

Data Literacy That Drives Results in the Era of Big Data

통계, 분석, 데이터 사이언스

실현가능

중요한 것은 앞으로 자신에게 필요한 기술인지 무엇인지 아는 것

데이터 활용에 필요한 3가지

생각하고 > 작업하고 > 생각한다

: 분석 전에 문제 및 목적을 정의하고 가설들을 구축 > 분석을 위한 기술과 지식 > 분석 결과에 대한 해석 및 스토리 구축

목적 사고력

데이터 안에 해답 따위가 존재하지 않는다.

우연히 찾은 데이터는 실제에서 극히 일부이기 때문이다.

데이터 분석 자체가 목적이 되기 십상이다.

목적 및 문제를 정의 > 지표를 결정 > 현재 상태를 파악 > 평가 > 요인을 분석 > 해결 방안을 모색

데이터를 가공한다 x

데이터에서 무엇을 알 수 있을지 생각한다 o

데이터를 모아 현재 상황을 파악한다.

목적과 문제에 대한 정의

- 사용된 언어가 구체적이고 명확한가
- 문제, 원인, 해결 방안을 구별하고 있는가
- 자신의 편견이 들어가지 않도록 주의

나는 문제를 명확하게 정의해 풀고 있는가?

Are you solving the right problem?

정의한 문제와 사용하는 데이터가 일치 하지 않을때

~ 짐작, 추측이 아닌 논리적인 설명을 통해 이용할 지표 결정

더 나은 목적과 문제를 정의하기

가치있는 정보

사실 및 결과 표시 != 평가

평가를 객관적으로 만드는 것은 바로 다른 것과의 비교

목적 중심 사고

- > 데이터 문해력의 본질 중 하나
- > 데이터 중심 x

추이와 변화

편차 고려

"당신이 직면하고 있는 문제나 그 배경에 대해 전혀 모르는 제3자가 바로 앞에 있다고 상상해보세요. 그 사람에게 당신이 다루는 문제를 데이터로 설명하려면 무슨 데이터를 어떻게 보여주면 좋을까요?"

원인 파악력 : 행동으로 이어지는 힘

현황 파악 및 평가 ~~ 행동 및 판단

→ 이 사이에는 간극이 존재한다

해결 방안은 원인에 실행되어야 한다는 것

원인 후보 열거 -> 지표 결정 -> 관련성을 확인

데이터 수집이 어려운 경우

: 유사한 데이터 파악 (정확도는 떨어져도 없는 것보다는 나음)

: 당장 수집 시작

: 정량이 아닌 정성적 정보로 대응

: 포기

활용하는 데이터 수와 조건에 따라 상관계수의 정확도는 큰 차이 존재

데이터를 뒤져보면 그 안에 반드시 답이 있을 것이다. 그것을 찾아내자 x

데이터를 보기 전에 가능성 따지기 o

원인은 한 가지가 아니라 여럿이거나 복잡할 수도 있다.

선형이 아닌 관계성도 존재

상관관계는 인과관계를 의미하는 것이 아니다.

전체 구성력 : 스토리 (논리)를 만드는 힘

문제 해결 프로세스에서 누구나 하는 실수

; 곧바로 해결 방안에 무턱대고 달려드는 것

; 결과와 유효성에 차이가 생기는 원인 파악 유무

결과 != 결론

시야 확대력

데이터 안에는 답이 없다는 전제로 시작

논리적 사고, 로지컬 핑킹

데이터 문해력 : 데이터에서 무언가를 읽어내는 능력이 아닌 스스로 정답에 대해 고민하고
데이터를 무기 삼아 합리적으로 논할 수 있는 능력