

Casual하지 않게 Causality 이해하기

Orientation

- Why Again?
- About Study
- To-do List
- Team Building

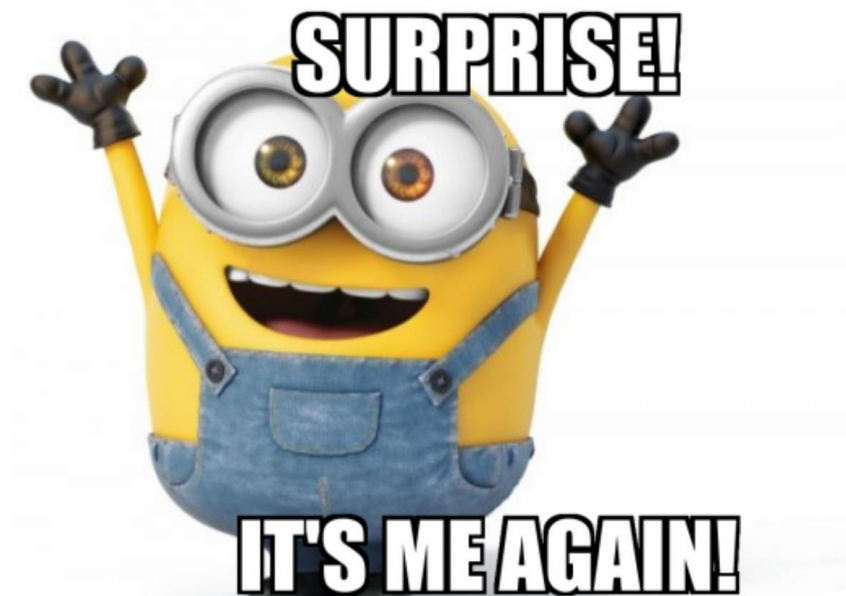


[범죄도시 2]



[Hermione Granger]

It's me again!



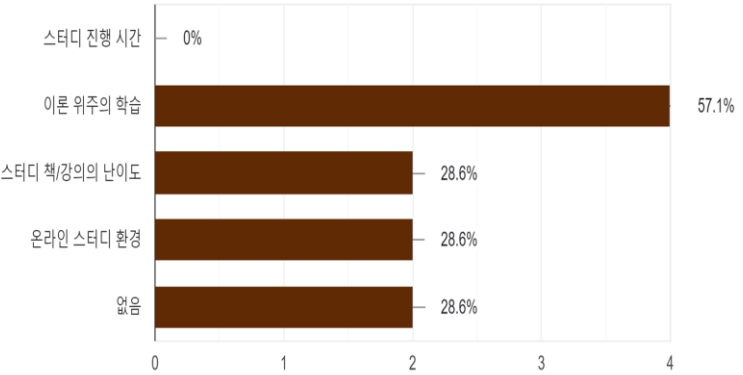
It's me again!





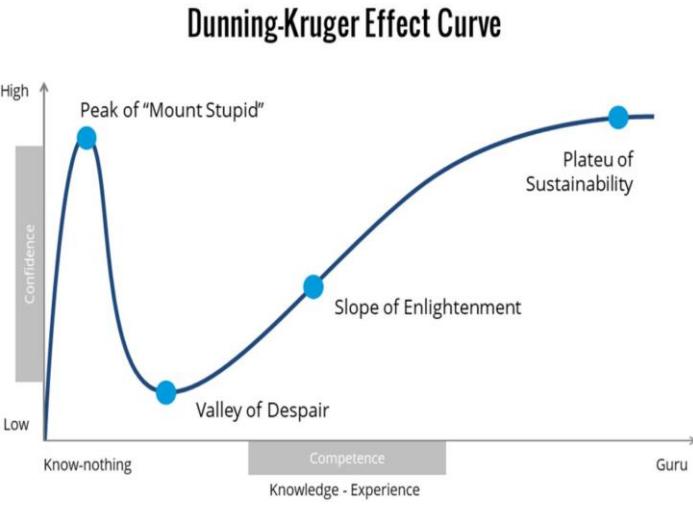
[우리에게 남은 과제]

3. 스터디를 하면서 아쉬웠던 점을 무엇인가요? (중복 가능)



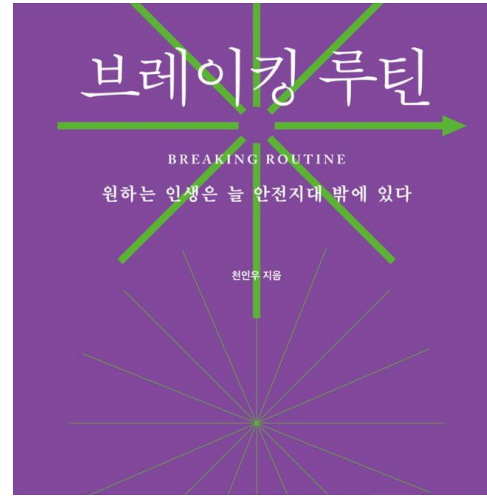
- 이론 위주의 학습
- 어려운 책의 난이도
- 온라인 스터디 환경

[아직 어려운 인과추론...]



- 알면 알수록 어려운 인과추론
- 언제, 어떻게 사용해야 할까?

[Outside of a Bubble]



- 계획만 세우면 아무것도 이뤄지지 않아요
- 22년 새해 목표 달성!

Why Again?



남은 절반의 과제를 해결!

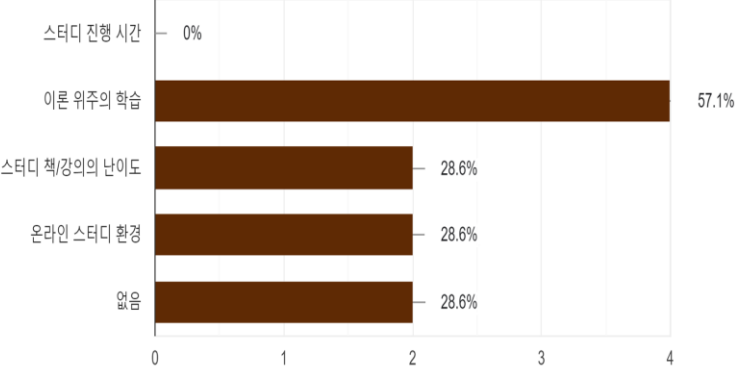
[우리에게 남은 과제]

[아직 어려운 인과추론...]

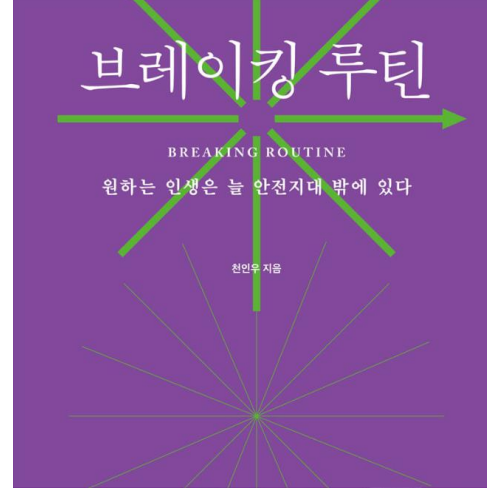
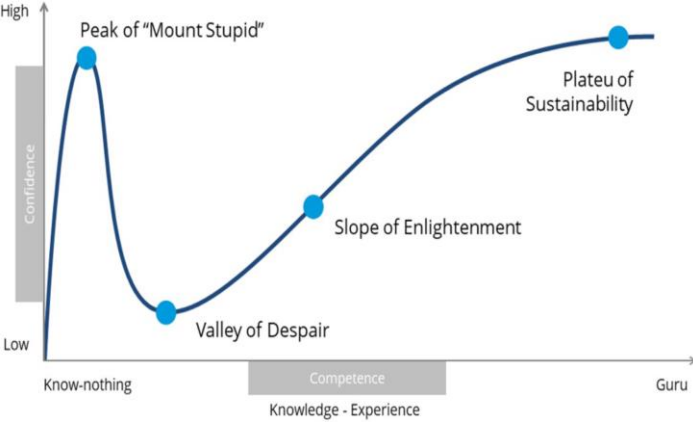
스터디를 통한 선순환!

[Outside of a Bubble]

3. 스터디를 하면서 아쉬웠던 점을 무엇인가요? (중복 가능)



Dunning-Kruger Effect Curve



➔ "Causal Inference의 실질적인 활용방법"
"데이터를 통한 문제해결 방법"

을 여러분과 함께 고민해보고자 해요!

- 브레이킹 루틴 <천인우 저> : [도서 링크](#)
- 더닝 크루거 효과 : [Quora 관련 이미지](#)

[4기 : Causal하게 Causality 이해하기]



[5기 : Causal하지 않게 Causality 이해하기]



[4기 : Causal하게 Causality 이해하기]



마법책을 통해서 주문을 배웠다면



[5기 : Causal하지 않게 Causality 이해하기]



이제는 직접 주문을 써봐야 합니다!

[스터디 목표]

Goal

1. Causal Inference 실습에 대한 한국어 자료 만들기
2. Causal Inference에 대한 개인 TASK를 설정 후, TASK 결과를 팀 블로그에 정리
3. 스터디에서의 의견 공유를 통해, 주 1회 1시간씩 깊은 고민을 해보는 습관 만들기



1. Data를 통한 문제 해결능력 키우기
2. Causal Inference에 대한 심화 이해
3. 인과추론 / 실험설계를 본인 연구/업무에 적용
4. 스펙업!!

◦ 스터디 Notion 페이지 : [Casual하지 않게 Causality 이해하기](#)

◦ 팀 Github : [가짜연구소 Causal Inference Lab](#)

[Team Blog]

가짜연구소 Causal Inference Lab

홈 태그 방명록

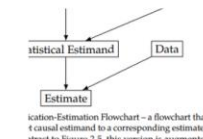
전체 글



Chapter 4. Causal Models

안녕하세요, 가짜연구소 Causal Inference 팀의 김성수, 정호재입니다. Introduction to Causal Inference 강의의 네 번째 챕터이며, 해당 챕터에서 다루는 내용은 아래와 같습니다. Contents Do-operator and Interventional Distributions Modularity Assumption Backdoor...

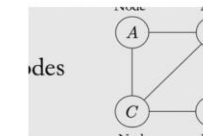
더보기 >



Chapter 3. Graphical Models

안녕하세요, 가짜연구소 Causal Inference 팀의 홍성철입니다. Introduction to Causal Inference 강의의 세 번째 챕터이며, 해당 챕터에서 다루는 내용은 아래와 같습니다. Contents Graphical models란 무엇인가요? Bayesian Networks Causal Graphs Basic building...

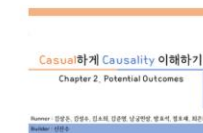
더보기 >



Chapter 2. Potential Outcomes

안녕하세요, 가짜연구소 Causal Inference 팀의 신진수입니다. Introduction to Causal Inference 강의의 두 번째 챕터이며, 해당 챕터에서 다루는 내용은 아래와 같습니다. Contents Potential Outcomes이란 무엇인가요? (aka. Neyman-Rubin Causal model) 인과추론의 근...

더보기 >



[스터디 계획]

계획표 (TBD)

Table List view Calendar +

날짜	Content	발표자
2022/08/06	OT (스터디 방향성 & 안건 논의)	J Jinsoo Shin
2022/08/13	광복절 연휴 잘 보내세요!	
2022/08/20	Chapter 0 Ice Breaking & Revision	J Jinsoo Shin
2022/08/27	Chapter 5, 6 Regression / 사례공유	
2022/09/03	Chapter 10, 11 Matching	
2022/09/10	메리메리 추석추석!	
2022/09/17	오프라인 리뷰1 / Chapter 16 Regression Discontinuity Design	
2022/09/24	Chapter 12 Doubly Robust Estimation	
2022/10/01	Chapter 13 DID	
2022/10/08	Chapter 14 Panel Data and Fixed Effects	
2022/10/15	오프라인 리뷰2	
2022/10/22	쉬어가는 한 주!	
2022/10/29	Chapter 15 Synthetic Control	
2022/11/05	Chapter 17 Meta-Learners	
2022/11/12	Chapter 18 Causal Model Evaluation	
2022/11/19	Chapter 8, 9 Instrumental variables	
2022/11/26	오프라인 리뷰3 & 종강 📄	

[Team Github]

TeamCausality Update README.md

d18ab76 on 28 Jun 30 commits

data	Data File Upload	last month
LICENSE	Initial commit	last month
README.md	Update README.md	last month

README.md

Python으로 하는 인과추론 : 개념부터 실습까지

Causal Inference for The Brave and True 책 번역

안녕하세요. 가짜연구소 Causal Inference 팀입니다.

이 책은 Matheus Facure (Nubank Data Scientist)의 [Causal Inference for The Brave and True](#)을 한국어로 번역한 자료입니다. 이 책은 인과추론에 대한 기본 개념과 Python 실습, 나아가 최신 사례까지 모두 다루고 있습니다. 인과추론에 대한 이론과 실습자료가 많지 않은 상황에서 이 책은 인과추론 처음 접하거나 인과추론에 대한 이해를 바탕으로 사례를 접하고 싶은 분 모두를 만족시킬 수 있는 책이라고 생각합니다. 한국에서 인과추론을 공부하시는 분들께 많은 도움이 되었으면 좋겠습니다.

가짜연구소 Causal Inference 팀은 전문 번역가가 아닌, 데이터분석가/과학자/연구자로 구성 되어 있습니다. 따라서, 한국어 번역 과정에서 이상하거나 어색한 부분이 발생할 수 있어요. 해당 부분에 대한 여러분의 피드백을 부탁드립니다.

마지막으로, 이 책의 한국어 번역에 동의해 주신 Matheus Facure에게 감사의 말씀을 전합니다. 그리고, 번역과 실습자료를 만들기 위해 같이 고생해준 Causal Inference 팀 멤버들에게도 진심으로 감사드립니다!

➡ 실습책 사례 논의 + 개인 Task 설정

- 스터디 Notion 페이지 : [Casual하지 않게 Causality 이해하기](#)
- 팀 Github : [가짜연구소 Causal Inference Lab](#)

- 개인별 TASK 예시

(1) 인과추론 or 데이터 분석 설계 관련 본인 연구/업무에 적용할 수 있는 사례 공유

(2) 일상생활에서 인과추론을 적용해볼 만한 사례 or
데이터로 문제를 해결해보면 좋은 사례를 블로그에 정리

(3) 인과추론 관련 본인만의 라이브러리 만들어 보기 or
CausalML, EconML, Pywhy 등과 같은 라이브러리 한국어 실습자료 만들기

(4) 데이터분석 or 인과추론 관련된 도서 후기 블로그에 정리

(5) 인과추론 관련 세미나 or Tech Blog 블로그 정리

업무 적용 사례 예시

★ 분석설계

- 문제 인식: ① 유로아이템을 무료로 일부 지급. < 업데이트 시 >
② 무료 이벤트로 ARPPU(반복구매)가 이전 판매 대비, 유의미하게 늘었는가?

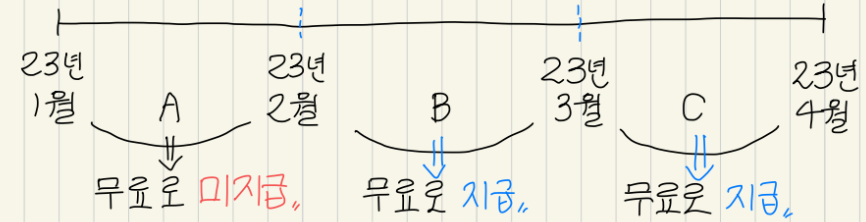
2. 구체화: 해당 이벤트 지속 여부 or 주요 매출원 or 이벤트 비용,
⇒ 이벤트 비용 대비, PU 전환 & ARPPU 증가가 있었는지

3. 가설설정
“ , 유의미한 증가는 없을 것이다,

4. 필요한 Dataset,

- 우편 정보 <아이템 획득> → Treatment group,
 - 유로아이템 구매 정보 <과금> → ARPPU
- 시점 < 무료 지급 이전 ~ 이후 >

5. Validation (동일기간): 업데이트 마다, 갱신되는 상품이 달라짐,



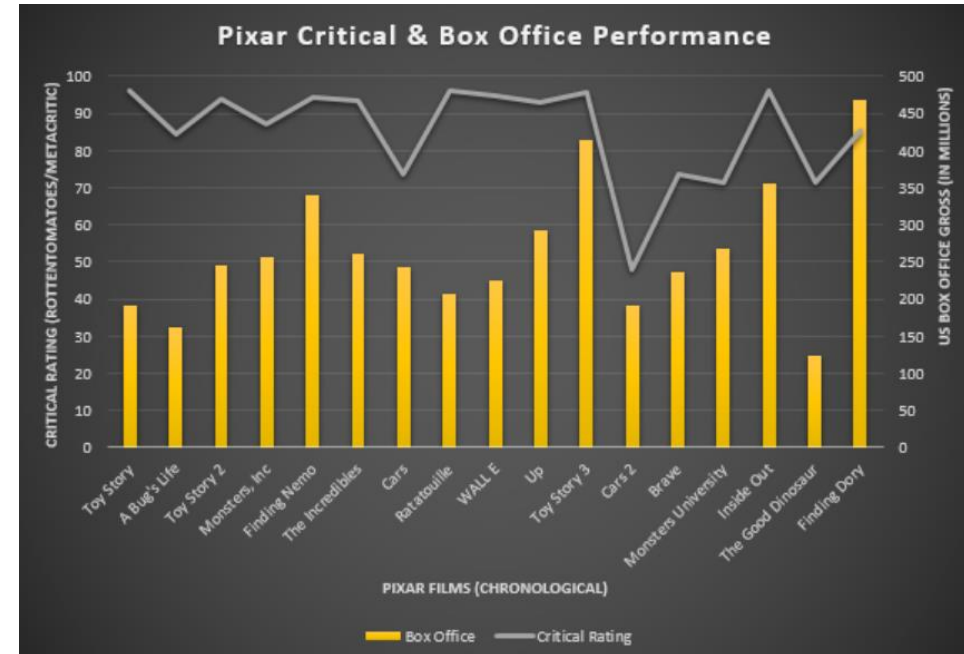
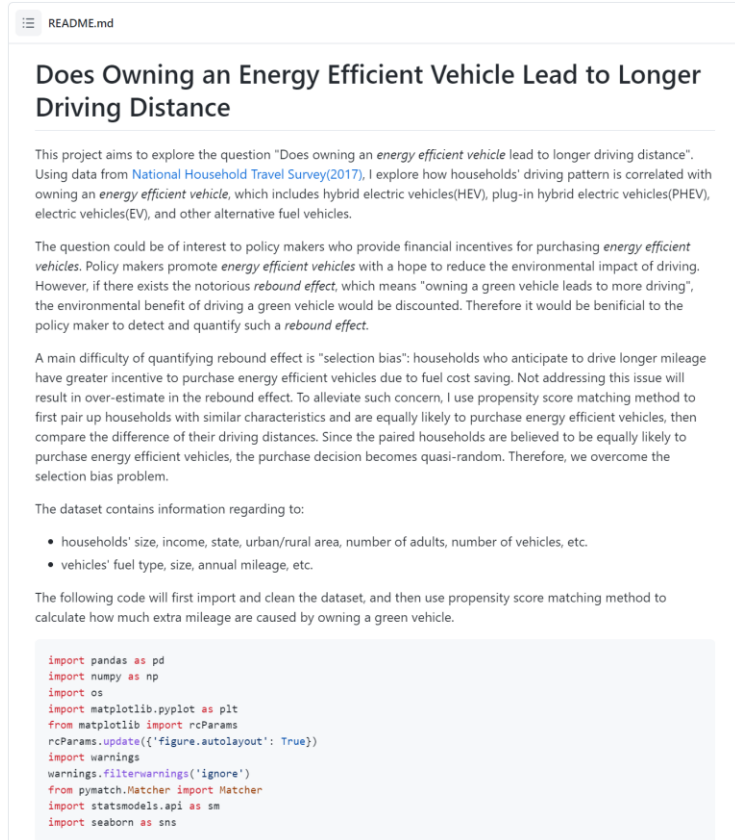
• 해당 프로모션 이외의 영향은 동일하다고 가정,
(외부영향)

- B-A: 무료 지급 영향 +
- C-B: 아이템 구성에 따른 차이,
- ↳ B-A에서 순수하게 무료 지급의 영향만을 보려면?

• Non-compliance,

↳ 우편은 받았는데 사용하지 않았음,

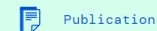
- 문제 해결 사례 예시



- Python/R 인과추론 관련 라이브러리 / 논문 예시

```
class SequenceProcessing :  
  
    def __init__(self, data, delimiter, category_col) :  
        self.data = data  
        self.delimiter = delimiter  
        self.category_col = category_col  
  
    def sequence_timeinterval(self):  
  
        ## Condition & sorting (sorting columns should be clarified)  
        df_except = self.data[self.data[self.category_col] != exception_col]  
        df_sorting = data_sorting(df_except, sorting_list)
```

[Python Class로 만들어 보기!]



Relating Graph Neural Networks to Structural Causal Models

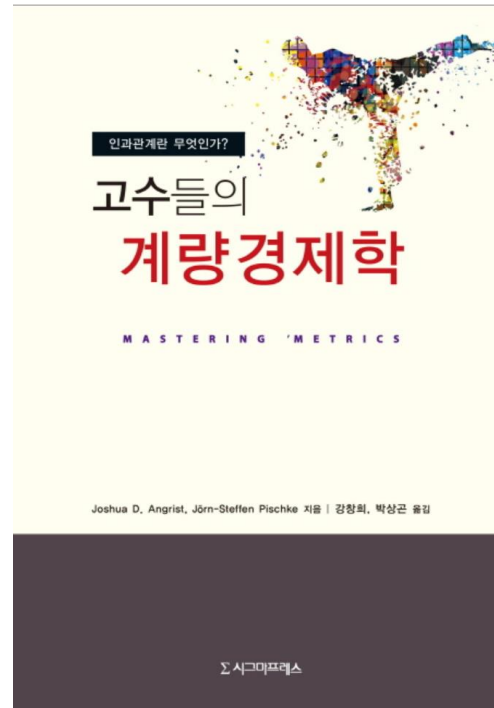
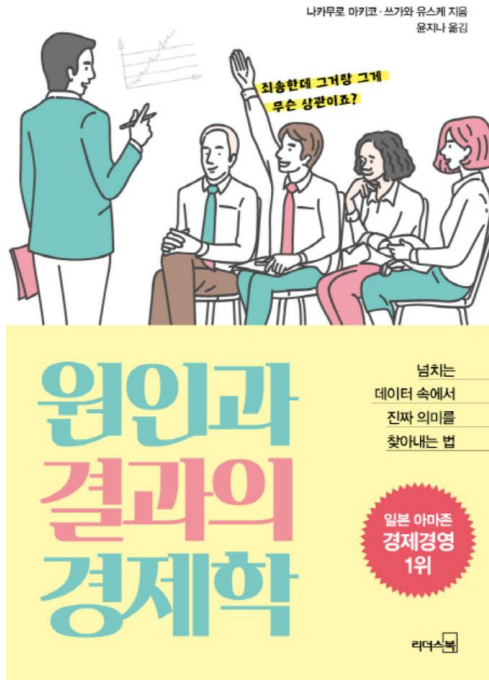
[Download](#)[View publication](#)

Abstract

Causality can be described in terms of a structural causal model (SCM) that carries information on the variables of interest and their mechanistic relations. For most processes of interest the underlying SCM will only be partially observable, thus causal inference tries to leverage any exposed information. Graph neural networks (GNN) as universal approximators on structured input pose a viable candidate for causal learning, suggesting a tighter integration with SCM. To this effect we present a theoretical analysis from first principles that establishes a novel connection between GNN and SCM while providing an extended view on general neural-causal models. We then establish a new model class for GNN-based causal inference that is necessary and sufficient for causal effect identification. Our empirical illustration on simulations and standard benchmarks validate our theoretical proofs.

[민상님 Nature Accept 기원]

- Data Analytics / Causal Inference 관련 도서



- 원인과 결과의 경제학 : [도서 교보문고 링크](#)
- 고수들의 계량경제학 : [도서 교보문고 링크](#)
- 데이터분석의 힘 : [도서 교보문고 링크](#)

- 팀 블로그 / Github 관련 논의

- (1) Contribution

- 챗터별 작성자 & 발표자 구분
 - 운영방법 : 탈중앙화(?) - 팀 블로그 & 개인 블로그

- (2) 작성

- 마무리 기간 설정
 - 썸네일 표지 및 양식관련

- (3) 팀 Github 안내

- 팀원 초대 (각자의 Github)
 - 실습책 번역에 대한 Commit 방식 논의

- 스터디 계획 / 방향성

- (1) 스터디 진행방식 논의

- 진행 시간
 - 온/오프라인 진행 방식
 - 스터디 계획 Feedback
 - 스터디 진행방식 Feedback 규칙

- (2) 스터디 운영

- 인과추론 이외 Data Engineering / ML / 등의 자유로운 공유 / 세미나
 - 출석 관련
 - 4기 수료증 or 궁금한 사항

[함께하는 즐거움]



[Luck favours the prepared]



[Exploration / Exploitation]



"이 스터디를 통해, 인과추론 뿐만 아니라 이외의 부분도
얻어갈 수 있는 좋은 스터디를 함께 만들어 봅시다!"



Causal-Lab
가짜연구소

Q & A



감사합니다
