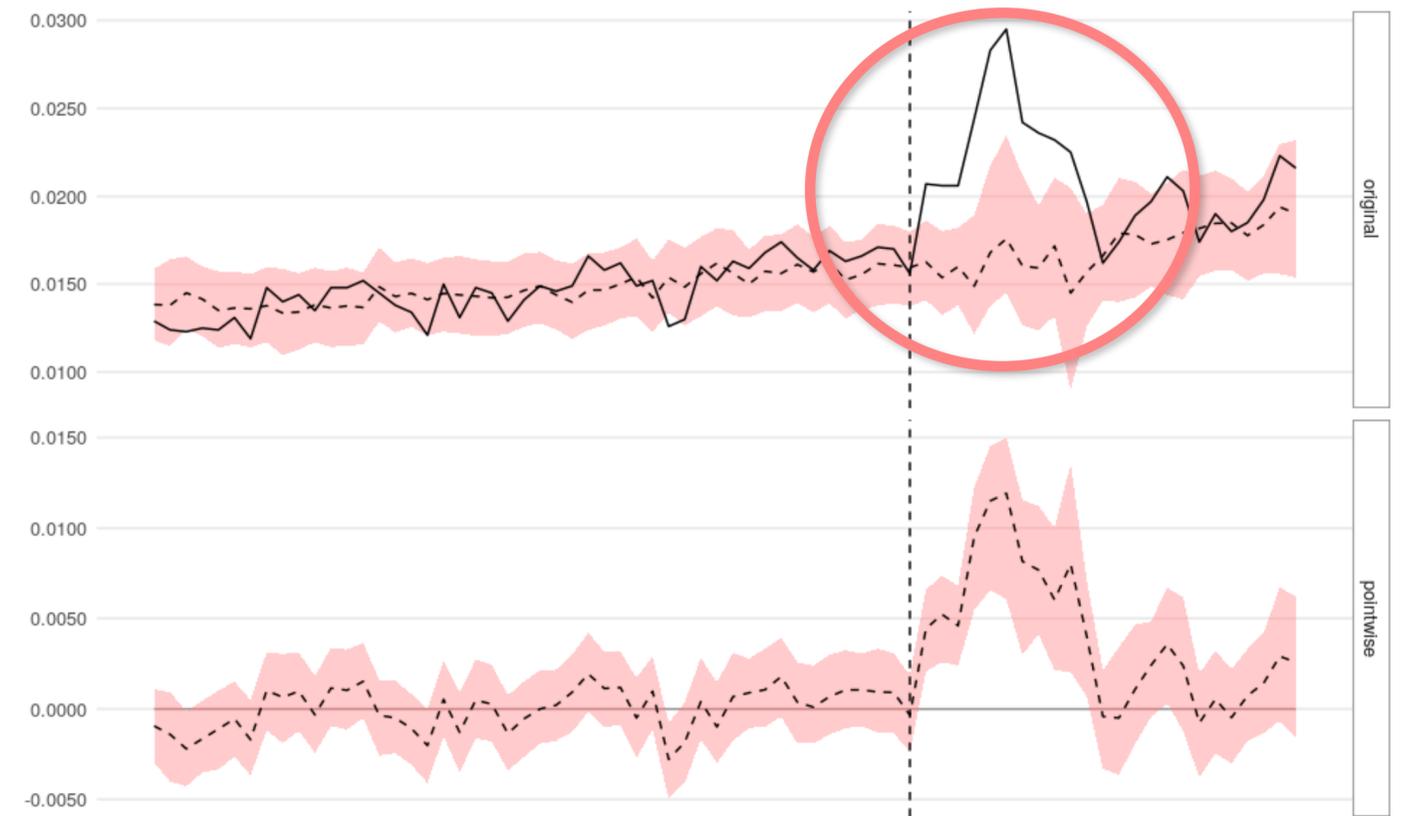
모델링 결과 해석 (2) Control 예측에 사용하는 변수 변경

이벤트의 영향을 받은 변수로 예측



추정한 인과 효과에서 이벤트의 영향이 필터링됨

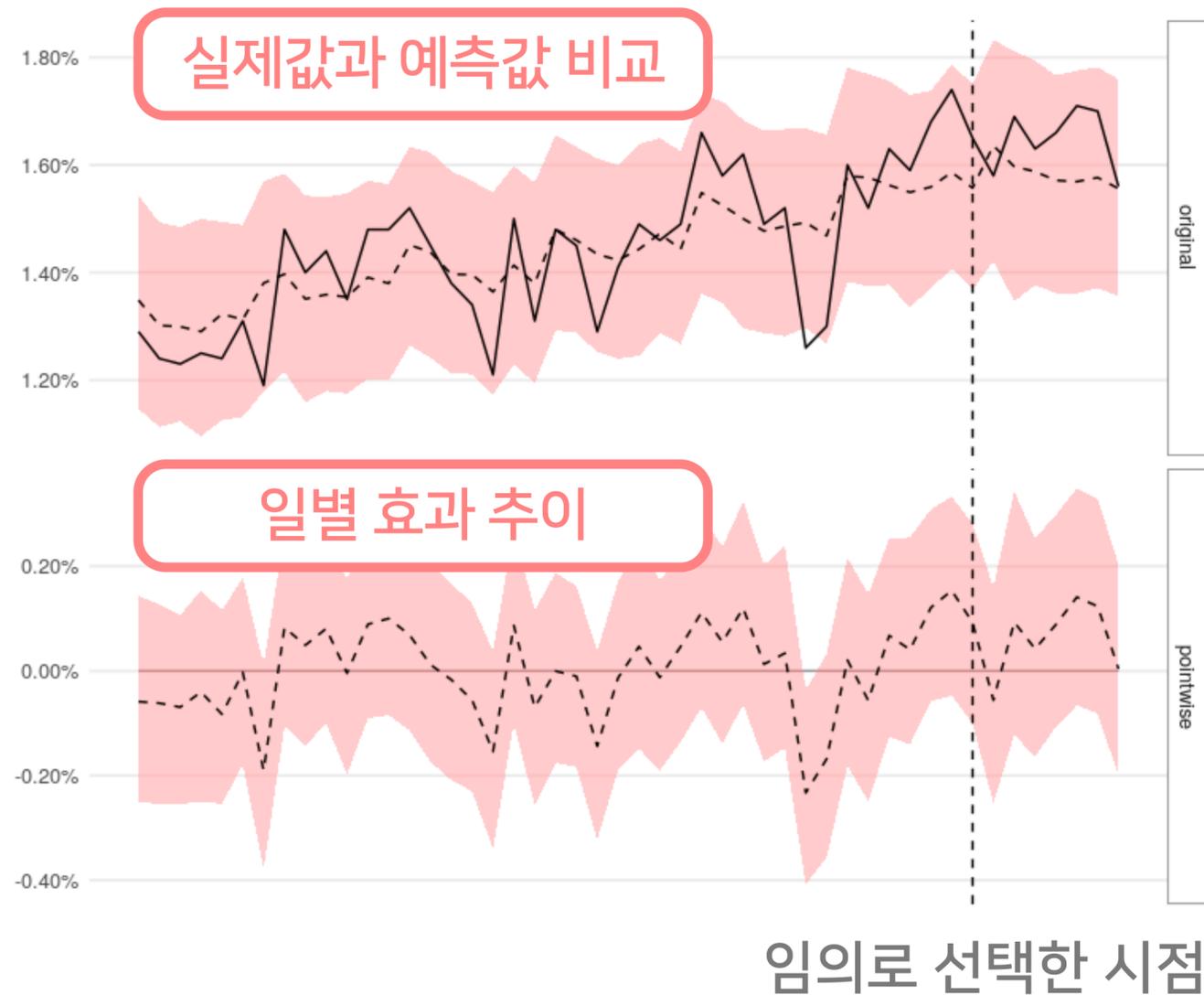
이벤트의 영향을 받지 않은 변수로 예측



추정한 인과 효과에 이벤트의 영향까지 포함됨

[CASE 1] 신규 기능 배포로 인한 영향

모델링 결과 해석 (3) 배포/이벤트 영향이 없는 시기



배포와 이벤트 영향이 없는 시기를 학습했을 때는 효과가 없는 것으로 추정되어야 한다.

동일한 변수로 배포 전 일주일 간의 지표를 학습한 결과, 통계적으로 유의미하지 않은 0.062%p (3.9%)의 변화를 확인할 수 있었다.

[CASE 2] 고객사의 검색광고 키워드 제거 영향

Background

- 고객사에서 검색 광고 키워드를 일부 제거한 이후 지표에 미친 영향을 확인해 달라고 요청이 왔다.
- 그런데 키워드를 제거하는 시점에 다양한 이벤트가 동시에 진행되고 있었다.
- 지표에는 어떤 영향이 있었을까?

목표

- 검색 광고 키워드를 제거한 것이 지표에 미친 영향을 확인한다.

[CASE 2] 고객사의 검색광고 키워드 제거 영향

데이터

목표 변수

검색 광고 키워드
제거 전/후 세션 수 차이

독립 변수

유사 사이트
(업종, 규모)

주요 지표

검색어 트렌드
(네이버/구글)

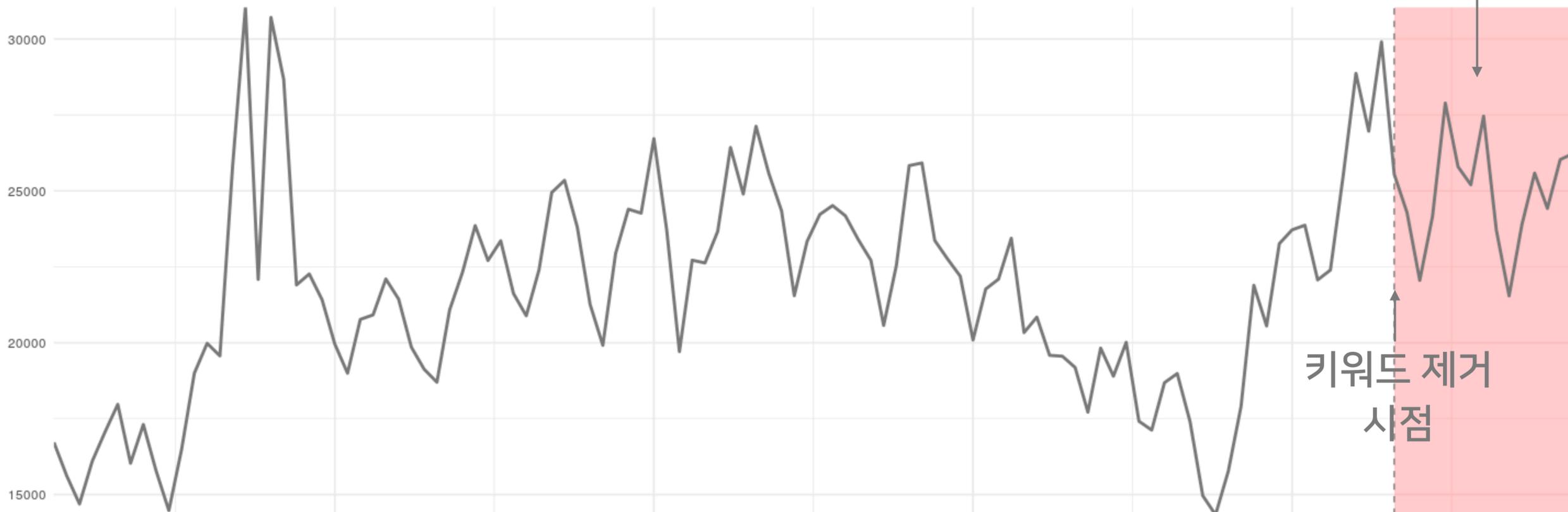
사이트 검색어

주요
유입 검색어

[CASE 2] 고객사의 검색광고 키워드 제거 영향

(1) 세션수 추이 확인 (인과 모형 적용 전)

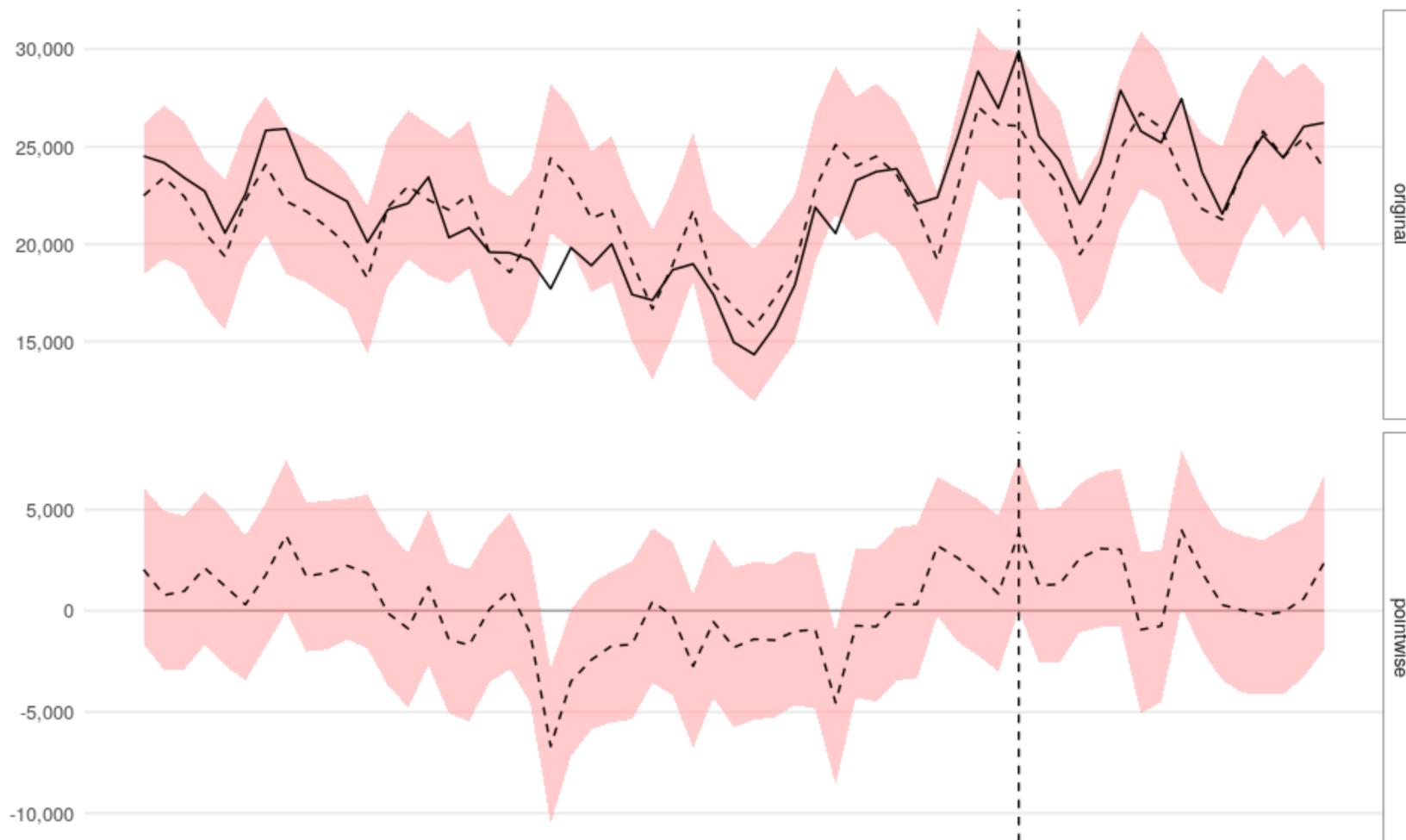
키워드 제거 전/후 세션수 추이



🤔 광고 키워드를 제거해서 세션 수가 줄어들었다고 봐야 할까?

[CASE 2] 고객사의 검색광고 키워드 제거 영향

(2) 광고 키워드 제거 효과 확인 (인과 모형 적용 후)



추정 결과!

세션수가 감소한 효과는 확인되지 않았다.

오히려 통계적으로 유의미하지는 않지만 세션수가 5.2% 상승한 것으로 나왔다.

광고 키워드를 제거한 이후 세션수가 줄었지만, 유입 키워드 등 다른 변수들의 변화를 통해

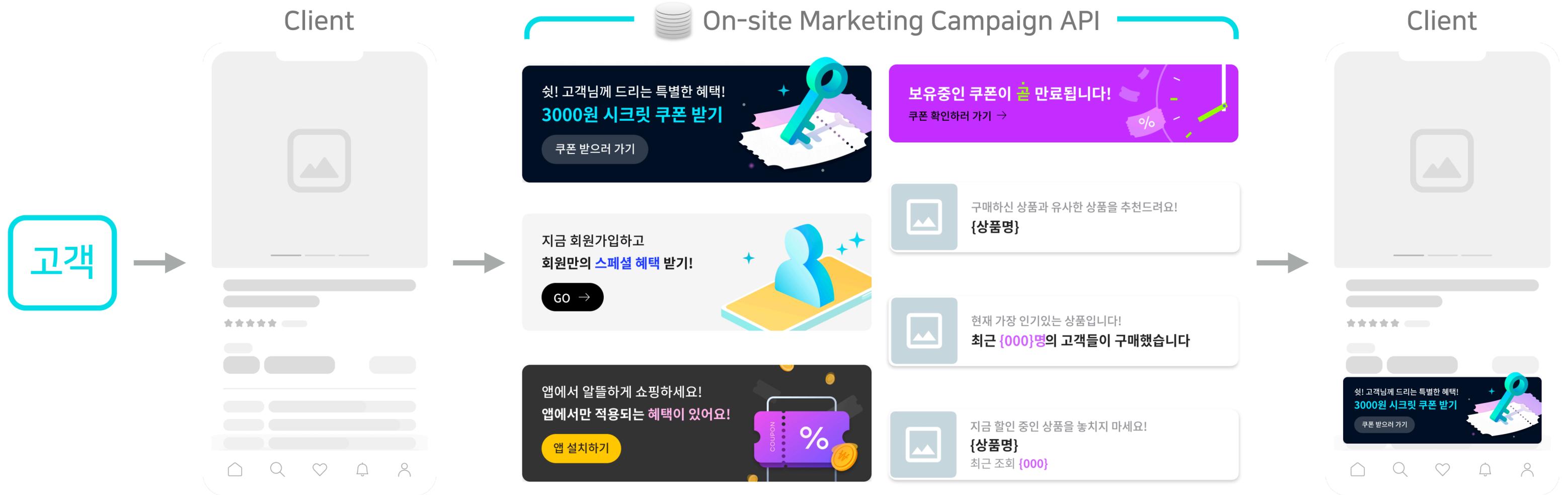
충분히 설명할 수 있는 상태로 보인다.

(업종 자체의 비수기일 가능성이 있다.)

광고 키워드 제거 이후 진행한 다른 이벤트로 세션수 감소를 방어했다고 해석할 수도 있다.

4.온사이트 마케팅 캠페인이 어떤 고객에 효과적이었는지 확인하기

Background : 온사이트 마케팅 캠페인

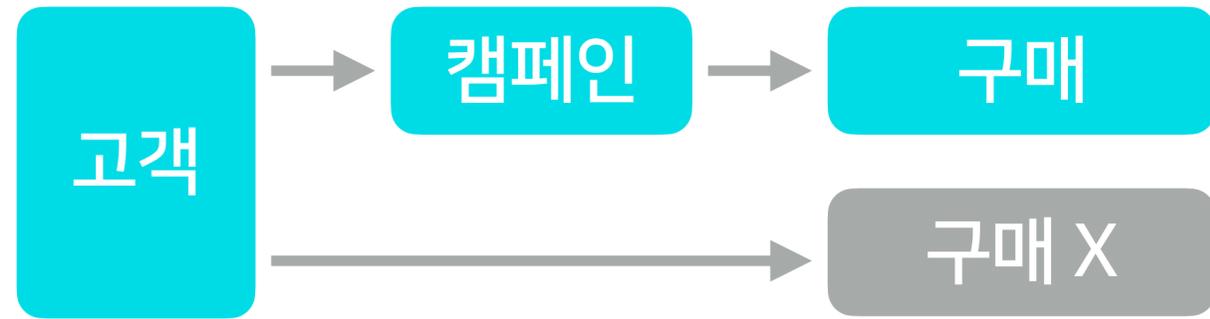


고객사 사이트
페이지뷰

고객에게 적용할 수 있는
온사이트 마케팅 캠페인 리스트 확인

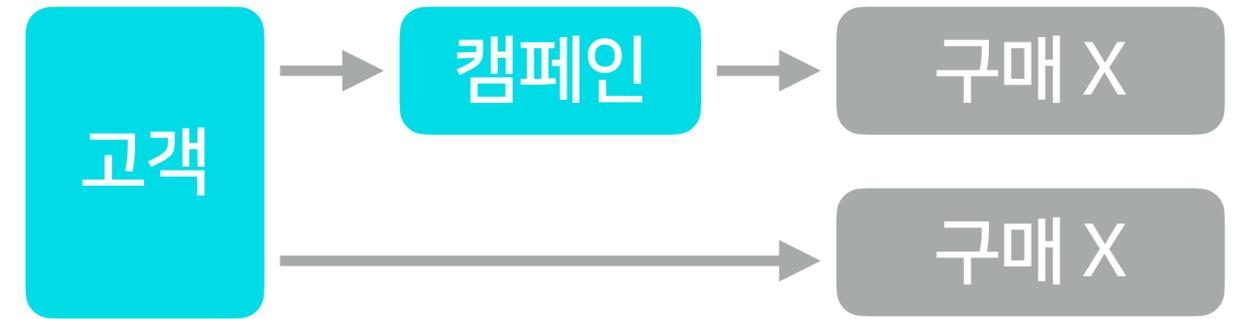
고객의 특성과
현재 상황에 가장 맞는
마케팅 캠페인 적용

고객별 캠페인 효과를 확인하는 방법



고객 A

캠페인 효과가 큰 경우
→ 캠페인을 지속적으로 적용



고객 B

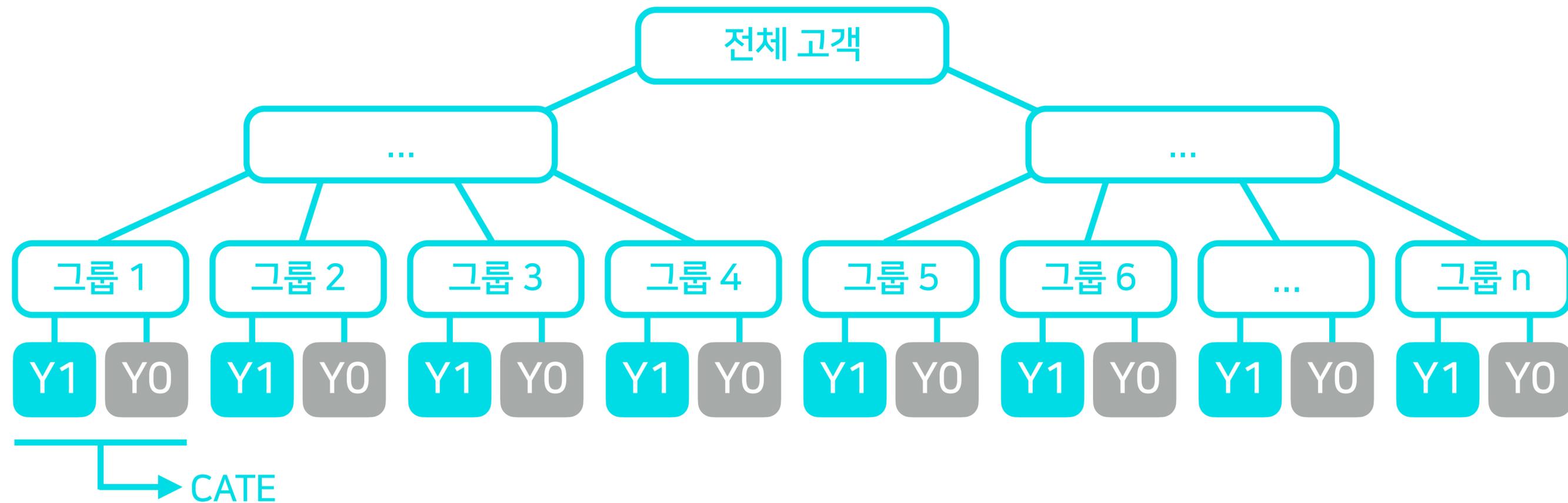
캠페인 효과가 적은 경우
→ 캠페인을 변경하거나 적용하지 않음

- 마케팅 캠페인으로 인한 구매 유도 효과는 **사람마다 다르게** 나타난다.
- 인과 효과를 바탕으로 마케팅 캠페인의 효율을 높이려면 고객마다 미치는 영향을 파악해야 한다.

고객별 캠페인 효과를 확인하는 방법

목표 : 마케팅 캠페인이 각 고객에 미치는 영향을 확인한다.

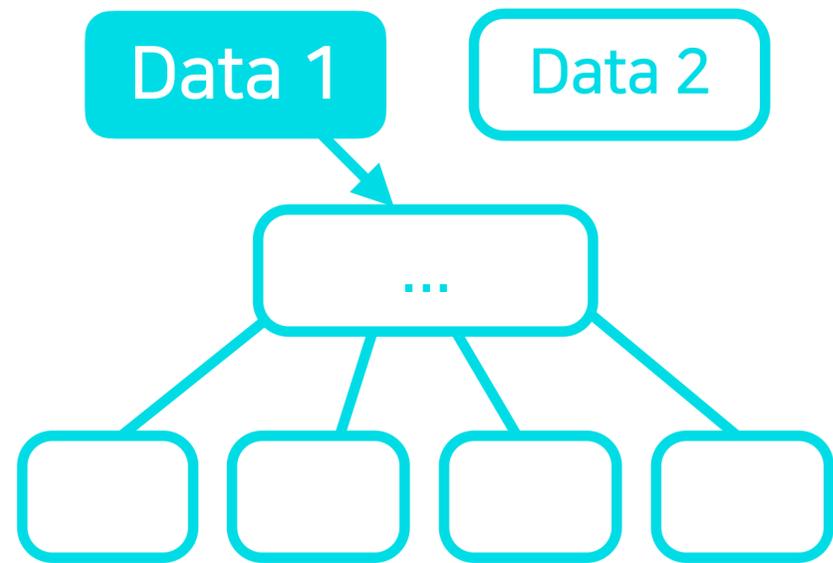
- 고객군을 나누고 CATE Conditional Average Treatment Effect 값을 계산한다.



개별 고객에 대한 캠페인 효과 추정하기

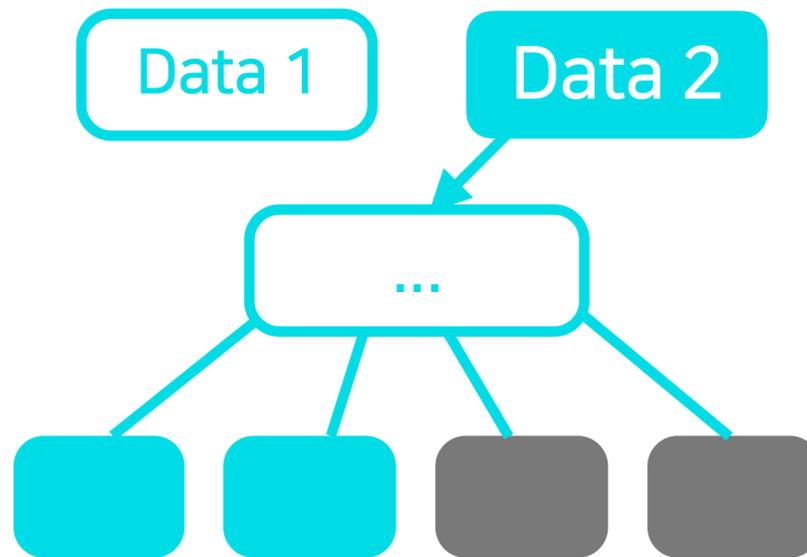
Causal Forest 모형 : key concept

(1) 데이터 중 일부를 사용해 Treatment Effect를 극대화하도록 트리를 생성한다.



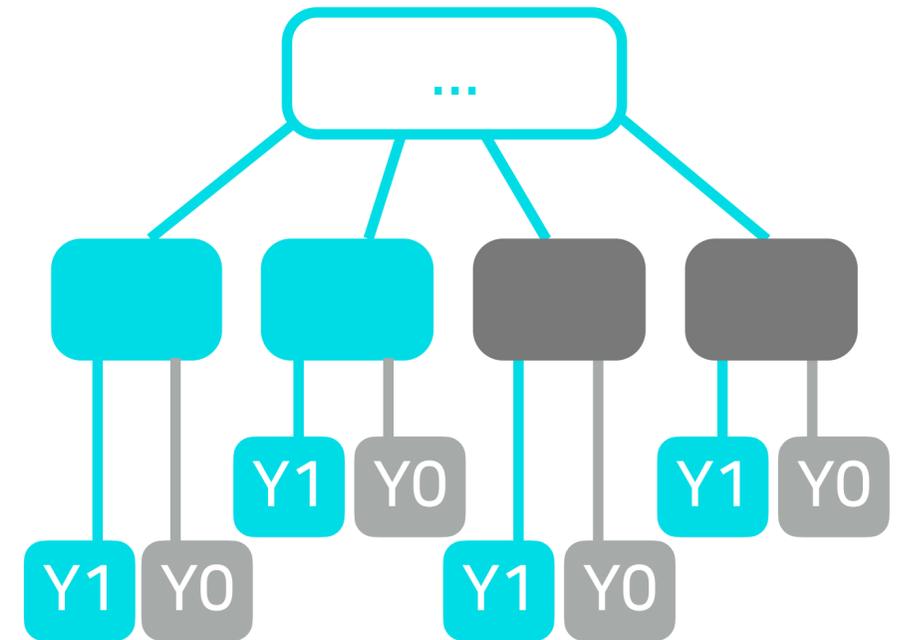
여기서는 캠페인으로 인한 효과를 극대화 시키도록 학습한다.

(2) 나머지 데이터를 트리로 내려보낸다.



Leaf Node에 비슷한 특성의 데이터가 모인다.

(3) 각 그룹별로 Treatment Effect를 계산한다.



[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인

Background

- 온라인 쇼핑몰에 방문한 고객은 다양한 온사이트 마케팅 캠페인에 노출된다.
- 온사이트 마케팅 캠페인끼리 서로 영향을 주고받을 수 있다.
- 어떤 고객은 캠페인에 잘 반응하는데, 그렇지 않은 고객도 있다.

목표

- 고객을 구매로 유도하는데 결정적인 영향을 미치는 캠페인을 확인한다.
- 온사이트 마케팅에 긍정적으로 반응하는 고객을 찾아낸다.

[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인

데이터

목표 변수

캠페인 노출 이후
구매 여부

독립 변수

고객 상태

신규 방문 여부

앱 사용 여부

CLV

이탈률

캠페인 노출 전
고객 데이터 요약

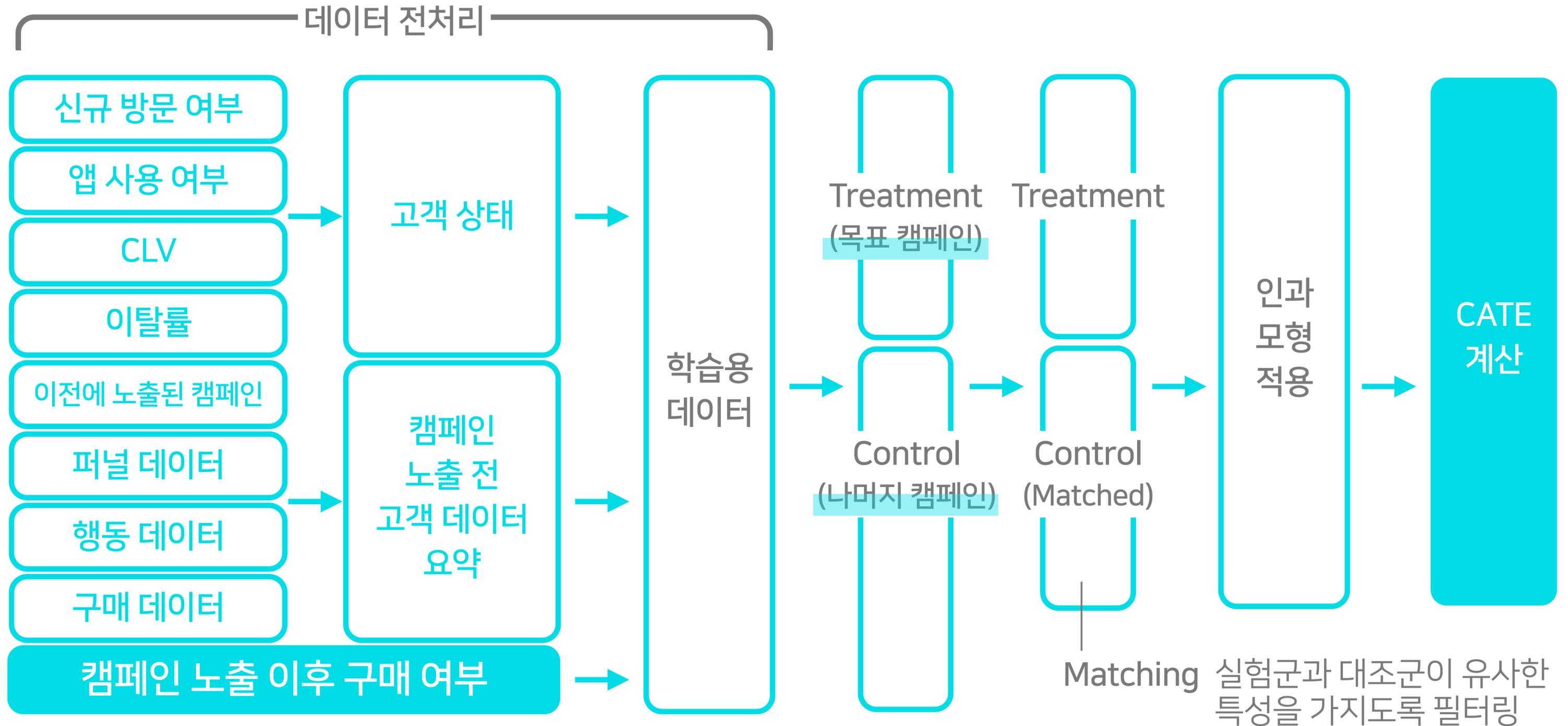
이전에 노출된 캠페인

고객이 도달한 퍼널

고객 행동 데이터

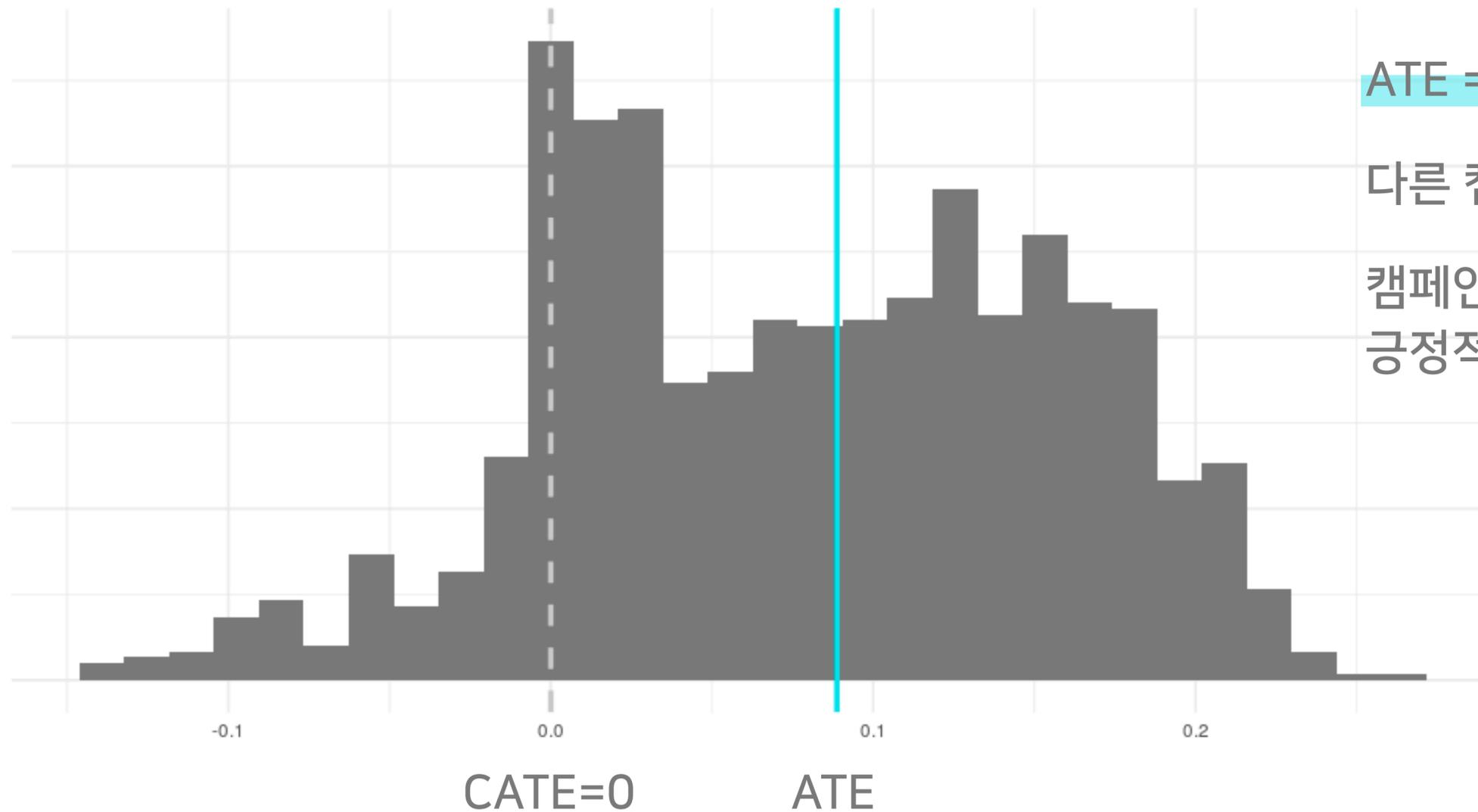
고객 구매 데이터

[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인



[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인

(1) 쿠폰 발급 캠페인



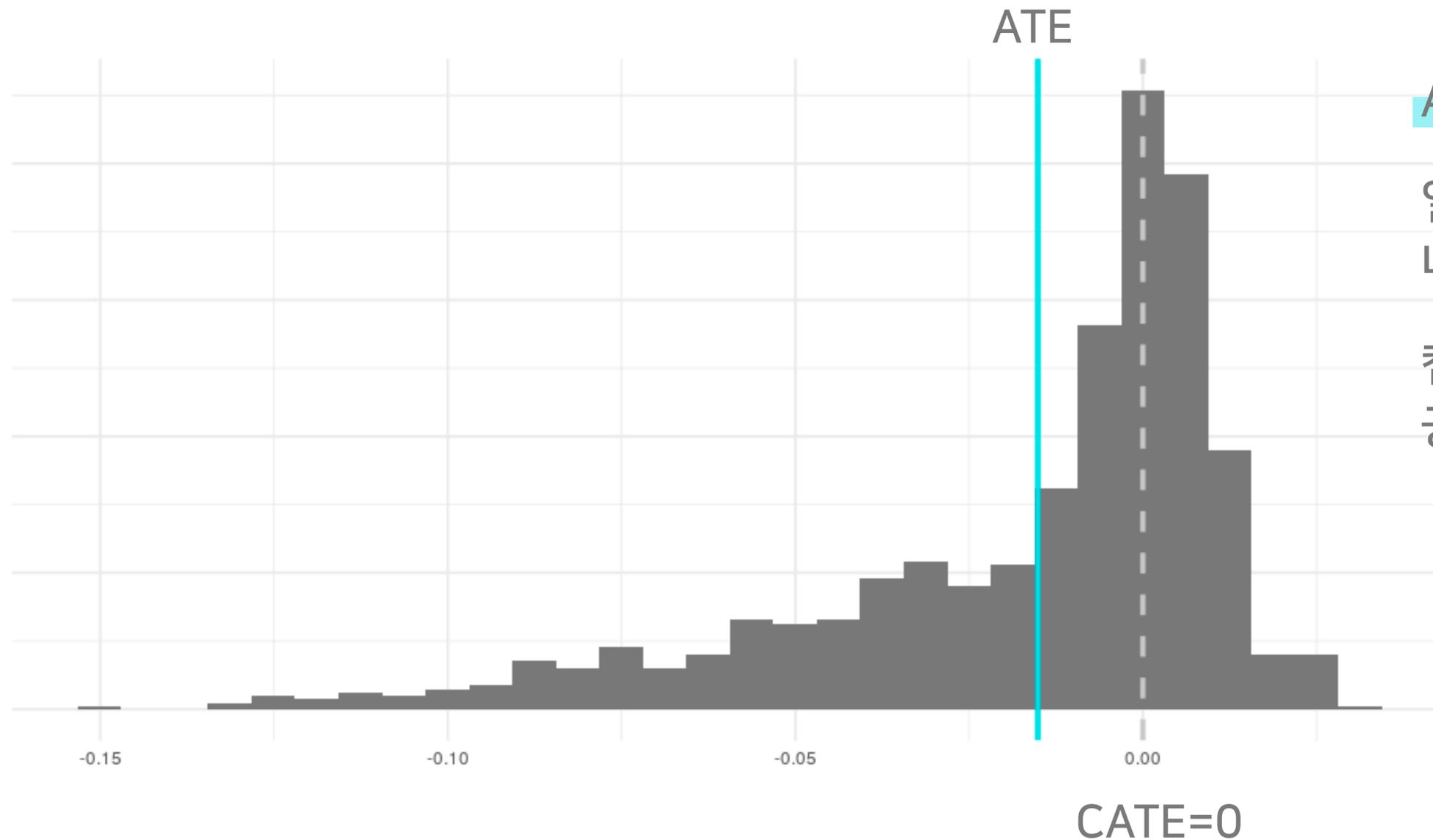
ATE = +0.088

다른 캠페인들에 비해 평균적인 효과가 높은 편

캠페인에 노출된 고객 중 85.3%에게
긍정적인 효과가 있었다.

[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인

(2) 연관 상품 추천 캠페인



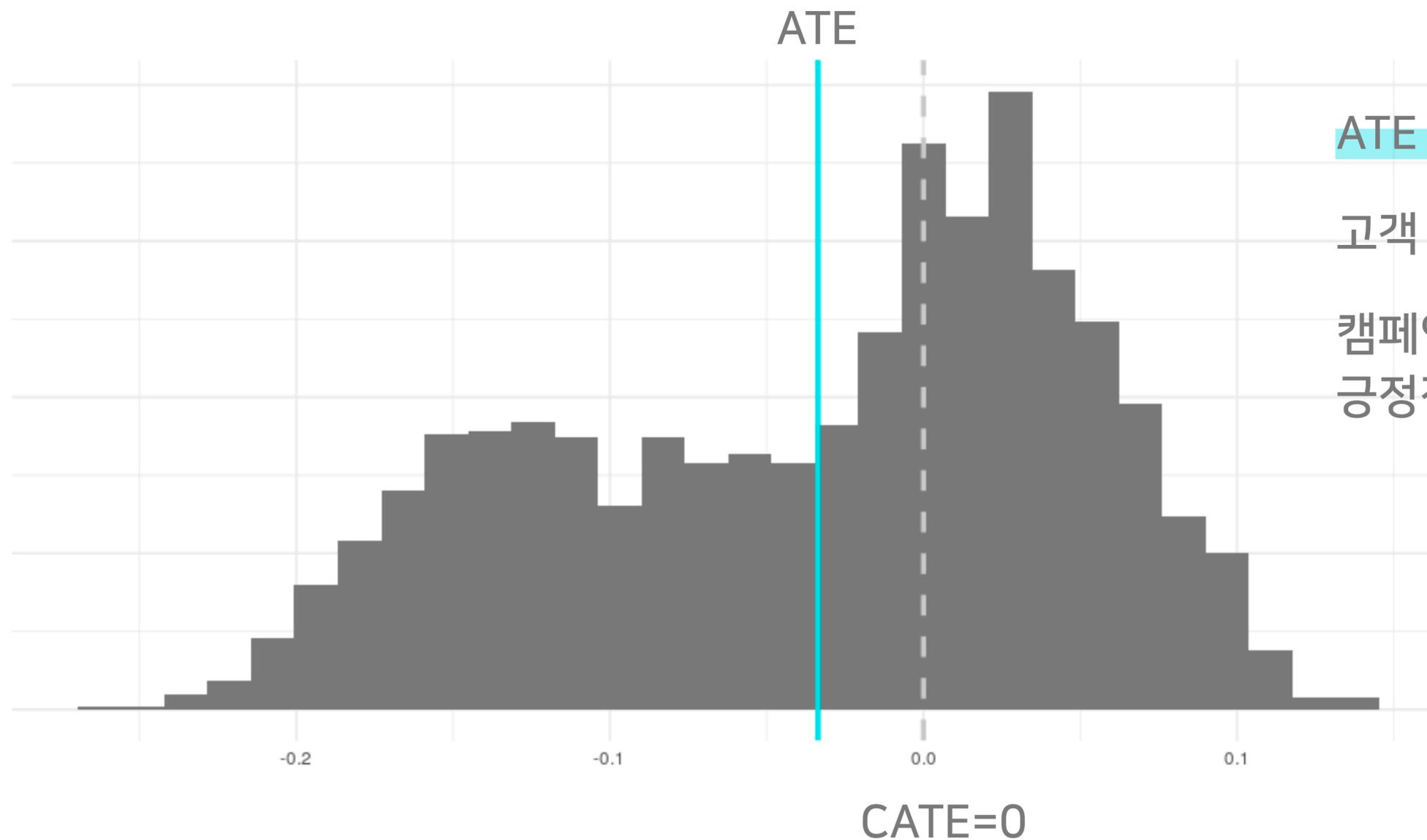
ATE = -0.015

일부 고객들에게서 상당히 부정적인 효과가 나타나고 있다.

캠페인에 노출된 고객 중 37.9%에게 긍정적인 효과가 있었다.

[CASE 3] 온사이트 마케팅의 고객별 인과 효과 확인

(3) 출석체크 혜택 알림 캠페인



ATE = -0.034

고객 특성에 따라 긍정/부정 영향이 갈리고 있다.

캠페인에 노출된 고객 중 43.4%에게
긍정적인 효과가 있었다.

[CASE 4] 각 캠페인이 집중해야 하는 고객군 확인

Background

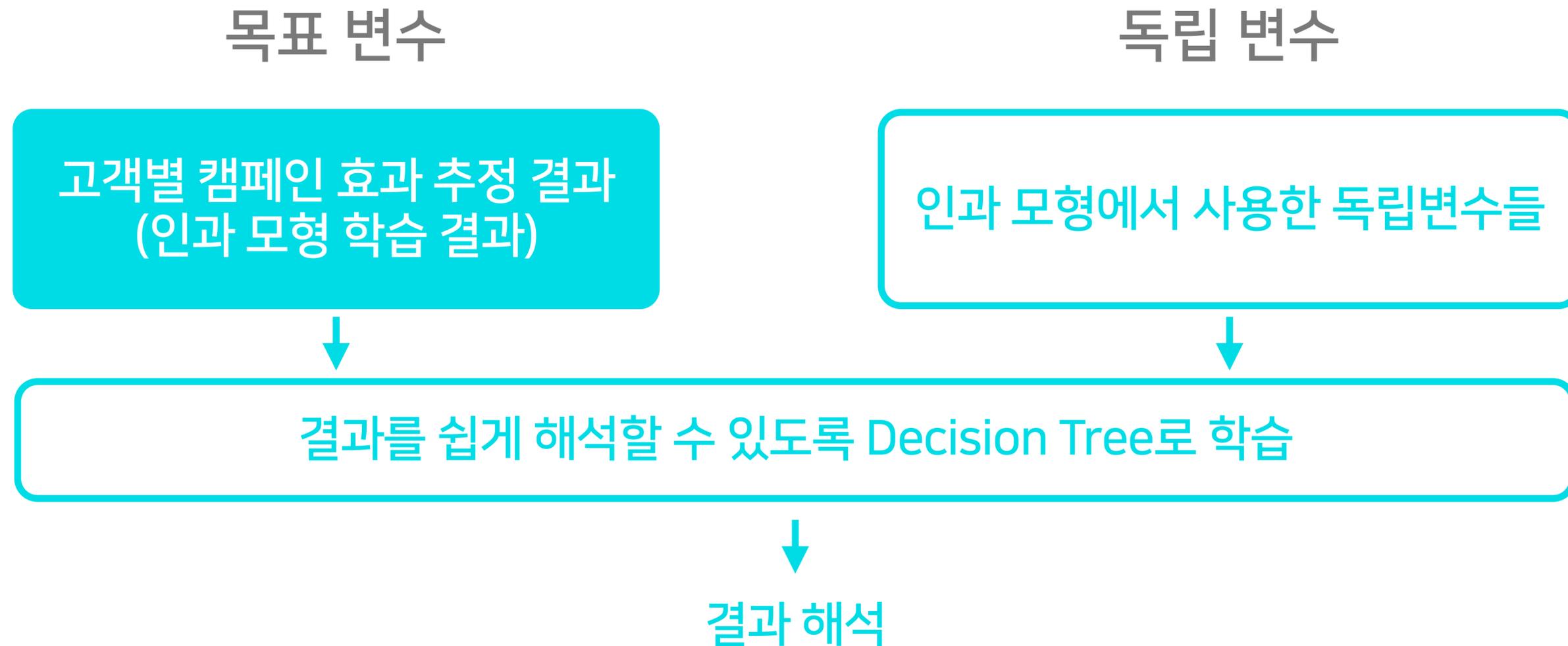
- 온사이트 마케팅 캠페인의 효율을 높이기 위해 캠페인 효과가 좋은 고객군을 찾고 싶다.
- 인과 모형의 결과를 마케팅 담당자도 직관적으로 이해할 수 있도록 설명하고 싶다.

목표

- 마케팅 캠페인이 어떤 고객군에 잘 동작했는지 확인한다.

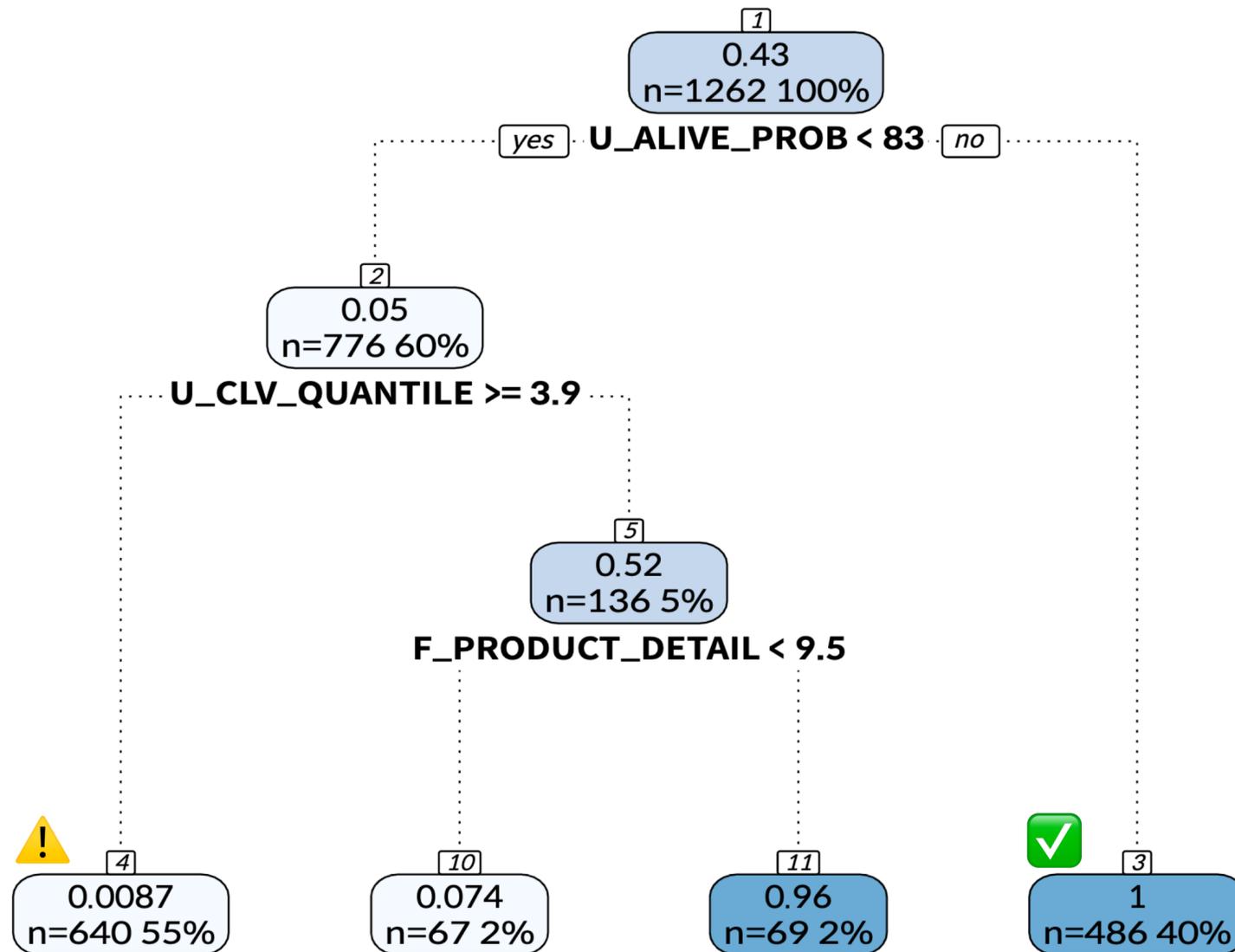
[CASE 4] 각 캠페인이 집중해야 하는 고객군 확인

데이터 및 모델링 프로세스



[CASE 4] 각 캠페인이 집중해야 하는 고객군 확인

(1) 쿠폰 발급 캠페인



✓ **효과가 좋은 고객군**

- 생존 확률 83% 이상

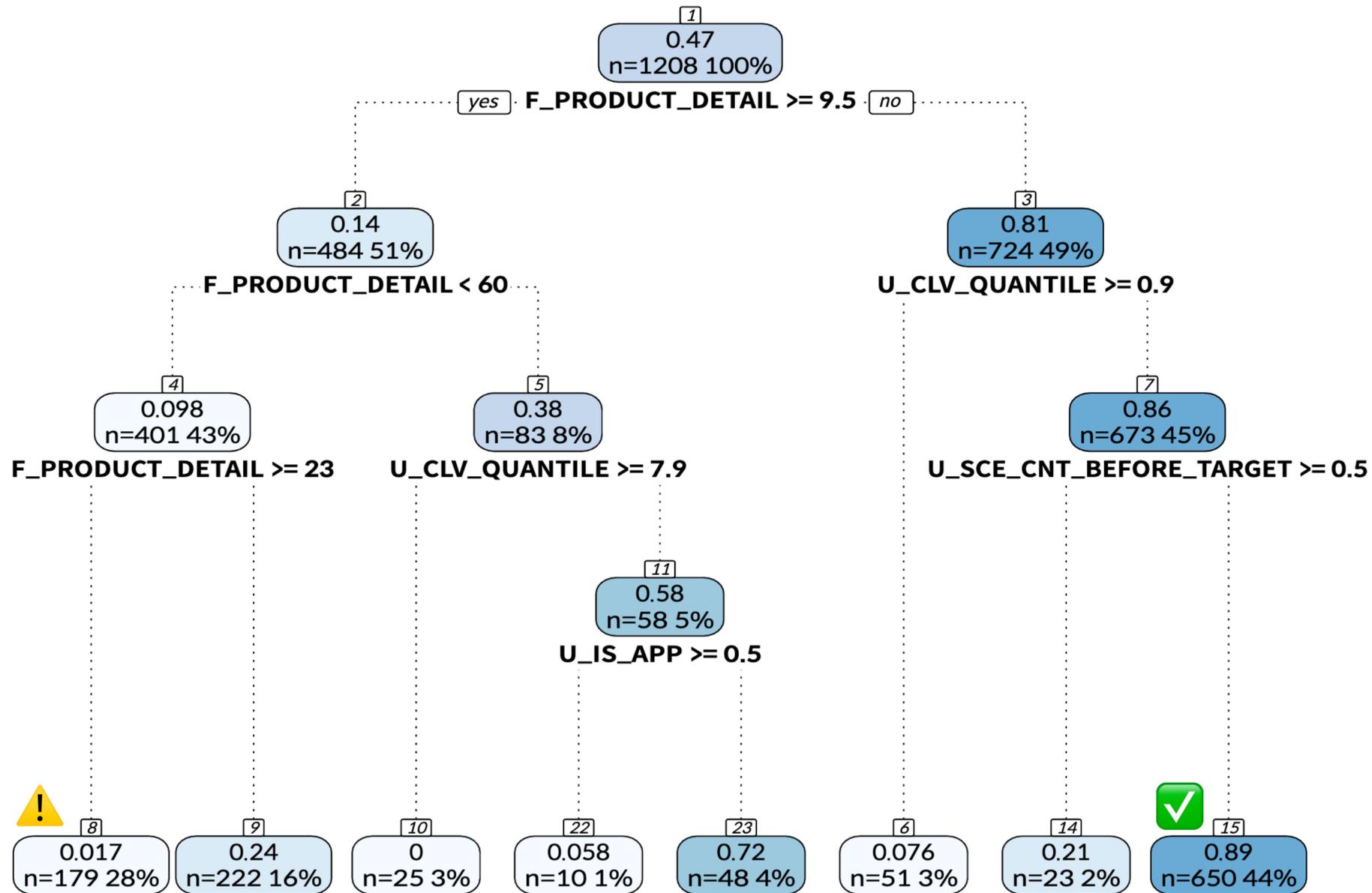
⚠ **효과가 좋지 않은 고객군**

- 생존 확률 83% 미만

- CLV가 하위 3.9%보다 높은 고객

[CASE 4] 각 캠페인이 집중해야 하는 고객군 확인

(2) 연관 상품 추천 캠페인



✓ **효과가 좋은 고객군**

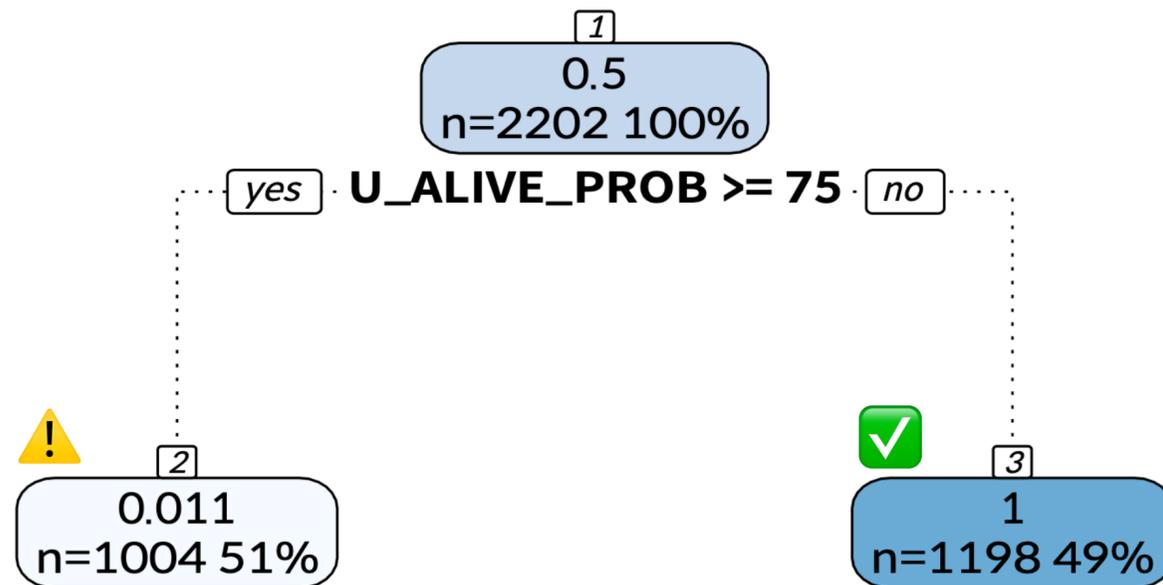
- 상품 상세페이지 9번 이하 조회
- CLV 하위 0.9%
- 24시간 이내에 다른 온사이트 캠페인에 노출되지 않은 경우

! **효과가 좋지 않은 고객군**

- 상품 상세페이지 23~60회 조회

[CASE 4] 각 캠페인이 집중해야 하는 고객군 확인

(3) 출석체크 혜택 알림 캠페인



✓ 효과가 좋은 고객군

- 생존 확률 75% 미만

! 효과가 좋지 않은 고객군

- 생존 확률 75% 이상

5. 마무리

개선이 필요한 부분

인과 모형의 결과물 평가하기

- 현실에 존재하지 않는 상황을 예측하다보니 정답셋이 존재하지 않는다.
- A/B 테스트할 때 인과 모형과 함께 비교해보기.

데이터를 남길 때부터 인과 모형을 염두에 두기

- Counterfactual Logging

어디까지 자동화할 수 있을까?

인과 모형이 중요한 이유

비즈니스 환경에서 올바른 의사 결정을 위해 성과/효과를 측정해야 한다.

- 모든 것을 "측정"할 수는 없다.
- 때로는 "추정"해야 한다.

일반적인 예측 모형으로는 인과 관계와 관련된 문제를 해결하기 어렵다.

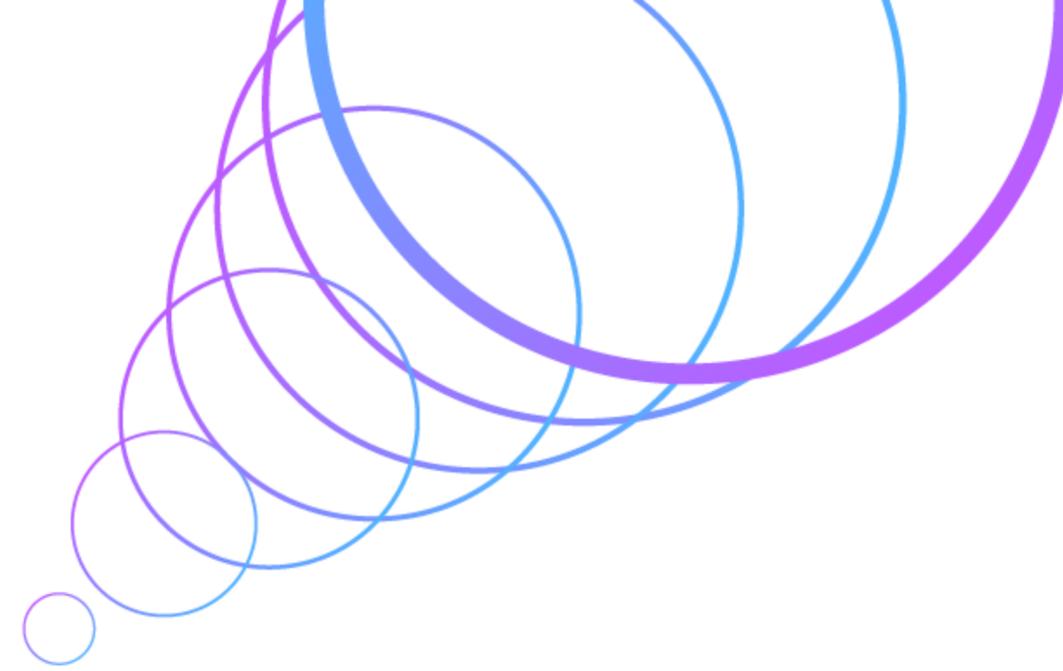
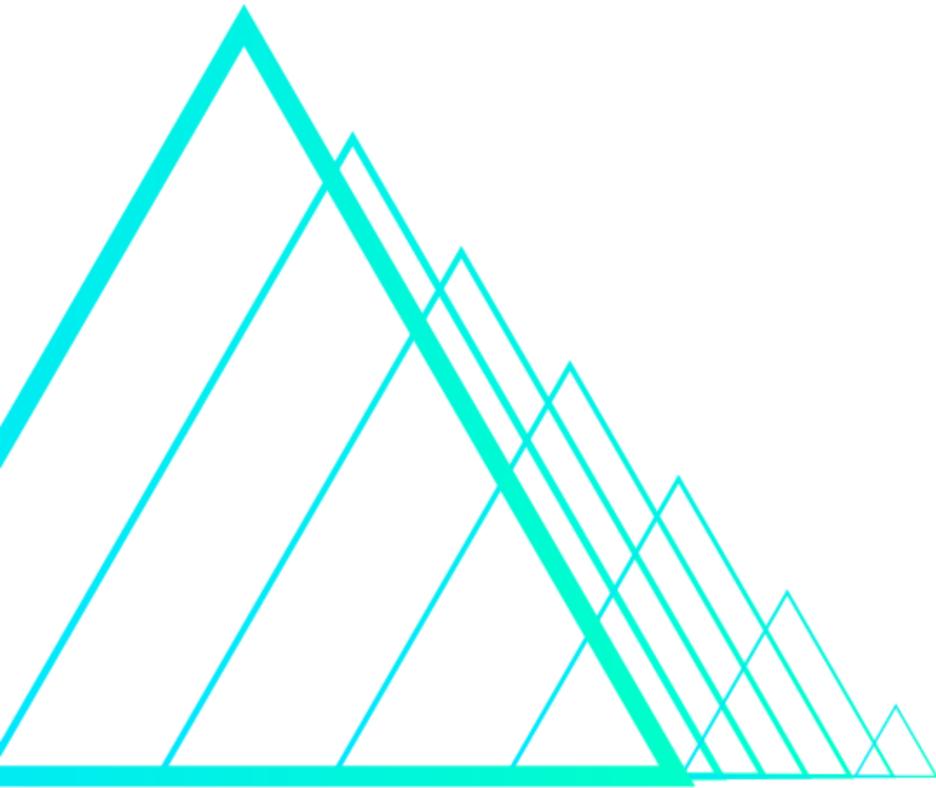
- 예측 모형 : 구매를 많이 하는 사람을 찾아낸다.
- 인과 모형 : 이벤트를 하면 더 많이 구매할 사람을 찾아낸다.

인과 모형으로 데이터 기반의 의사결정 프로세스를 고도화시킬 수 있다.

- 실험 : 시행착오를 겪으며 의사 결정을 객관적으로 할 수 있는 방법을 제공한다.
- 인과 모형 : What if ?

데이터라이즈는 **이커머스를 위한 올인원 그로스 솔루션**을 만듭니다.
데이터로 쏘아올릴 🚀 로켓에 함께 탑승하실 분을 기다립니다.

<https://team.datarize.ai>



Thank You

