

# 『Marketing Science』

---

Week 2 : 마케팅에서의 인과추론

2025 가짜연구소 인과추론팀 Conference




## Meetup for Product & Data

25.03.29(토) 오후 1시 30분  
한빛미디어 리더스홀

Product / Data / Decision Making에  
관심이 있는 실무자라면 참여 가능!

**일시**  
2025.03.29 (토) 오후 1시 30분 ~ 5시 30분

**장소**  
서울 서대문구 연희로2길 62 한빛미디어 사옥

**참가비**  
참가비가 없는 무료 행사!

**신청방법**  
(1) 인과추론팀 구성원 / 24년 행사 참여자라면? Eventus 바로 등록!  
(2) 인과추론팀이 처음이라면? 가짜연구소 Discord #인과추론과-실험 채널에 이름과 소속이 담긴 인사말 / 포스트 작성 후, Eventus 등록

**프로그램**  
(1) 워크샵 - 알고싶다! 인과추론팀 & 후원사  
(2) 세미나 - 알고싶다! 데이터의 모든 것  
(마케팅 분석 / 데이터 엔지니어링 / 인과추론과 머신러닝 / 임상통계)  
(3) 네트워킹 : MBTI가 I인 사람도 할 수 있다!

행사 소개 페이지

주관 : 가짜연구소 인과추론팀  
후원 : 한빛미디어




2025 가짜연구소 인과추론팀 Conference




## Meetup for Product & Data

25.03.29(토) 오후 1시 30분  
한빛미디어 리더스홀

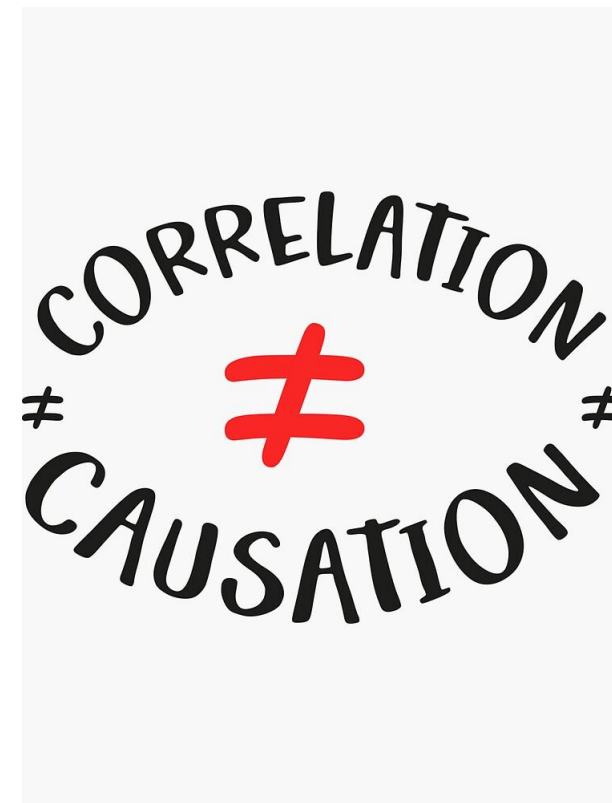
시간	컨텐츠	연사자
13:30 ~ 13:40	행사 등록 & 아이스브레이킹	
13:40 ~ 13:55	행사 Keynote	신진수 (크래프톤)
13:55 ~ 14:10	후원사 소개 세션	한빛미디어
14:15 ~ 14:30	세미나 1 - Causal Inference with CausalML	박이삭 (하이브 IM)
14:30 ~ 14:45	세미나 2 - Marketing Data Analytics	권남택 (넷마블)
14:45 ~ 15:00	세미나 3 - Lang2SQL	이동욱 (WEMADE)
15:00 ~ 15:15	세미나 4 - Adaptive Designs and Master Protocols	박상호 (성균관대)
15:20 ~ 16:20	"I"를 위한 네트워킹 (feat. 커피 시음회)	
16:20 ~ 17:20	대화 속의 성장 (#채용 #이직 #고민 #커리어)	
17:20 ~ 17:30	럭키 드로우 with 한빛미디어	

STAFF : 신진수, 정호재, 김지연, 박성수, 이동욱, 박상호, 권남택, 박이삭

1. 문제해결과 인과추론
2. 질문으로 시작하기
3. 인과추론의 근본적인 문제
4. 인과추론을 바라보는 관점
5. 올바른 질문하기


인과추론에 대해 설명하기 전에..

- 이름만 들어보면 어렵지 않습니다 (Causal Inference)
- 연관관계로부터 **인과관계**를 찾는 학문 / 방법론
- 연관관계가 인과관계와 다른 이유는? Bias (편향)



인과추론은 알고 보면 사람의 인지 과정과 비슷하다

## Three hierarchy in human cognition

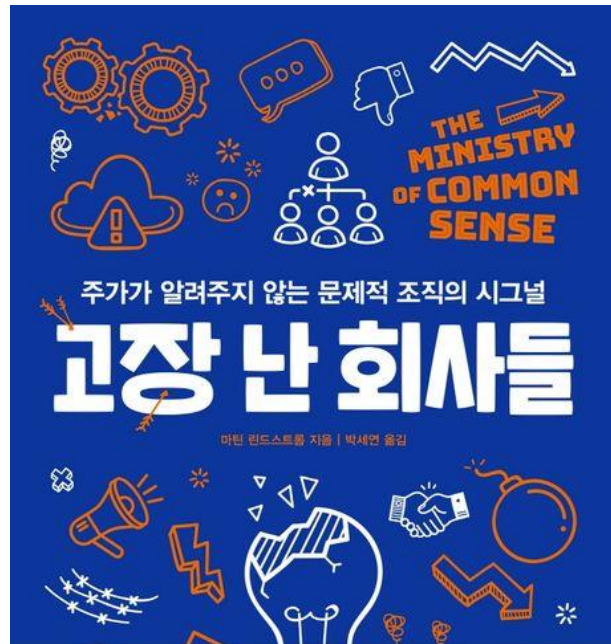
Human cognition	Task	Quantity	Question
L1 (Association)	Classification Regression Observational study	$P(y   x)$	What does the symptom tells about my headache?
L2 (Intervention)	Reinforcement Learning Randomized trial	$P(y   do(x))$	What if I took the aspirin, will my headache be cured?
L3 (Counterfactual)	 Retrospect	$P(y_x   x', y')$	Given that I didn't take the aspirin and didn't get cured, what if I did?

## 사회의 문제를 해결하기 위해서는





## 조직의 문제를 해결하기 위해서는



★★★★★  
〈타임〉 선정  
가장 영향력 있는  
인물 100

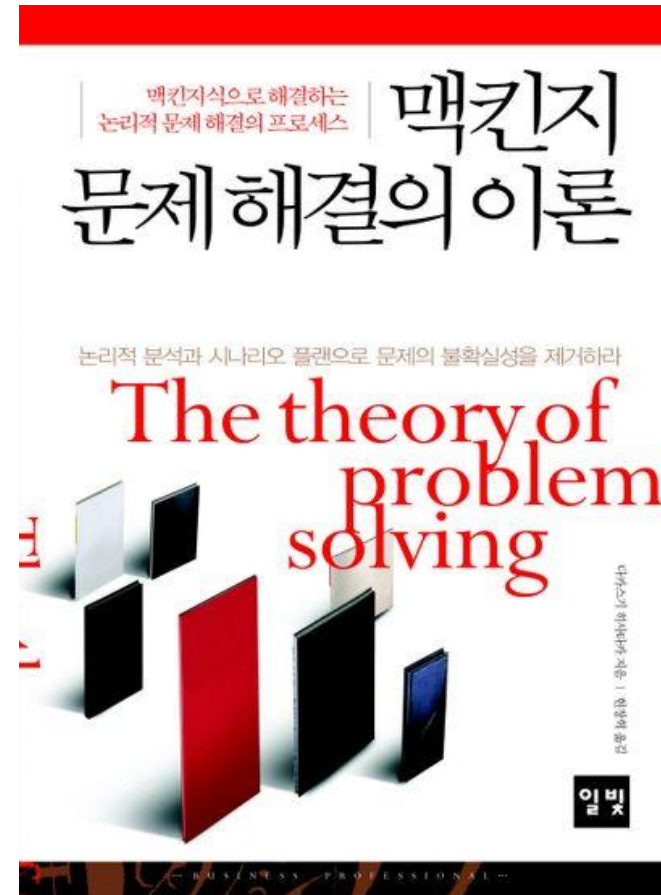
★★★★★  
〈실커스50〉 선정  
세계 최고의  
경영 사상가

★★★★★  
2021  
폴스트리트저널  
베스트셀러

★★★★★  
파이낸셜타임스  
이달의  
비즈니스북

“이 회사, 믿고 투자해도 괜찮은 걸까?”

어크로스



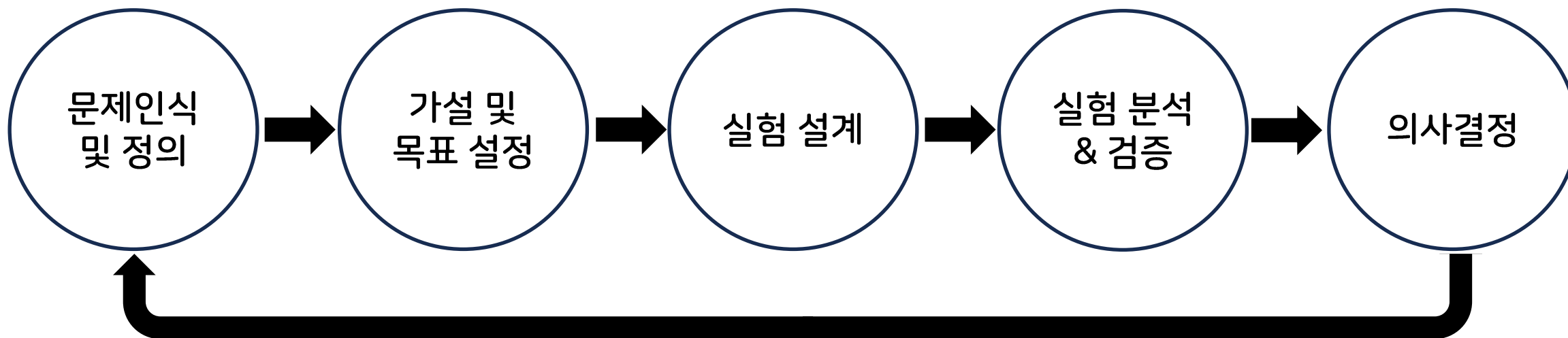




결국 인과추론은 'why'에 대한 답을 찾아나가는 과정



## 문제해결 Framework




## 간단한 마케팅 사례

- Situation : 기존 A가 아닌 B라는 Paid 캠페인을 집행했는데, CPI가 낮아짐  
(CPI = Cost Per Install)
- Decision : 바로 B 캠페인을 바꿔봅시다!

무엇이 이상한가요?

에이, 설마! 그럴 리가~



에이, 설마. 그럴 리가~

## 무엇이 문제일까?

(1) 질문이 이상하다 - Performance Marketing의 대전제 "**LTV**" > "**CAC**"  
(CAC = Customer Acquisition Cost)

(2) 의사결정 시, 비전 (중장기 지표)과 마일스톤 목표 (단기 지표)를 같이 바라봐야 한다

(3) 결과인 CPI에 영향을 줄 수 있는 요인은 다양하다 (ex. Media / Creative)  
→ 다양한 요인 : **Confounding Factors**

이로 인해, **개입 (Intervention)**에 따른 **효과 (Effect)**를 추정할 수 없다



데이터 분석가는 무엇을 해야 할까?

(1) Incrementality Test (A/B 테스트)가 이상적이다

(2) 할 수 없는 경우라면, 최대한 실험 설계를 이용하여 효과를 추정해볼 수 있다

→ **처치 (캠페인) 배정 매커니즘**에 기반하여 실험을 설계한다

→ **디자인 기반 인과추론**

## 이상적인 실험이란?

동일한 대상에게 시간을 거슬러, 다른 처치/개입을 할 수 있는 경우



## 인과추론의 근본적인 문제

(1) 시간을 돌릴 수 없으므로, **동일 대상에 대해 서로 다른 처치를 동시에 관측 불가**

(2) 즉, **반사실 결과**(Counterfactual Outcome)이 생긴다  
→ 반사실 결과 : 실험군이 처치 받지 않았을 때의 결과

(3) 우리가 할 수 있는 최선은? **대조군을 실험군의 반사실과 최대한 비슷하게 모델링**

(4) 인과추론에 대한 근본적인 문제가 있으니, 가정이 늘 수반됨

## 대조군을 실험군의 반사실과 비슷하게 만들기

↔ **편향**을 제거하기

$$\underbrace{E[Y \mid T=1] - E[Y \mid T=0]}_{\text{관측된 처치 효과}} = \underbrace{E[Y_1 - Y_0 \mid T=1]}_{\text{인과 효과}} + \underbrace{\left\{ E[Y_0 \mid T=1] - E[Y_0 \mid T=0] \right\}}_{\text{편향}}$$

## 편향 제거 후 결과

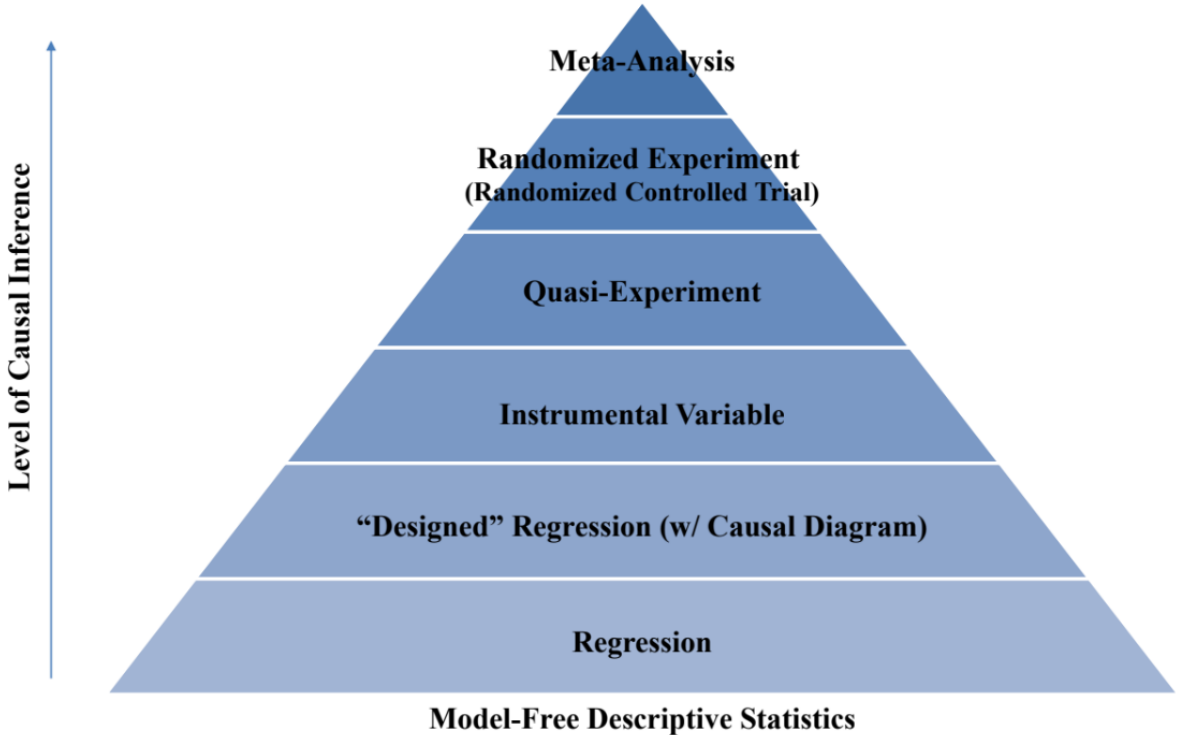
$$E[Y \mid T=1] - E[Y \mid T=0] = ATT = ATE = E[Y_1 - Y_0]$$

## 편향을 제거하는 쉬운 방법? 동전 던지기 (Random Assignment)

### Causal Hierarchy of Research Design for Causal Inference



인과추론의 데이터과학



Korea Summer Session on Causal Inference 2021

Session 2. Overview of Research Design for Causal Inference

2



왜 편향을 제거할까? 개입에 따른 인과효과를 식별하기 위해서

(1) 식별 : 관측 가능한 데이터에서 인과 추정량을 찾아내는 방법

(2) 식별을 하기 위해서는 : 데이터 생성 과정을 이해하고 있어야 함

(3) 연관관계와 인과관계가 다른 이유는 **편향 때문**이며, 이를 제거함으로써 우리가 원하는 인과 추정량을 **식별**하고, 이를 **추정**할 수 있다!

## 인과 추정량

- 개별 처치효과 (ITE, Individual Treatment Effect)

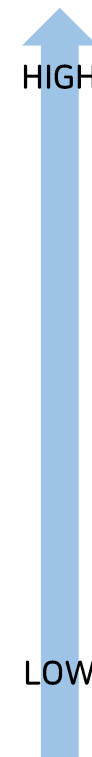
- 조건부 평균 처치효과 (CATE, Conditional Treatment Effect)

$$CATE = E[Y_{1i} - Y_{0i} \mid X = x]$$

- 평균 처치효과 (ATE, Average Treatment Effect)

$$ATE = E[Y_{1i} - Y_{0i}]$$

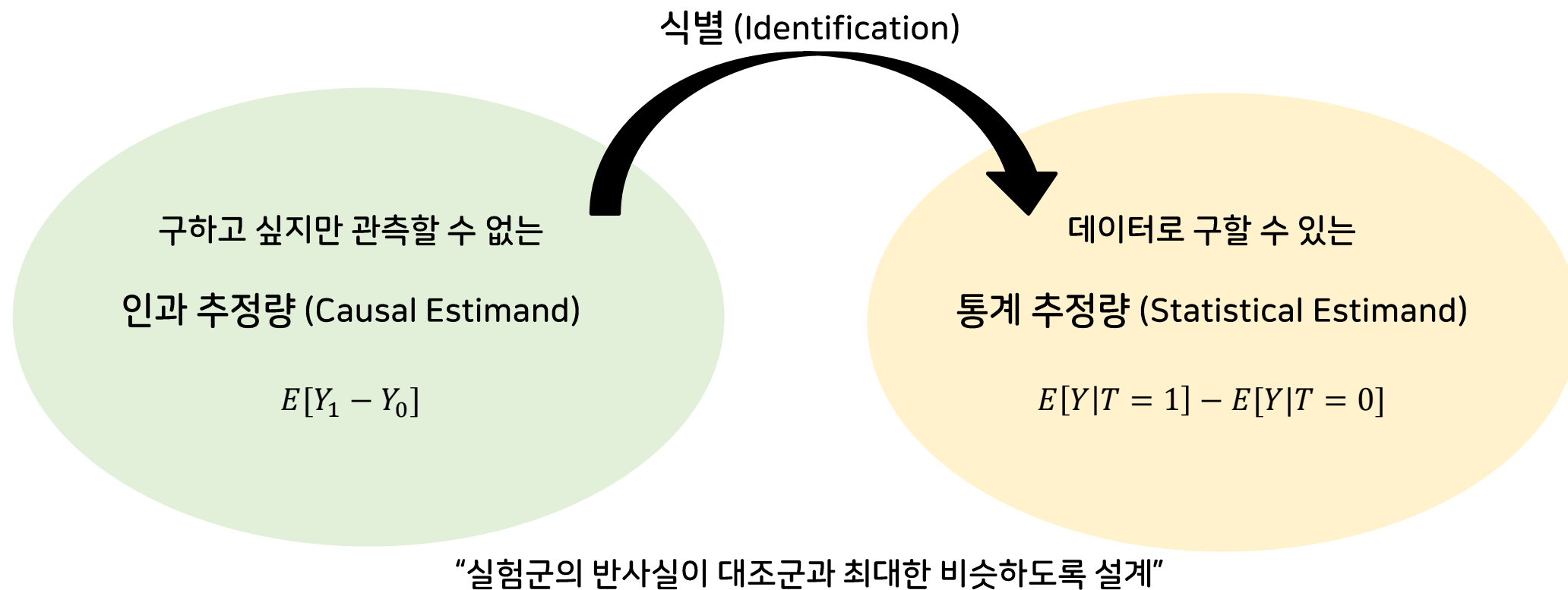
개인화



추정 비용/어려움



## Identification → Estimation



# 인과추론을 바라보는 관점

## 지난시간 남택님 베이지안 리뷰

### Link with Causal Inference



#### 베이지안과 인과추론

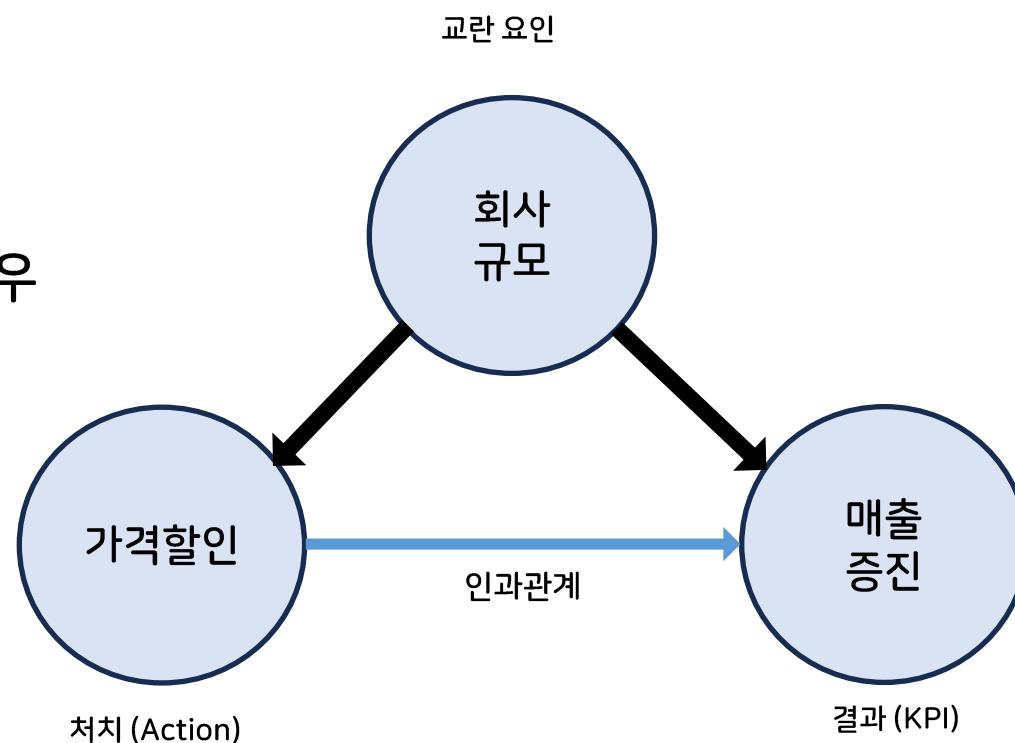
- 베이지안과 인과추론은 데이터를 바라보는 관점 (Framework)으로, **충돌되지 않음**
- 빈도주의 + 인과추론은 보통 잠재적 결과 (Potential Outcome; PO)에서만 이해하지만,  
베이지안 + 인과추론은 PO + Graph로 문제를 정의하는 경향이 있음 (베이지안 네트워크)
- 미국에는 베이지안 기반 인과추론 연구가 매우 활발
  - [Richard Hahn](#): 베이지안 트리앙상블 기반 인과추론 (Bayesian Causal Forest, ...)
  - 베이지안 트리앙상블은 인과추론계의 SOTA

## 인과추론에서 편향을 바라보는 3가지 관점

(1) **잠재적 결과** : 앞서 봤던 편향 (Missing Value)

(2) **구조적 인과 모형** : Backdoor Path

(3) **회귀분석** : 내생성 (Endogeneity)  
> 오차항과 설명변수간의 Correlation이 있는 경우





## 인과추론에서 편향을 바라보는 3가지 해석

- 올바른 실험 디자인을 통해, 인과 추정량을 식별하고 추정해야 함
  - 선택 편향이 없다
  - 비인과적으로 흐르는 뒷문 경로가 없다
  - 설명변수가 오차항과 외생적이다

## 베이지안 + 인과추론

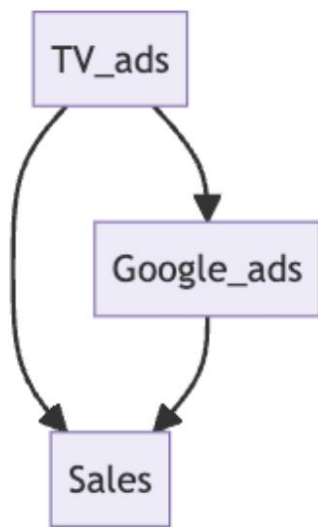


Figure 4.23– DAG for Google Ads versus TV ads

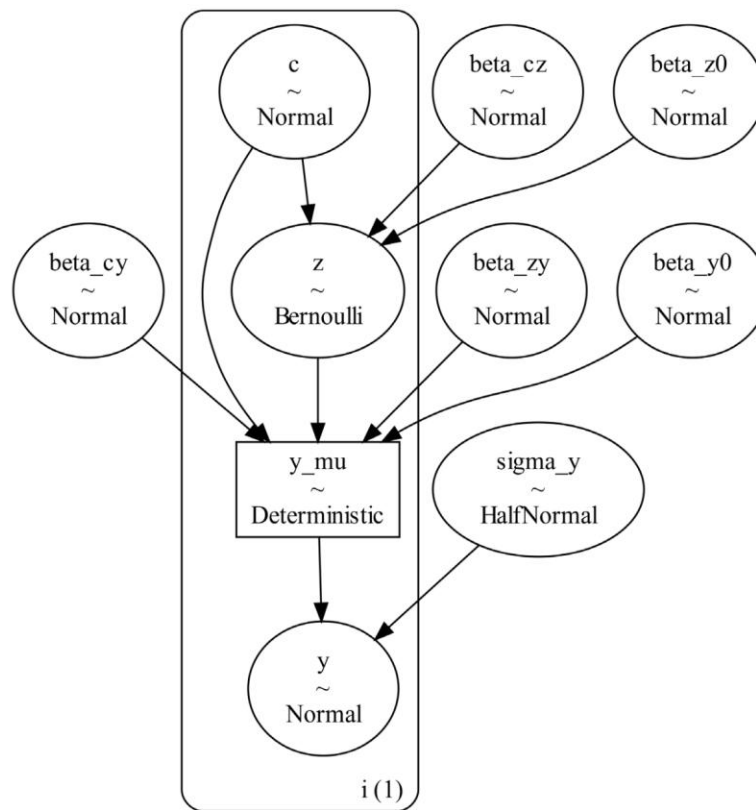


Figure 4.24 – Graphviz representation of the PyMC model

## 관점에 따른 식별과 추정

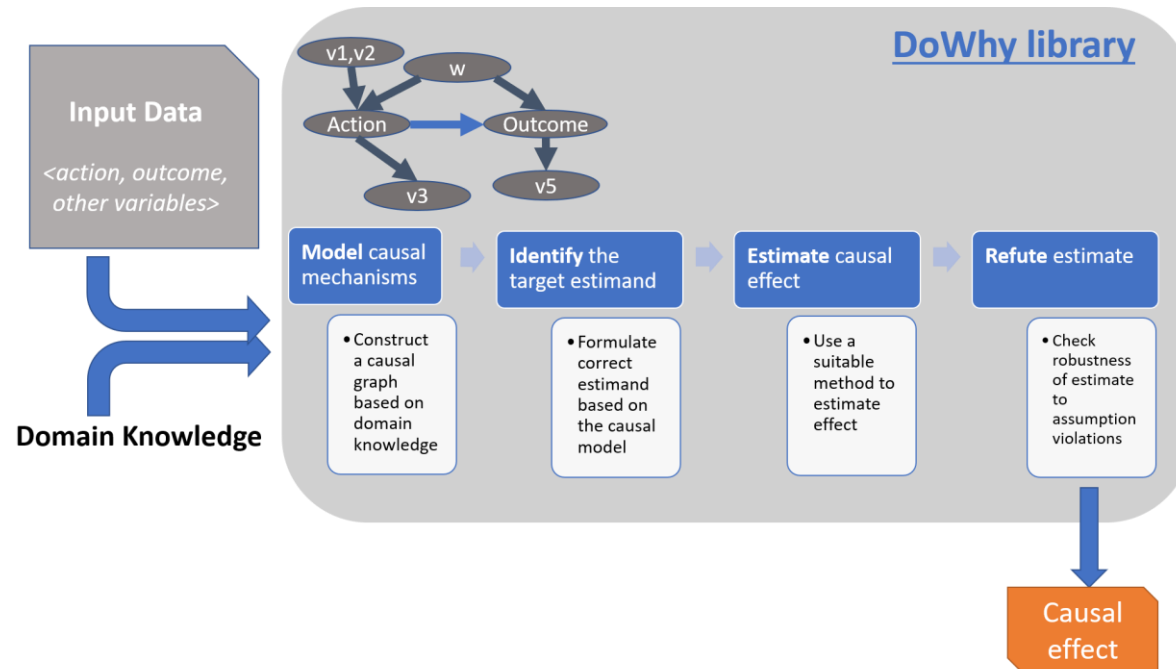
### Potential Outcome Framework vs. Structural Causal Model



인과추론의 데이터과학

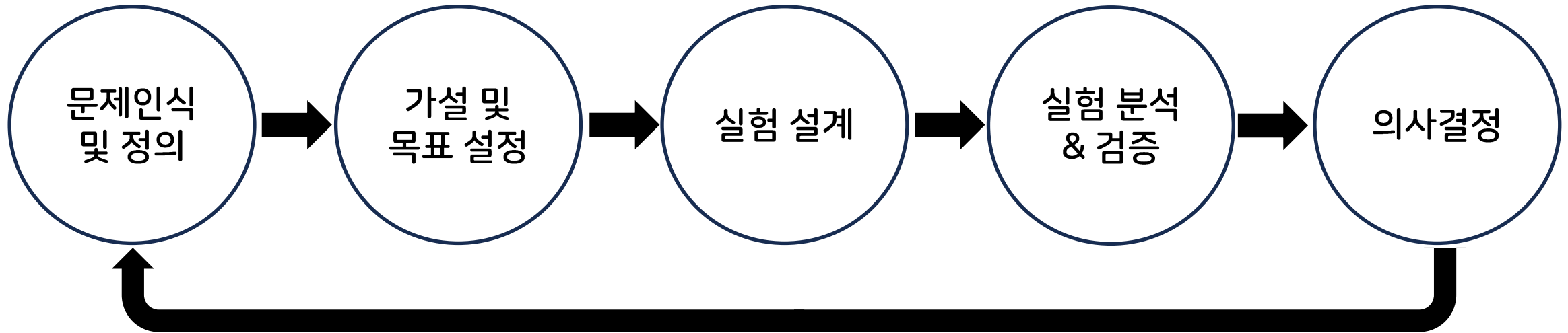
	Potential Outcome Framework	Structural Causal Model
Gold Standard	Random Assignment	
Causal Inference Using Observational Data		
(1) Identification (Is it possible to estimate a causal effect?)	Research Design (Quasi-Experiment / Natural Experiment)	Backdoor Criterion / do-Calculus (based on Bayesian Network/DAG)
(2) Estimation (How to estimate a causal effect using data?)	Statistical Methods (DID, RD, Matching, IV, SC, etc.)	Statistical/Computational Methods (IPW, Doubly Robust Estimators, Double ML, etc.)

## 인과 효과 추정을 위한 Framework



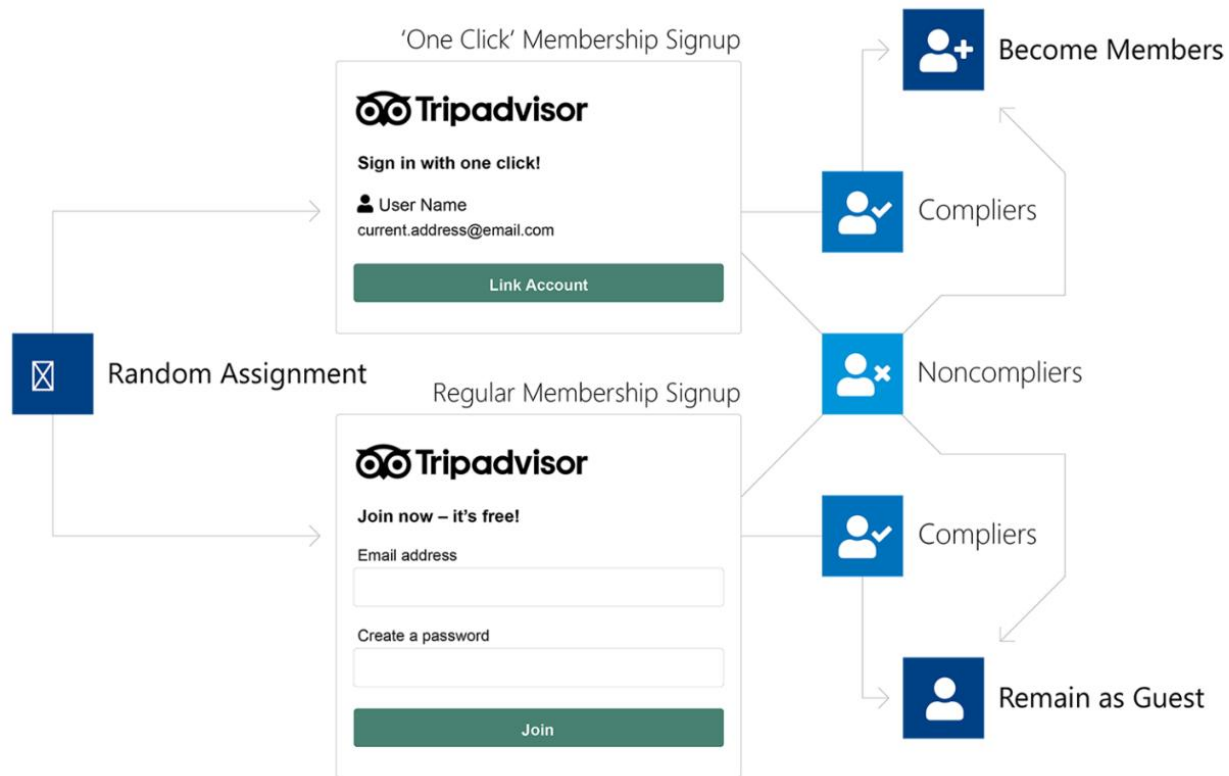
- DGP : 우리의 데이터는 어떻게 생성되었는가?
- 식별 : 해당 데이터 기반으로 인과 효과를 식별할 수 있는가?
- 추정 : 식별이 가능하다면, 어떻게 효과를 추정할 수 있는가?
- 검증 : 해당 인과효과는 통계적으로 / 비즈니스적으로 타당한 결과인가?

## 문제해결 Framework



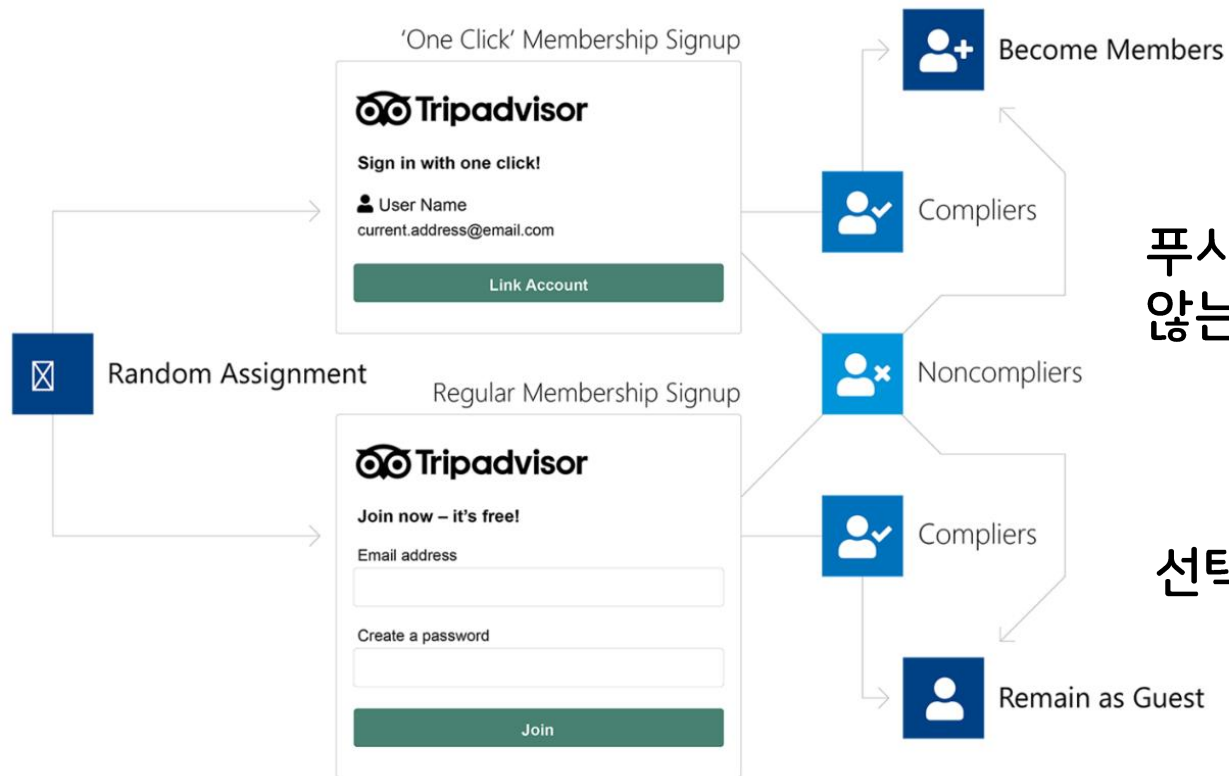


A/B 테스트가 **Gold Standard**인데... 이런 상황이라면? 만능일까?



# 인과추론을 바라보는 관점

- 푸시 발송 : 실험군과 대조군에게 랜덤하게 할당
- 푸시 오픈 : 푸시 발송이 랜덤하게 되었지만, 푸시 오픈은 유저가 선택 가능



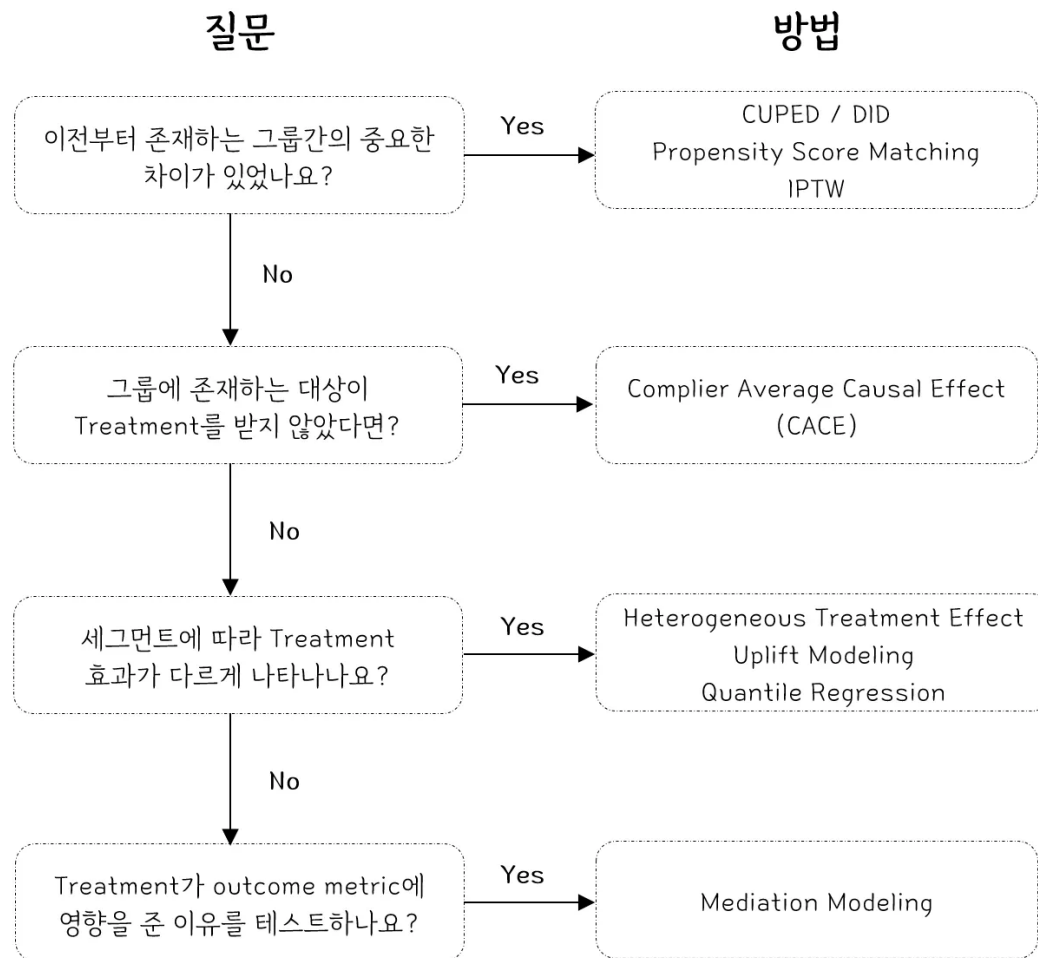
푸시 발송 (처치) 받은 모든 유저가 푸시 오픈하지 않는 **불응 (Non-compliance)** 문제가 발생

선택 편향을 제거할 수 있는 실험 디자인이 필요!

## 돌아와서...

- Situation : 기존 A가 아닌 B라는 Paid 캠페인을 집행했는데, CPI가 낮아짐  
(CPI = Cost Per Install)
- ~~Decision : 바로 B 캠페인을 바꿔봅시다!~~
- 분석가가 하는 일 : 그럼 B 캠페인의 효과를 보기 위해 다음과 같이 실험 설계 (데이터 생성 과정)을 만들어볼까요?
- 예시 : T1 기간에는 A만 비용을 증액하고, T2는 B만, T3는 A+B 동일 비용으로 증액해주실 수 있나요?

## 올바른 질문하기 - 데이터 생성 과정과 인과추론



## 모바일 마케팅과 인과추론

- 모바일 마케팅과 인과추론
  - (1) 모바일 프로덕트 출시 할 때, UA는 필수적인 상황
  - (2) UA 비용에 대한 데이터는 실험 데이터보다는 **관측 데이터**가 일반적
  - (3) 해당 관측 데이터 기반으로, 다음 Action을 위한 인사이트 축적이 필요

## 활용 사례

- ATT 동의율을 높이기 위한 실험 설계
- Retargeting 캠페인 효과 측정
- 도구변수와 개인화 메세징

인과추론은 문제 해결을 위해 올바른 질문을 할 수 있도록 안내

- 조직에서 문제를 해결하는 방법은 다양
- 가장 좋은 방법은 어느 누구의 리소스도 사용하지 않고 바로 해결할 수 있는 방법
- 만약 그렇지 않다면, 인과추론은 문제 해결을 위한 좋은 접근 방법 중 하나
- 인과추론은 단순 방법론 관점이 아닌, 문제 해결의 관점에서 활용되어야 함
- 인생은 늘 선택의 연속. 올바른 의사결정을 위해 인과적 사고는 도움되지 않을까요?



감사합니다.

---