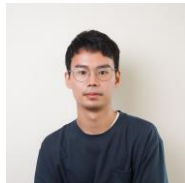


Casual하게 Causality 이해하기

신진수 | 크래프톤

발표자 소개

본캐



Data
Analyst

KRAFTON

Product



데이터야. 놀자 2024

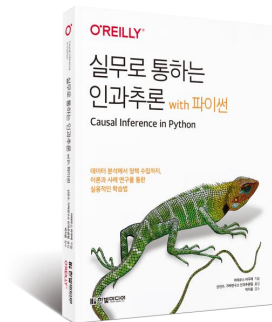
부캐

가짜연구소
인과추론팀



Causal-Lab

번역가



인과추론팀

신진수 (크래프톤)

실무로 통하는 인과추론 특강

신진수 (크래프톤)

- 김소희 (TVING)
- 김지연 (EA Korea)
- 박병수 (넥슨코리아)
- 박시온 (넥슨코리아)
- 박이삭 (하이브IM)
- 임종언 (넥슨코리아)
- 유정현 (넥슨코리아)
- 최은희 (넥슨코리아)
- 김성수 (LG CNS)
- 김준영 (컬럼비아대)
- 박상우 (한국은행)
- 박상호 (성균관대)
- 방태모 (Gmarket)
- 정호재 (롯데캐피탈)
- 조인서 (KOICA)

OCE 연구자로 거듭나기

방태모 (Gmarket)

- 문정하
- 박혜지 (CLS)
- 양유승 (우아한형제들)
- 유연승 (올빅넷)
- 이재호 (라인플러스)
- 조동민 (넥슨코리아)
- 지서영 (넥슨코리아)
- 황영산 (프라이데이즈랩)
- 황의림 (더블다운인터랙티브)

Smart Causal Inference

박상호 (성균관대)

- 김상돈
- 김혜민 (111 퍼센트)
- 안시형 (아우름플래닛)
- 이명석 (텔레픽스)
- 최의관 (하나은행)
- 하소희 (OP.GG)

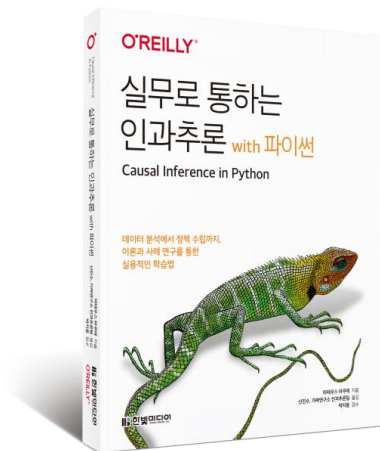
인과추론과 실무

김성수 (LG CNS)

- 김지연 (EA)
- 박수영 (크립토헤프)
- 손영주
- 손지영 (쿠팡)
- 손지우 (LG CNS)
- 심소현 (Imweb)
- 유도영 (SKT)
- 이은경 (온택트헬스)
- 최지환 (교보문고)

한국어 자료가 많지 않은 인과추론과 온라인 통제실험에 대한

접근성을 높일 수 있도록 기여



24.05.25(토) 오후 2시 OP.GG사육

발표를 하게 된 배경은요!



궤도 & 김상욱 교수님의 공통점?



대중에게 어려울 수 있는 학문인 과학 학습의

접근성을 높일 수 있도록 기여



(1) 너두 할 수 있어

O'REILLY®

실무로 통하는 인과추론 with 파이썬

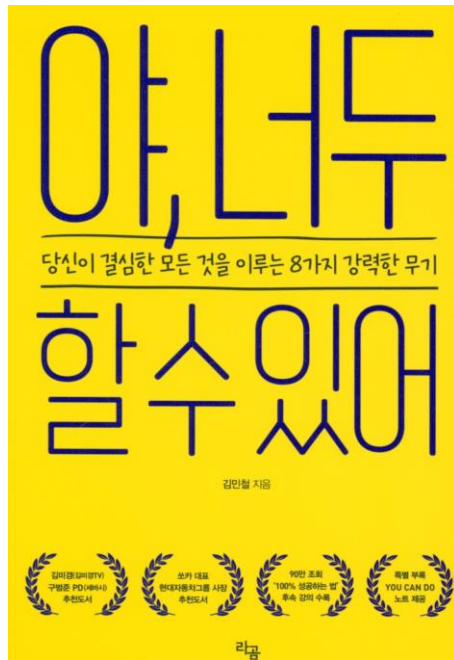
Causal Inference in Python

데이터 분석에서 정책 수립까지,
이론과 사례 연구를 통한
실용적인 학습법



인벡미디어

마태우스 파무레 지음
신진수, 가짜연구소 인과추론팀 옮김
박지용 감수



(2) 어려운 인과추론

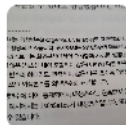
전체 리뷰 구매 리뷰

좋아요 순 ▼

종이책 ar***** | 2024.03.30 | 신고/차단

🌟🌟🌟🌟 / “ 최고예요

진심으로 추천하는 책입니다. 오렐리라서 믿을 수 있고, 번역가의 열정이 더해져서 정말 좋은 내용으로 잘 쓰여진 책입니다. 다만, 저는 리뷰어들이 이리 많을 줄 몰랐어요. 이유는 생각보다 내용이 처음부터 엄청 어려웠거든요. 오랜 시간 빅데이터 분석을 해온 저에게 쉽지 않은 책이었기에 저 혼자만 쓸 줄 알았어요 ㅎㅎㅎㅎ...



펼치기 ▼

👍 1 💬 답글 0

종이책 ar***** | 2024.03.27 | 신고/차단

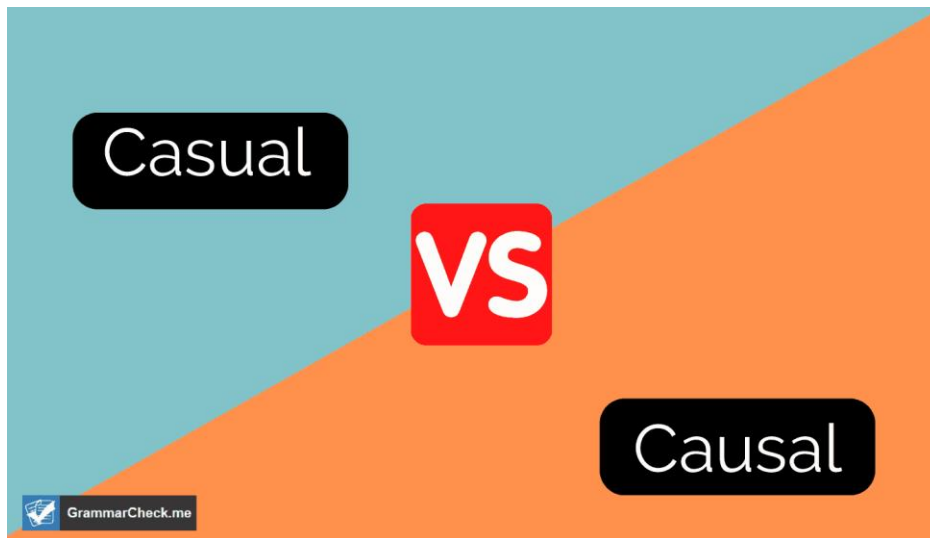
🌟🌟🌟🌟 / “ 최고예요

결코 쉽지 않아요. 인과관계와 연관관계를 시작으로 조금은 제너럴한 빅데이터 방법들이 나오기도 하지만, 확실히 기본 지식은 있어야 보기 쉽습니다.

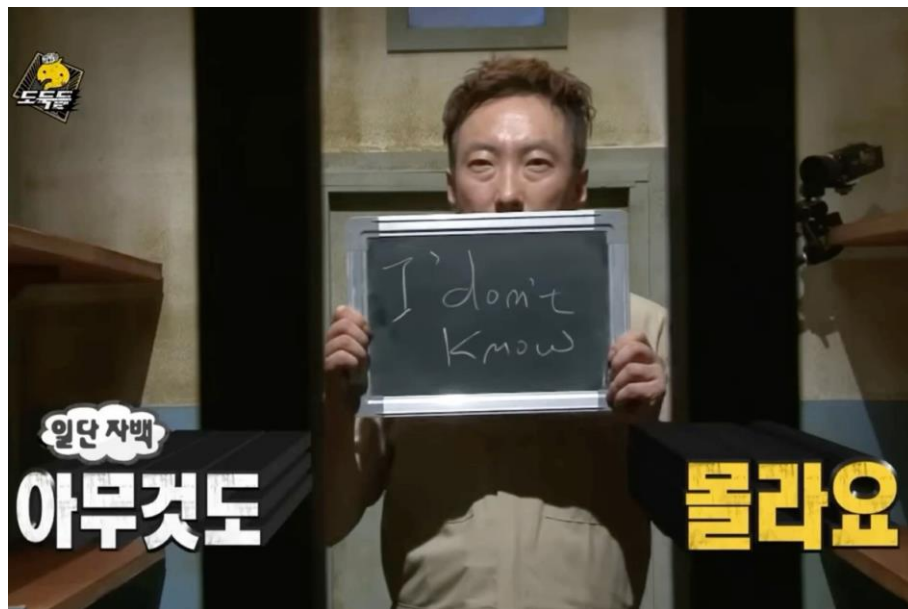


Casual하게 Causality 이해하기

Casual vs Causal



인과추론에 대해 아무것도 모르셔도 됩니다!



부탁드리고 싶은 점



어렵고 실무적인 부분을 더 학습하고 싶다면

특강 소개

01) 인과추론 기본서 출간
02) 인과추론 실무 방법 및 2024 공표 자료 발표
03) 인과추론 활용하기

실무로 통하는 인과추론 with 파이썬

인과추론의 데이터과학
동영상 5개 조회수 214회 4일 전에 업데이트됨

모두 재생 | 서플

1 [실무로 통하는 인과추론] 오리엔테이션
인과추론의 데이터과학 · 조회수 2천회 · 3주 전
... 지금 재생 중

2 [실무로 통하는 인과추론] 1-1. 인과추론 기초
인과추론의 데이터과학 · 조회수 626회 · 10일 전
28:55

3 [실무로 통하는 인과추론] 1-2. 인과추론의 목적
인과추론의 데이터과학 · 조회수 212회 · 10일 전
16:17

4 [실무로 통하는 인과추론] 1-3. 가격할인과 인과추론
인과추론의 데이터과학 · 조회수 334회 · 10일 전
38:56

5 [실무로 통하는 인과추론] 2-1. 온라인 통제 실험 소개 및 통계 분석 기초
인과추론의 데이터과학 · 조회수 266회 · 4일 전
47:26

제 발표와 함께하시는 분들은

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

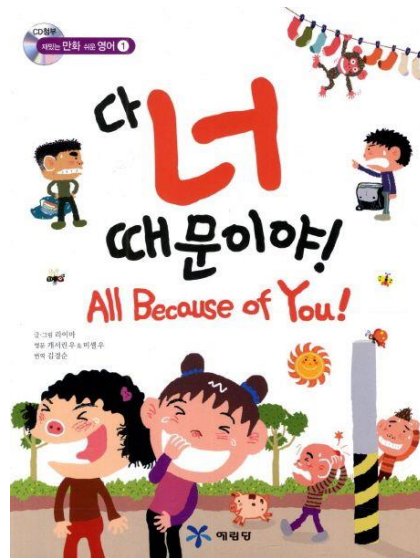
발표를 도와주시는 분들



Do you know ... 인과추론?



우리에게 너무나 익숙한 인과추론 (원인 찾기)



인과추론 조기교육(?)

6단원 일이 일어난 까닭

「원인과 결과를 안다」

1차시 3-1 국어(나) 170~173쪽

원인과 결과에 대해서 알아 볼까요?

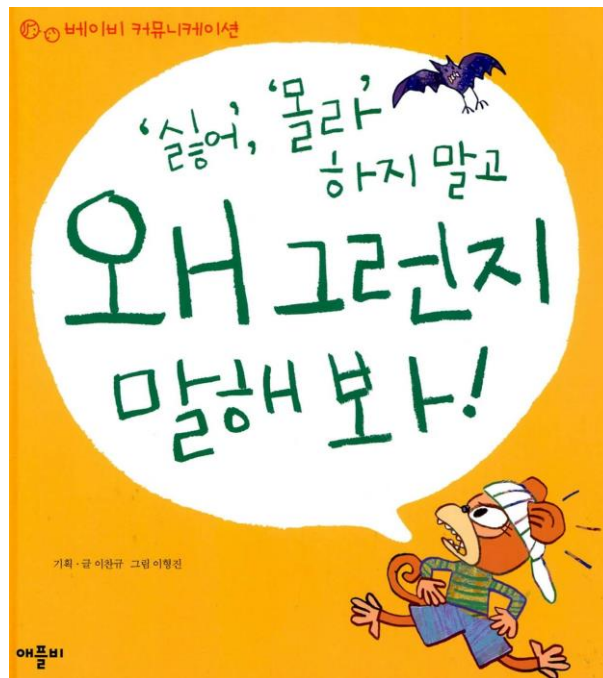
A cartoon illustration of a young boy with dark, spiky hair, wearing a blue and white striped shirt. He is shown from the chest up, looking upwards with a thoughtful expression, his hand near his chin. A speech bubble above him contains the text '원인과 결과에 대해서 알아 볼까요?'.

인과추론 공부하면 “가장 많이” 듣는 말

상관관계는 인과관계가 아니다.

“Correlation does not imply causation”

회사에서 “가장 많이” 받는 질문



인과관계는 중요하지만...



그럼에도, 우리가 인과추론을 배워야 하는 이유

- 조직 : 진수님! 혹시 상품 구성을 바꾸면, 구매 지표가 늘까요?
- 개인 : 매일 마스크 팩을 붙이고 자면, 피부가 좋아질까?

우리가 인과추론을 배워야 하는 이유

- 조직 : 진수님! 혹시 상품 구성을 바꾸면, 구매 지표가 늘까요?
처치 결과
- 개인 : 매일 마스크 팩을 붙이고 자면, 피부가 좋아질까?
처치 결과

(조직) 유저 경험 개선을 바탕으로 제품 성장에 기여



Engineering ~

Data / ML

Using Causal Inference to Improve the Uber User Experience



(개인) 어제보다 더 나은 오늘을 만들기 위해

어제보다
더 나은
오늘을
살자



생각한 대로
꿈꾸는 대로
노력한 대로

오늘만큼 가치 있는 것은 없다.
하루하루의 소소한 행복이 모여
삶은 변화하는 것이다. 오늘이리
는 날이 행복이 될지 불행이 될
지는 당신의 마음에 달려 있다.

오리슨스 마든 박진배 옮김

HAPPINESS
IN YOUR LIFE

그리고 ChatGPT says...

AI 시대에서 데이터 실무자에게 중요한 역량은?



AI 시대에서 데이터 실무자에게 중요한 역량은 다음과 같습니다:

1. 프로그래밍 및 코딩 기술:

- **Python, R, SQL:** 데이터 분석, 머신러닝 모델 구축, 데이터베이스 관리에 필수적인 언어들입니다.
- **버전 관리 시스템(Git):** 코드의 변경 사항을 체계적으로 관리하고 협업하는 데 중요합니다.

2. 데이터 분석 및 시각화:

- **통계적 분석:** 데이터의 패턴과 추세를 이해하고 유의미한 인사이트를 도출하는 능력입니다.
- **시각화 도구(Tableau, Power BI, matplotlib, seaborn):** 데이터를 시각적으로 표현하여 이해하기 쉽게 만드는 기술입니다.

3. 머신러닝 및 AI:

- **머신러닝 알고리즘 이해:** 회귀, 분류, 군집화, 시계열 분석 등의 알고리즘을 이해하고 적용할 수 있는 능력입니다.
- **딥러닝 프레임워크(TensorFlow, PyTorch):** 복잡한 데이터 모델링을 위한 딥러닝 기술입니다.

인과추론:

- **인과관계와 상관관계 구분:** 데이터를 분석할 때 단순히 상관관계를 찾는 것에서 나아가, 변수들 간의 인과관계를 명확히 이해하고 설명할 수 있는 능력입니다.
- **도구와 방법론:** 자연 실험, 무작위 대조 실험(RCT), 도구변수법, 이중차분법 등의 인과추론 기법을 활용하여 데이터에서 인과관계를 추론할 수 있는 능력입니다.
- **상황 이해:** 실제 비즈니스 문제에서 어떤 변수들이 인과관계에 영향을 미치는지 이해하고, 이를 기반으로 전략적 결정을 내릴 수 있는 능력입니다.

여러분은 인과추론에서 벗어날 수 없습니다



Key Questions

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

Summary

- 인과적 사고를 하는 것은 사람에게 너무나 자연스러운 일
- 본질적인 인과관계를 찾아, 개인은 어제보다 나은 성장을 이룩하고
- 조직(데이터 분석가)은 유저 경험 개선을 통해 제품 성장에 기여

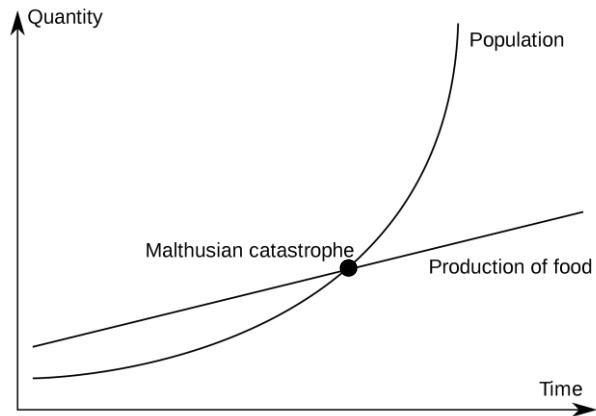
Key Questions

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

근본적인 문제를 이해해야

인과추론을 제대로 이해할 수 있음

타노스의 가설 검증



말서스 트랩 (Malthusian trap)



타노스 가설 : 인구를 절반으로 줄이면,

삶의 질(행복)을 개선할 수 있다!

처치
(Treatment)

타노스 가설 : 인구를 절반으로 줄이면,

삶의 질(행복)을 개선할 수 있다!

결과
(Outcome)



실험 대상 배정

- 실험군 : You were slain by Thanos, for the good of the Universe. (사망)
- 대조군 : You were spared by Thanos. (생존)

방금 우리는 무엇을 했을까요?

실험 대상을 무작위로 배정 (처치 배정 매커니즘)

Aka. A/B 테스트

실험군에 대한 잠재적 결과

- 사실적 결과 : 핑거 스냅으로 사망한 실험군
- 반사실적 결과 : 핑거 스냅에서 생존한 실험군

동일 대상에 대해 동시에

서로 다른 잠재적 결과를 관측할 수 있을까요?

선택 2014



핑거 스냅으로 실험군은 이미 사망



핑거 스냅으로 실험군은 이미 사망



살려주세요..

개별 실험 대상에 대한 처치 효과를 알 수 없음

인과추론의 근본적인 문제

(Fundamental Problem of Causal Inference)

인과추론의 근본적인 문제가 발생하는 이유



Summary

- 동일 대상에 대해 서로 다른 잠재적 결과를 관측할 수 없음
- 우리는 타임 스톤이 없기에 인과추론의 근본적인 문제가 항상 존재



우린 답을 찾을 것이다 늘 그랬듯이

<인셉션>, <다크나이트> 시리즈 감독 작품

인과추론의 근본적인 문제 접근하기

우리가 원하는 것 : 실험군의 반사실



우리가 가진 것 : 대조군



대조군을 실험군의 반사실과 최대한 비슷하게 디자인

실험군과 대조군이 비교 가능하면 됩니다!



VS



비교 가능하지 않다면?



VS

SAMSUNG

편향이 발생...

실험군 반사실 : 내가 원하는 머리



대조군 : 그 머리를 한 나



대조군을 실험군의 반사실과 최대한 비슷하도록 만들기



핑거 스냅으로

무작위로 처치를 배정해서

편향을 제거



A/B 테스트처럼 무작위 배정만

편향을 제거할 수 있을까?

선택 2014

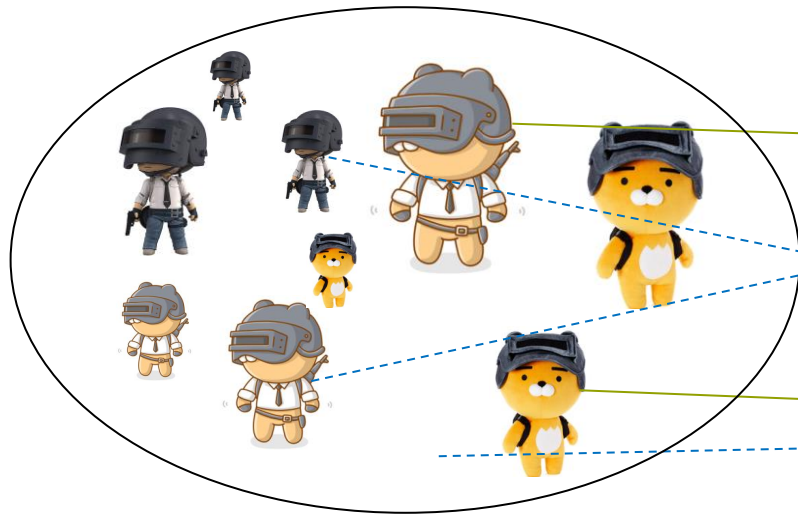


편향 제거를 위한 실험 디자인 - 매칭

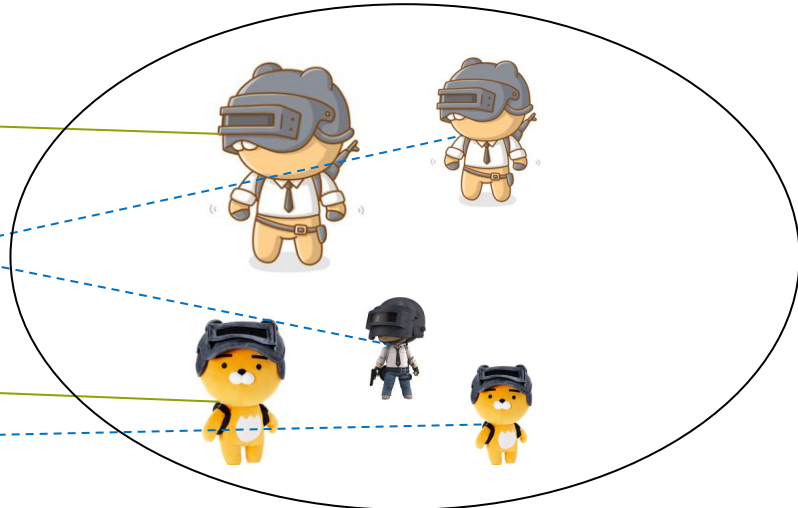
- 실험군과 대조군에서 유사한 대상끼리 짝을 지어 비교하는 방법
- 실험 대상의 “유사도”에 대한 기준에 따라 다양한 방법이 존재

편향 제거를 위한 실험 디자인 - 매칭

실험군



대조군



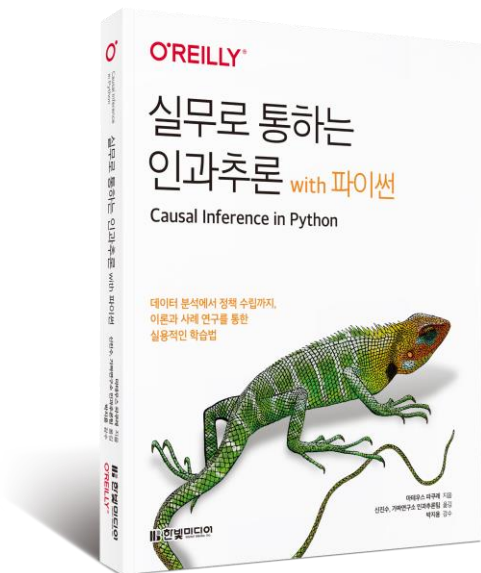
편향 제거를 위한 실험 디자인 - 통제집단합성법

- 대조군을 결합해서 **처치가 없을 때 실험군**과 비슷한 가상의 대조군 생성
- **패널데이터**에서 활용하는 인과추론 방법론

(광고)

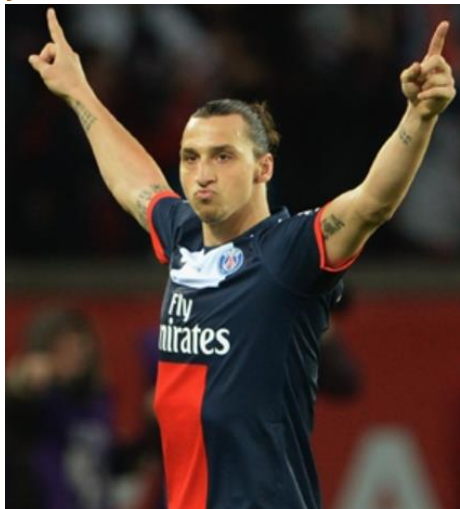
데이터야. 놀자 2024

다양한 인과추론 기반 디자인에 대해 알고 싶다면



Summary

실험군



대조군



얼굴을 바꿔도 즐라탄이 보이는 변신술(?)

인과추론 참 쉽죠?





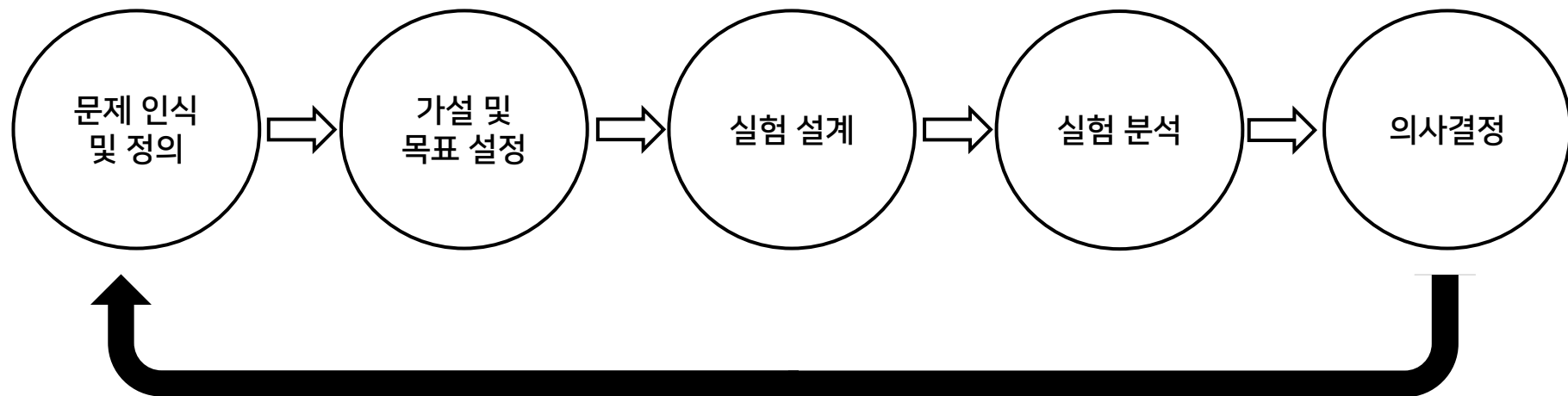
Key Questions

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

기안84님과 함께하는 인과추론



문제 해결을 위한 Cycle

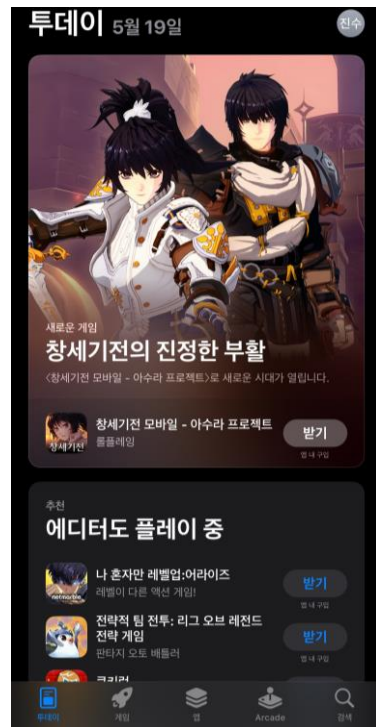


- 목표 : AOMG에서는 자체 앱 다운로드(전환/인스톨)을 늘리고 싶음
- 상황 : 그러나 마케팅에 투입할 예산이 거의 소진되어 없음

* 인과추론 이해를 돕기 위한 가상의 시나리오입니다.

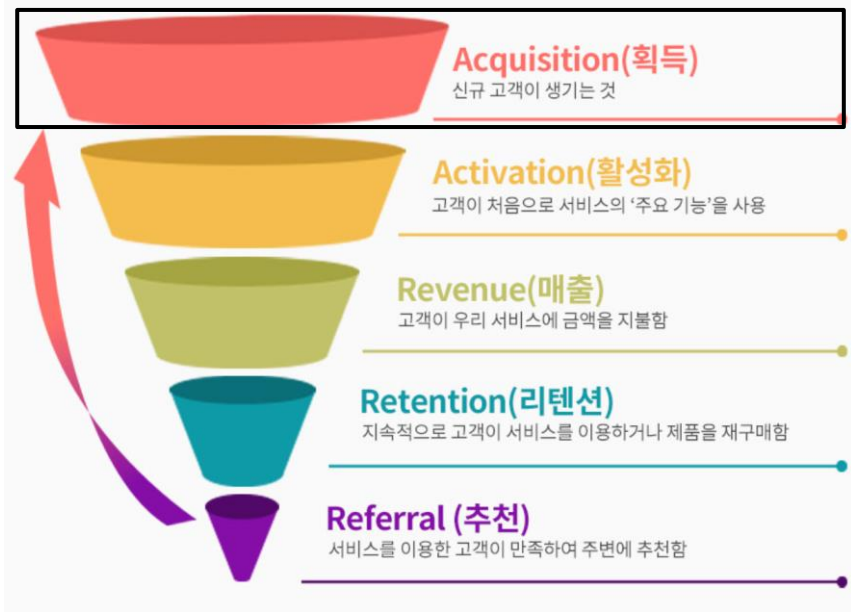
0) 시나리오

데이터야.놀자 2024



0) 시나리오

데이터야.놀자 2024



고객과의 첫 접점

0) 시나리오

데이터야. 놀자 2024



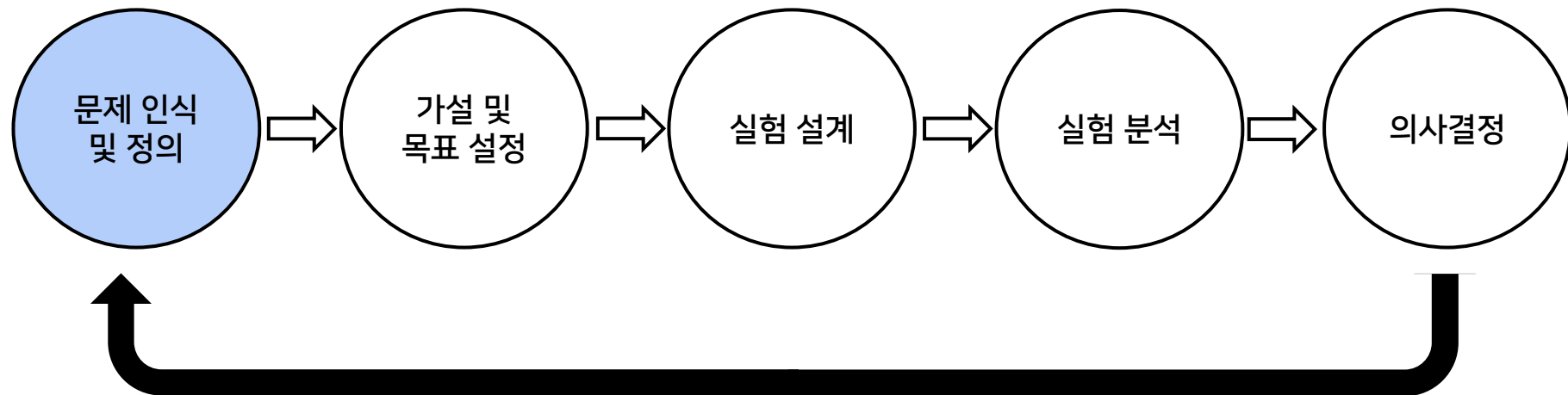
0) 시나리오

데이터야. 놀자 2024

AOMG 로고를 본 여러분의 느낌?

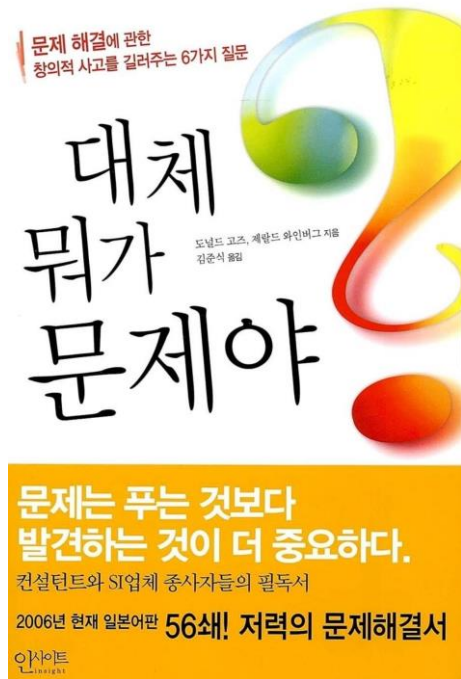


문제 해결을 위한 Cycle



1) 문제 인식 & 정의 - 지표 모니터링 & 호기심

데이터야. 놀자 2024



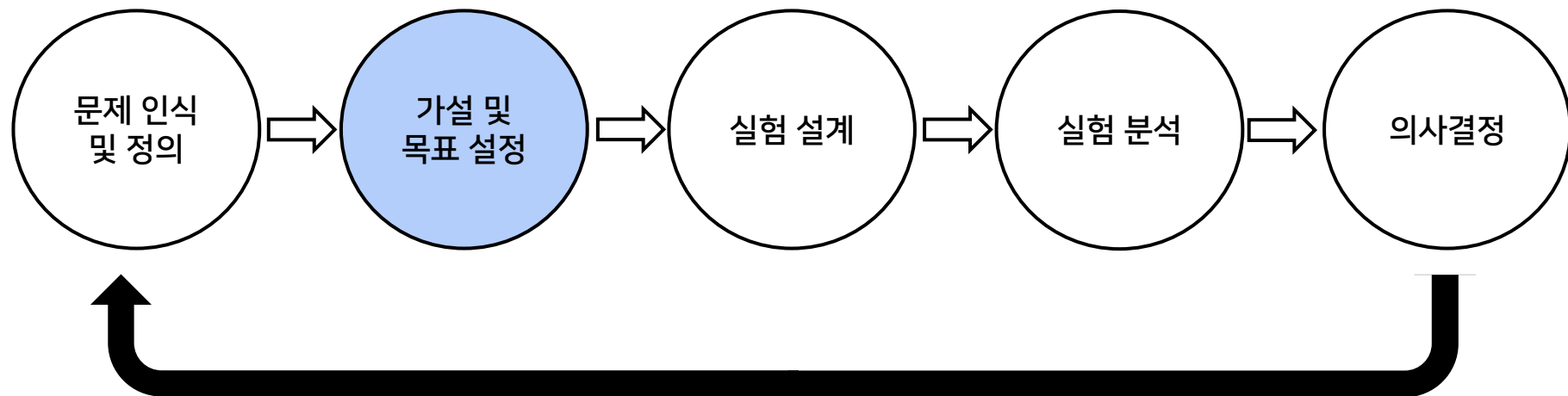
1) 문제 인식 & 정의 - 지표 모니터링 & 호기심

데이터야. 놀자 2024



- 기획자 : 기안84 (AOMG 디자이너)
- 문제 : 아싸, INTP에게는 위화감을 주므로 앱 다운로드에 부정적 영향
- 근거 : 경험, SNS 모니터링(데이터)

문제 해결을 위한 Cycle



AOMG 로고를 바꾸면, 유저가 앱을

더 많이 설치하지 않을까?

AOMG 로고를 바꾸면, 앱 스토어 방문 유저에 대한

스토어 전환율이 더 높아지지 않을까?

2) 가설 설정 - Actionable & Quantifiable

데이터야. 놀자 2024

처치 (개입)

실험 대상

AOMG 로고를 바꾸면, 앱 스토어 방문 유저에 대한

결과

스토어 전환율이 더 높아지지 않을까?

모든 MBTI를 만족시킬 수 있는 로고로 변경하면



인싸/아싸 모두의 이목이 집중되어



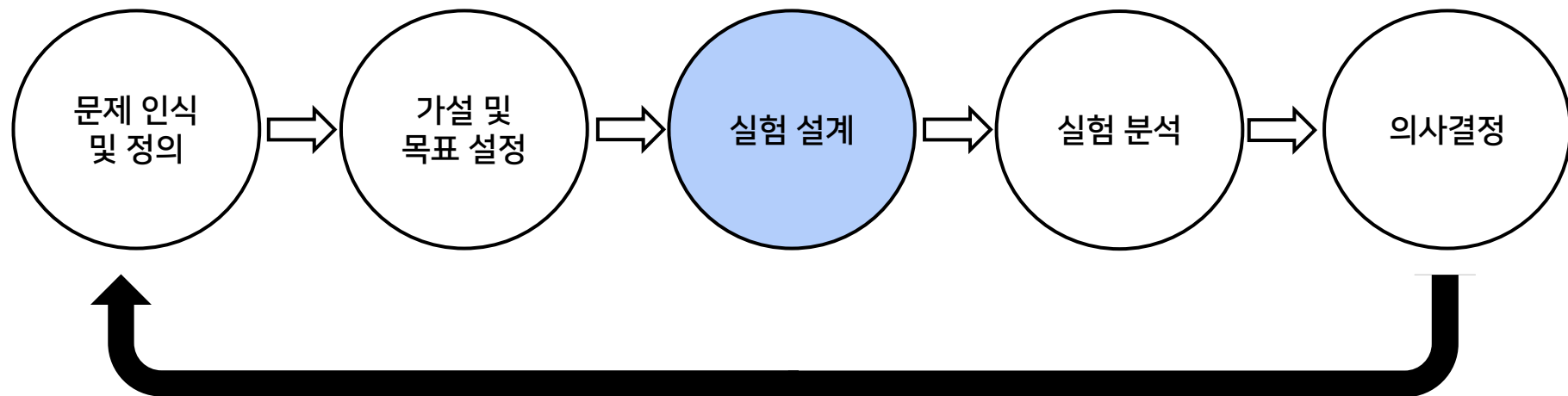
AOMG 스토어 전환율이 더 늘어날 것이다!

2) 가설 설정 - Actionable & Quantifiable

데이터야. 놀자 2024



문제 해결을 위한 Cycle



왜 / 누가 / 언제 / 어디서 / 무엇을 / 어떻게

3) 실험 설계 - 편향 제거

- 왜 : 실험/가설 검증의 목적
- 누가 : 실험 대상 (앱스토어 방문 유저)
- 언제 : 실험 기간 (1주간)
- 어디서 : 실험 장소 (앱스토어)
- 무엇을 : 처치 (로고 변경)
- 어떻게 : 처치 배정 매커니즘 (무작위 배정)

3) 실험 설계 - 편향 제거

데이터야.놀자 2024



3) 실험 설계 - 편향 제거

데이터야. 놀자 2024

실험군



대조군

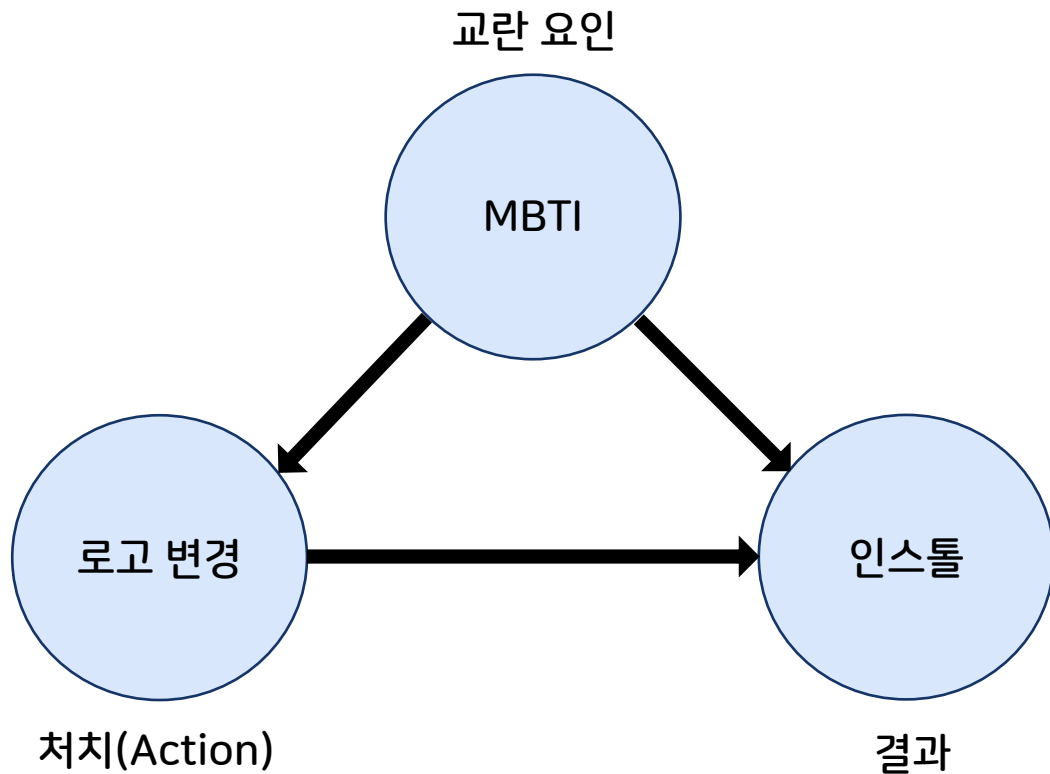


문제는 바로...



3) 실험 설계 - 편향 제거

데이터야.놀자 2024



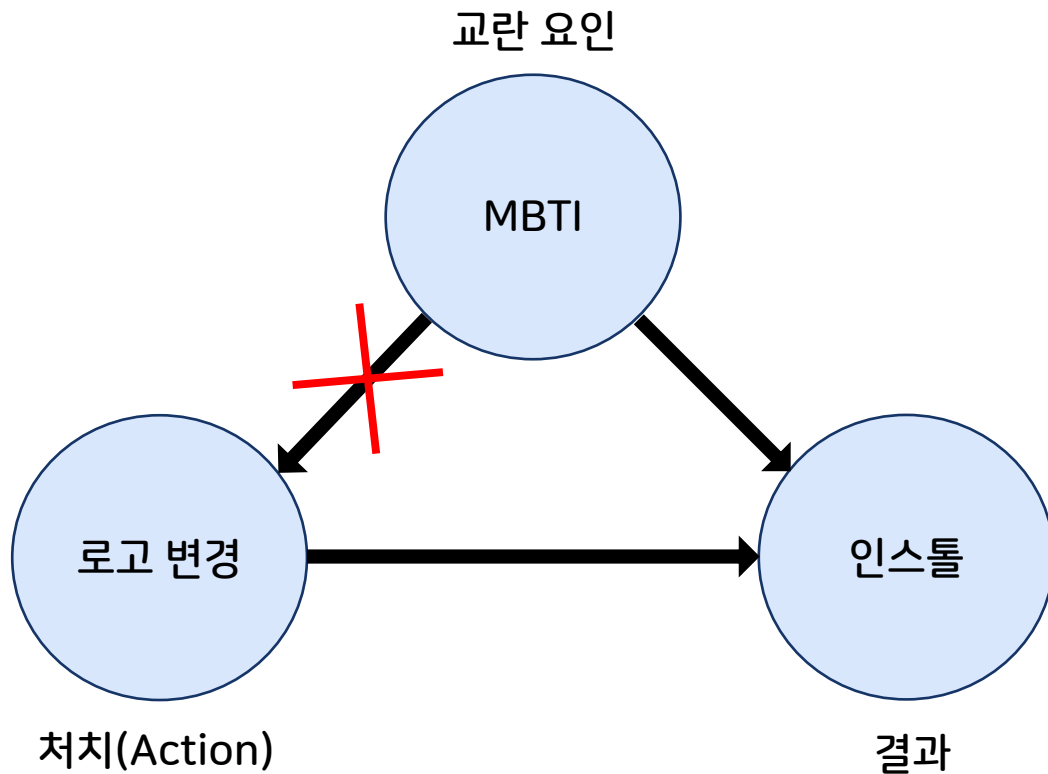
3) 실험 설계 - 편향 제거

데이터야. 놀자 2024

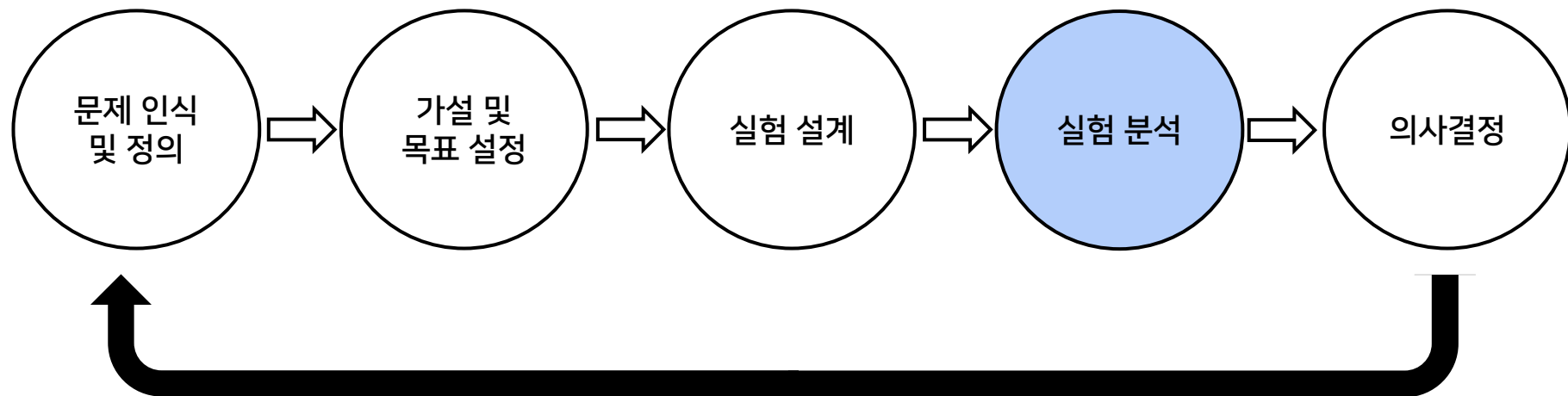


3) 실험 설계 - 편향 제거

데이터야.놀자 2024



문제 해결을 위한 Cycle



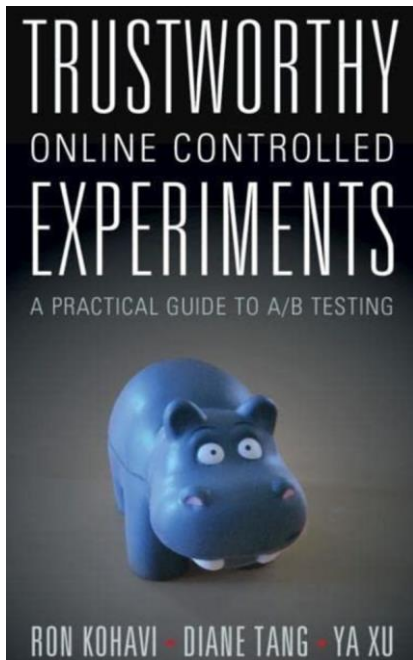
로고 변경에 따른 효과(처치 효과) 파악



+ 10.9%p

4) 실험 분석 - 결과의 신뢰성 & 리포트/시각화

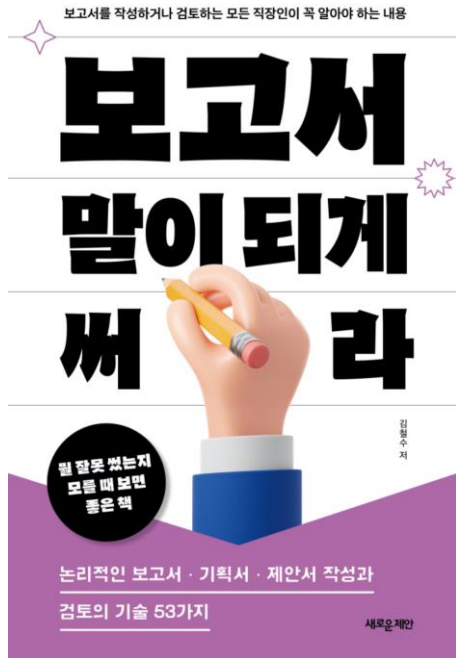
데이터야. 놀자 2024



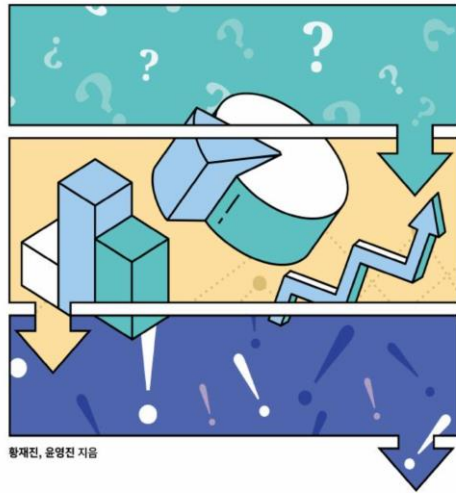
- 10.9%p 증가는 믿을 만한 수치인가?
- 무작위 배정은 잘 이뤄졌는가?
- MBTI 이외에 실험에 영향을 주는 요소?

4) 실험 분석 - 결과의 신뢰성 & 리포트/시각화

데이터야. 놀자 2024



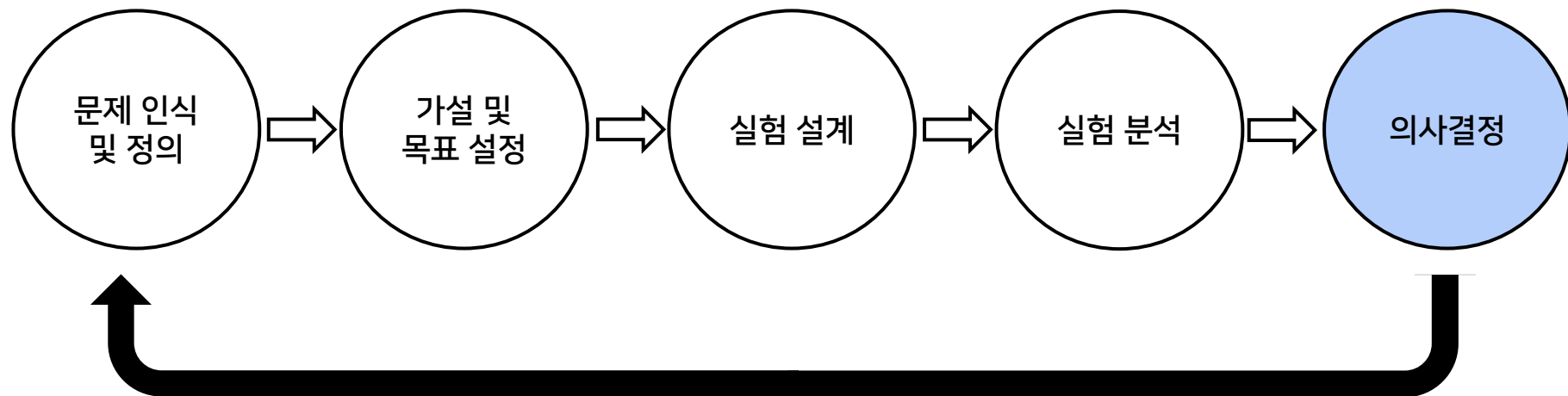
막대 차트부터 대시보드까지 태블로로 실습하며 배우는 인사이트 도출법



사례 분석으로 배우는
데이터 시각화

한빛미디어

문제 해결을 위한 Cycle



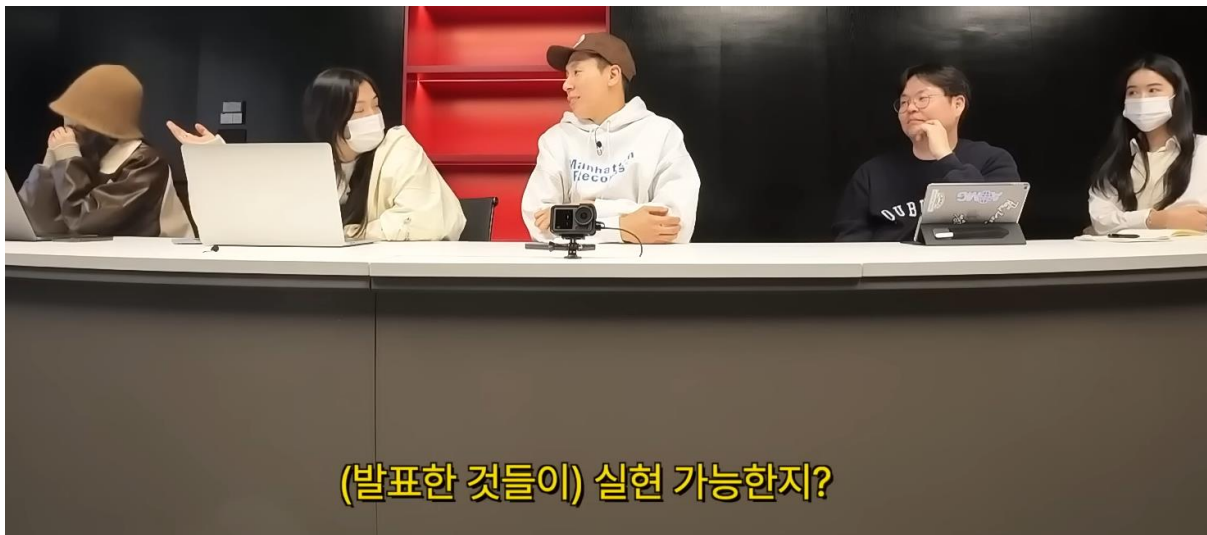
5) 의사결정

데이터야. 놀자 2024



5) 의사결정

데이터야. 놀자 2024



5) 의사결정

데이터야. 놀자 2024



Key Questions

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

Summary

- 인과추론은 데이터로 조직의 문제를 해결하기 위한 여러 방법 중 하나
- 우리 제품의 문제는 무엇인가? 그 문제의 크기는 어떠한가?
- 인과추론은 해당 문제를 효율적으로 풀기 위한 방법인가?



인과추론과 문제 해결 과정에서 꼭 필요한 2가지?

도메인 지식 = 제품 + 시장 + 조직

제품



+

시장

최다 플레이 게임				
최다 플레이 게임 100개 현재 플레이어 수 기준				
순위	게임	가격	현재 플레이어 수	오늘의 최다 동시 접속자 수
1	Counter-Strike 2	무료 플레이	1,100,938	1,527,573
2	Dota 2	무료 플레이	478,432	716,906
3	PUBG: BATTLEGROUNDS	무료 플레이	400,864	631,467
4	Apex 레전드™	무료 플레이	231,960	357,225
5	NARAKA: BLADEPOINT	무료 플레이	180,438	278,864

출처 : STEAM, 24.05.18(토)

커뮤니케이션 = Impact / Time



“이 질문이 결국 당신을 ‘성공’으로 이끌어줄 것이다!”

마이크로소프트 재도약의 비밀 공개
기존의 공식을 다시 쓰는 필독서!



끝나기 전, 마지막으로요!

우리가 인과추론을 배워야 하는 이유

- 조직 : 진수님! 혹시 상품 구성을 바꾸면, 구매 지표가 늘까요?
- 개인 : 매일 마스크 팩을 붙이고 자면, 피부가 좋아질까?

Growth Mindset

Growth



Key Questions

1. 우리가 인과추론을 공부해야 하는 이유
2. 인과추론의 근본적인 문제 이해하기 (feat. 타노스님)
3. 인과추론으로 문제해결하기 (feat. 기안84님)

혹시 저에게 궁금하신 부분이 있다면..



SCAN ME



jsshin2013@gmail.com

감사합니다