

Data Science

데이터 과학 개요

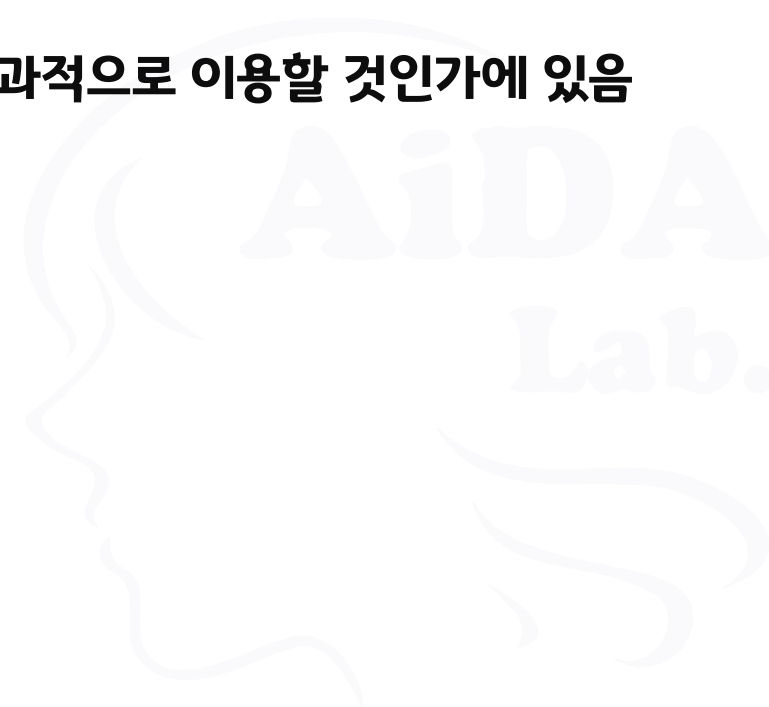
강사 양석환



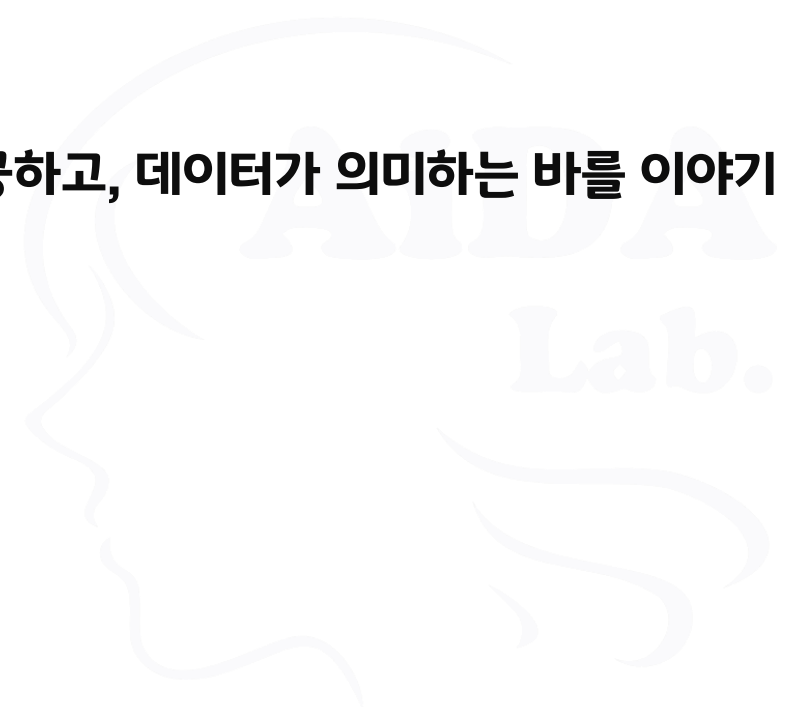
데이터 과학과 데이터 과학자



- 정리되지 않은 데이터에서 인사이트(insight), 즉 유용한 규칙을 발견하고자 하는 학문
- 데이터 과학은 실무적인 필요에 의해 성립된 학문이므로
 - 빅데이터 환경에서 영리를 목적으로 하는 기업이건 정부나 민간 단체와 같은 비영리 조직이건
 - 공통적인 관심은 어떻게 하면 대량의 데이터로부터 가치를 창출해 효과적으로 이용할 것인가에 있음



- 데이터 과학과 관련된 분야를 전공하고 데이터 분석과 관련된 업무에 종사하는 사람
 - 흔히 컴퓨터 과학자보다는 통계학을 더 잘 알고, 통계학자보다는 컴퓨터 과학을 더 잘 아는 사람이라는 식으로 표현하기도 함
- 데이터 과학자의 역할
 - 현장에 존재하는 대량의 데이터를 모으고, 분석에 적합한 형태로 가공하고, 데이터가 의미하는 바를 이야기 (story)에 담아 다른 사람에게 효과적으로 전달하는 역할

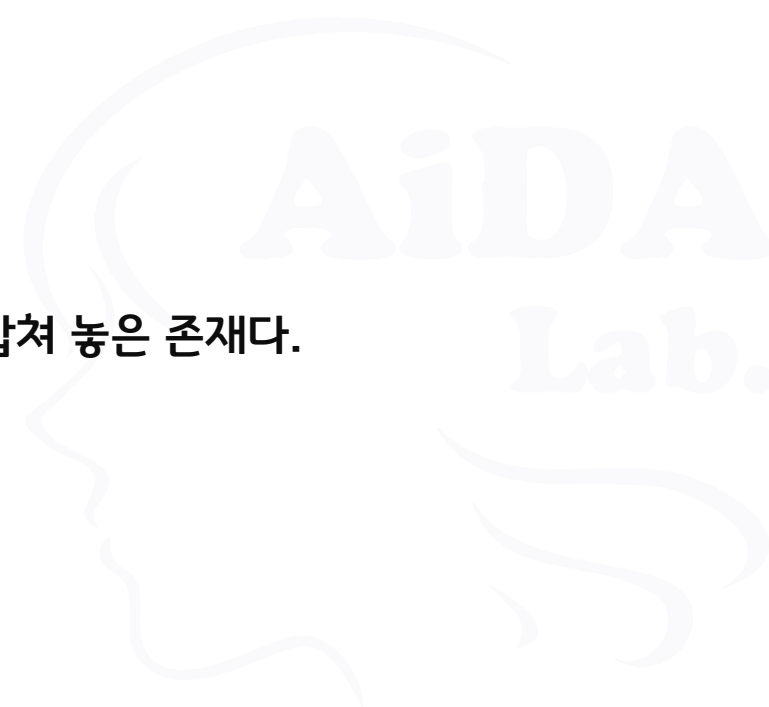


- 데이터 과학과 데이터 분석은 어떻게 다른가?
 - 차이는 거의 없다고 보면 됨
 - 데이터 과학, 데이터 분석 등의 영역은 워낙 광범위하게 퍼져 있기때문에 명확한 분류 기준 등이 없음
- 통계학과 비교한다면?
 - 통계학이 정형화된 실험데이터를 분석 대상으로 하는 것에 비해
 - 데이터 과학은 기업의 실무 현장에서 쌓이는 빅데이터를 대상으로 하는 정도
 - 통계학과는 달리 데이터 과학은 총체적(holistic) 접근법을 사용한다는 정도

- 데이터 과학이 전망이 좋다고 하는데 이것만 잘 하면 되는가?
 - 데이터 과학은 컴퓨터 공학과 통계학 등 다양한 관련 학문이 통합된 의미
 - 실제로 이러한 모든 분야에 정통한 전문가는 존재하기 어려움
 - 따라서 실무에서 데이터 분석은 다양한 전문가로 구성된 팀을 구성해 진행하는 것이 일반적임
 - 그렇다고 하더라도 데이터 과학자는 태생적으로 최소한 2, 3개의 분야에 정통한 학제적 배경을 가지고 있어야 다양한 분야에 활용할 수 있음

• 데이터 과학이란?

- 데이터 과학, 데이터 과학자는 비교적 최근에 등장한 개념이기 때문에 정의에 대해서는 다양한 견해가 존재함
- 링크드인(LinkedIn)의 수석 과학자인 로가티(Rogati)의 의견
 - 모든 과학자는 데이터 과학자다.
 - 데이터 과학자는 반은 해커이고 반은 분석가다.
 - 마치 반짝이는 눈을 가진 탐험가 콜럼버스와 의심 많은 형사 콜롬보를 합쳐 놓은 존재다.



- 데이터에 기반한 기업의 활동 범위가 점차 넓어짐에 따라
 - 기업의 주요 의사결정, 비즈니스 인텔리전스(BI, Business Intelligence)
 - 생산과 마케팅 분석
 - 사기 방지
 - 위험관리, 보안
 - 데이터 서비스와 운용, 데이터 인프라 등
 - 그 외에도 다양한 영역에서 활동 중

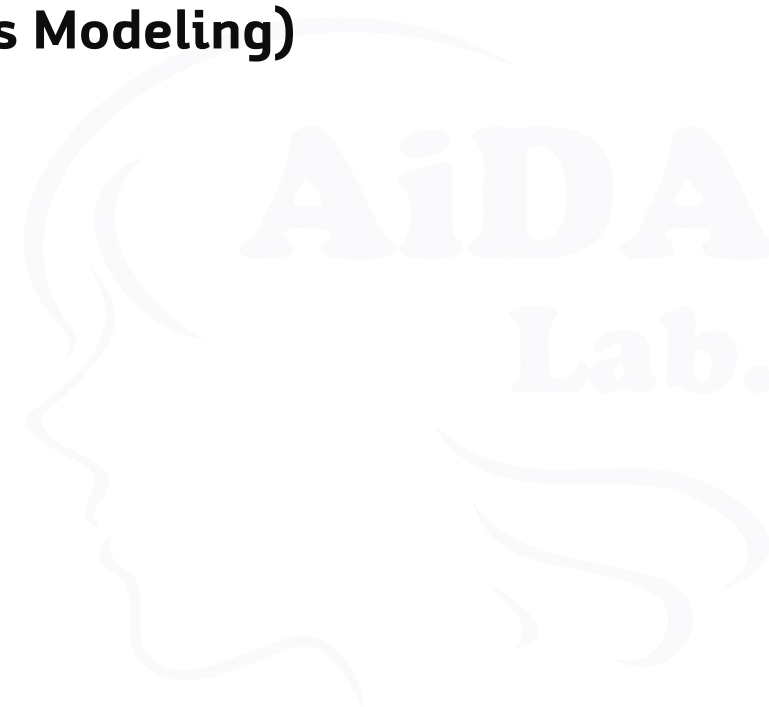


- 대표적인 3가지 핵심 기술

- 기본적으로 여기저기에 산재되어 있는 분석에 필요한 데이터를 모으고 가공하는 데이터 처리(Data Management)
- 분석에 필요한 모형을 만들고 결과를 도출하는 분석 능력(Analytics Modeling)
- 해당 업종에 대한 이해(Business Analysis)

- 데이터 과학이 예술의 경지로 진화하려면

- 의사소통 능력, 협업, 리더십, 창의력, 규율, 열정이라는 요소



- 데이터 과학자가 지녀야 할 덕목

- 기술 전문성(Technical expertise)
- 호기심(Curiosity)
- 데이터로부터 이야기를 만들어내고 이를 효과적으로 전달하는 능력(Storytelling)
- 문제 해결을 위해 창의적인 관점에서 접근하는 능력(Cleverness)



**THANK
YOU**

