

Tableau DataFest



Self-BI : 혁신으로 가는 큰 걸음

HONG SEONGBONG

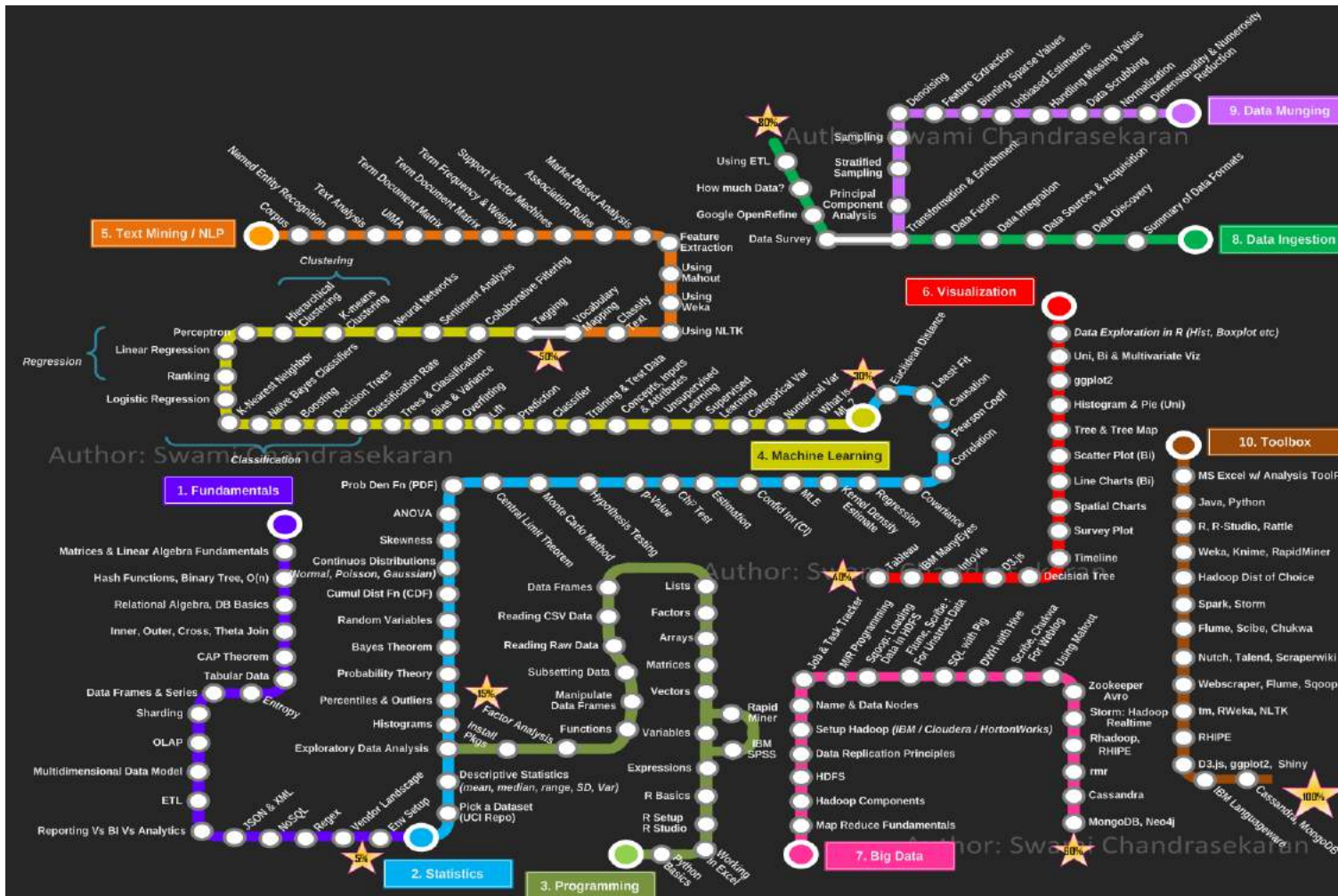
Chief Digital Technology Officer
Digital Technology Development Division
AMOREPACIFIC

#Data22



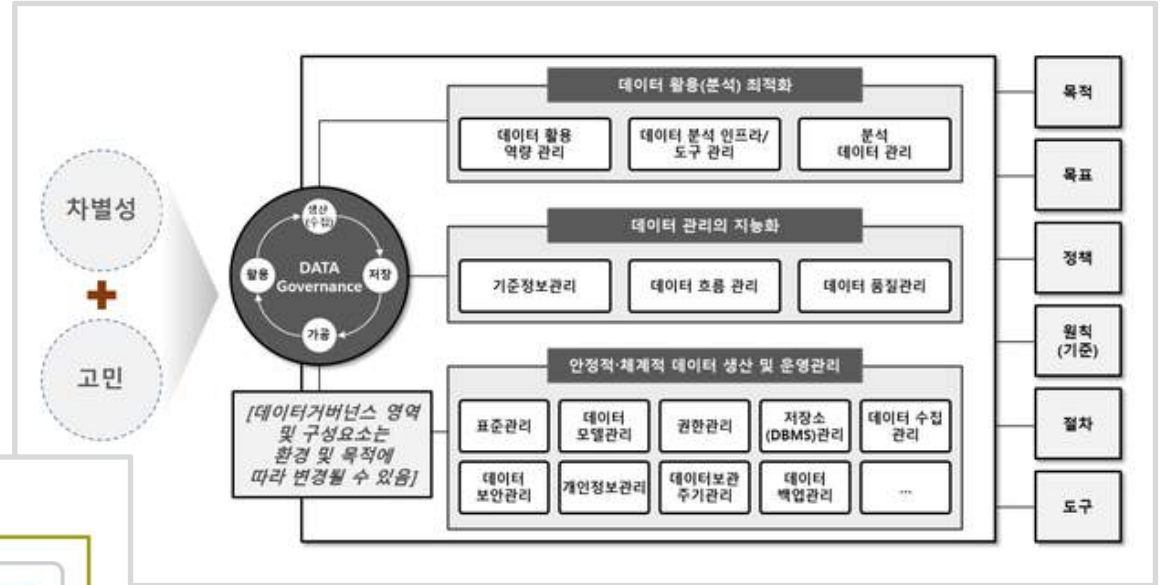
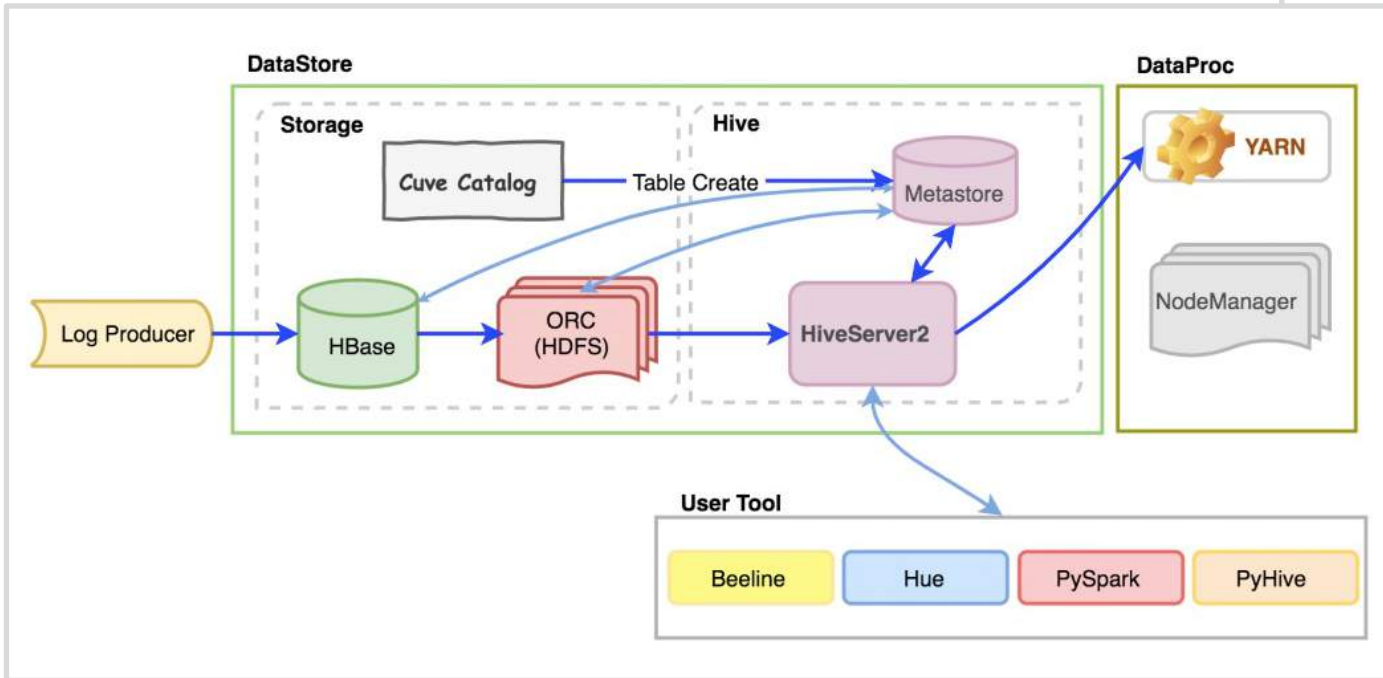
들어가며

데이터 사이언스 지하철맵 - 데이터 관련해서 준비해야 할 것들이 많아 보인다



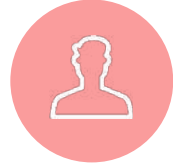
오늘 잠시 잊어야 할 것들

데이터 거버넌스, 빅데이터 플랫폼 같은 것들!



데이터를 너무 어렵게 보지 않기

데이터 전문가 vs 현장 전문가



빅데이터 vs 쓸모 있는 데이터



통계적 분석과 AI vs 데이터 탐색



Technology vs 상식과 논리



데이터는 어디서 확보해야 하나..?

보고서는 일일이 수동으로..?

그 많은 데이터에서 어떻게,
어떠한 인사이트를 뽑아내야 하는가?

내 고객이 누구인지 찾아내고
그들에게 타겟팅 할 수 있는 방법은?

...



그런데 현실은 어렵다

늘어만 가는 마케터의 고민...



쉬울 것 같지만 어려운 질문 :

Data는 정말 중요한가



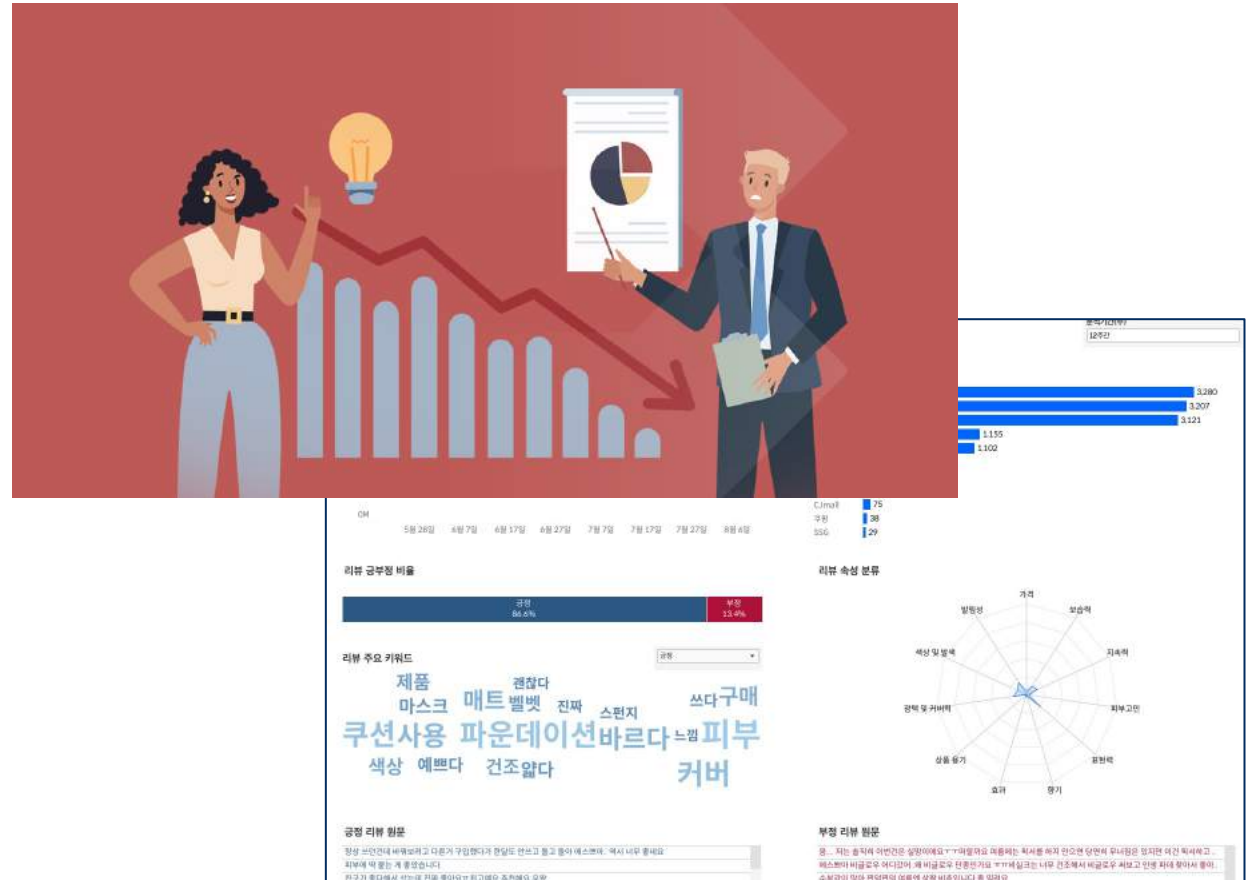
데이터는 (존재하지만) 잘 보이지 않는 것을 보이게 한다

물류 : 고객 주문을 24시간 내에 처리할 방법은?



마케팅 & CRM : 낮아지는 마케팅 효율을 어떻게 개선할 것인가?

세일즈 : 최근 3개월간 우리 제품의 시장 경쟁력은 어떠한가?

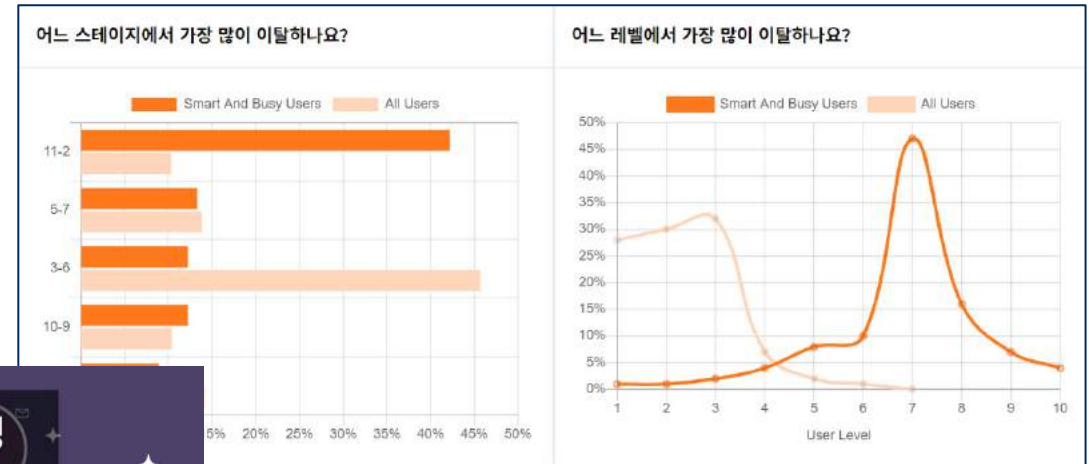


데이터는 보이지 않는 관계성을 알려준다

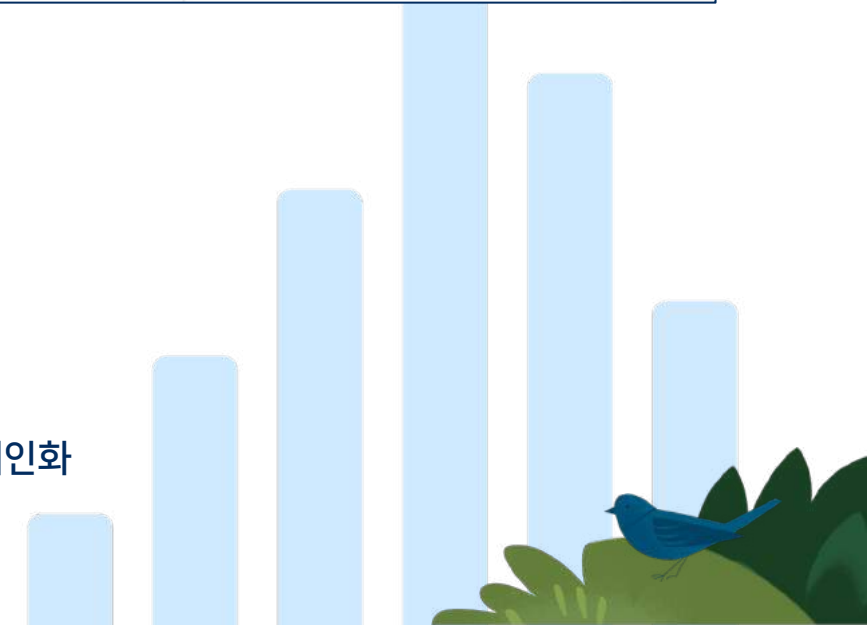
수요 예측



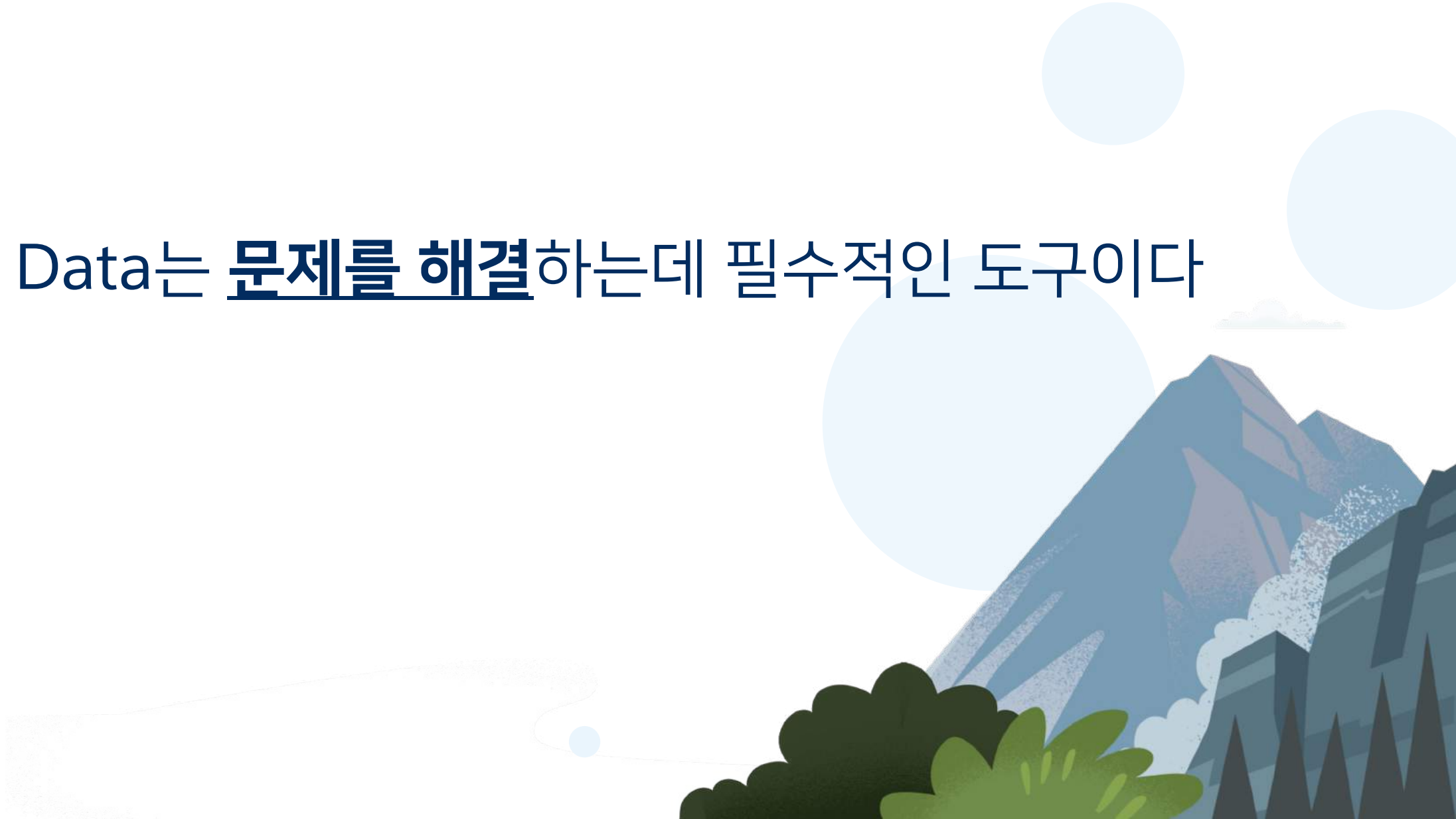
사용자 이탈 예측



콘텐츠 개인화



이제 Data는 문제를 해결하는데 필수적인 도구이다



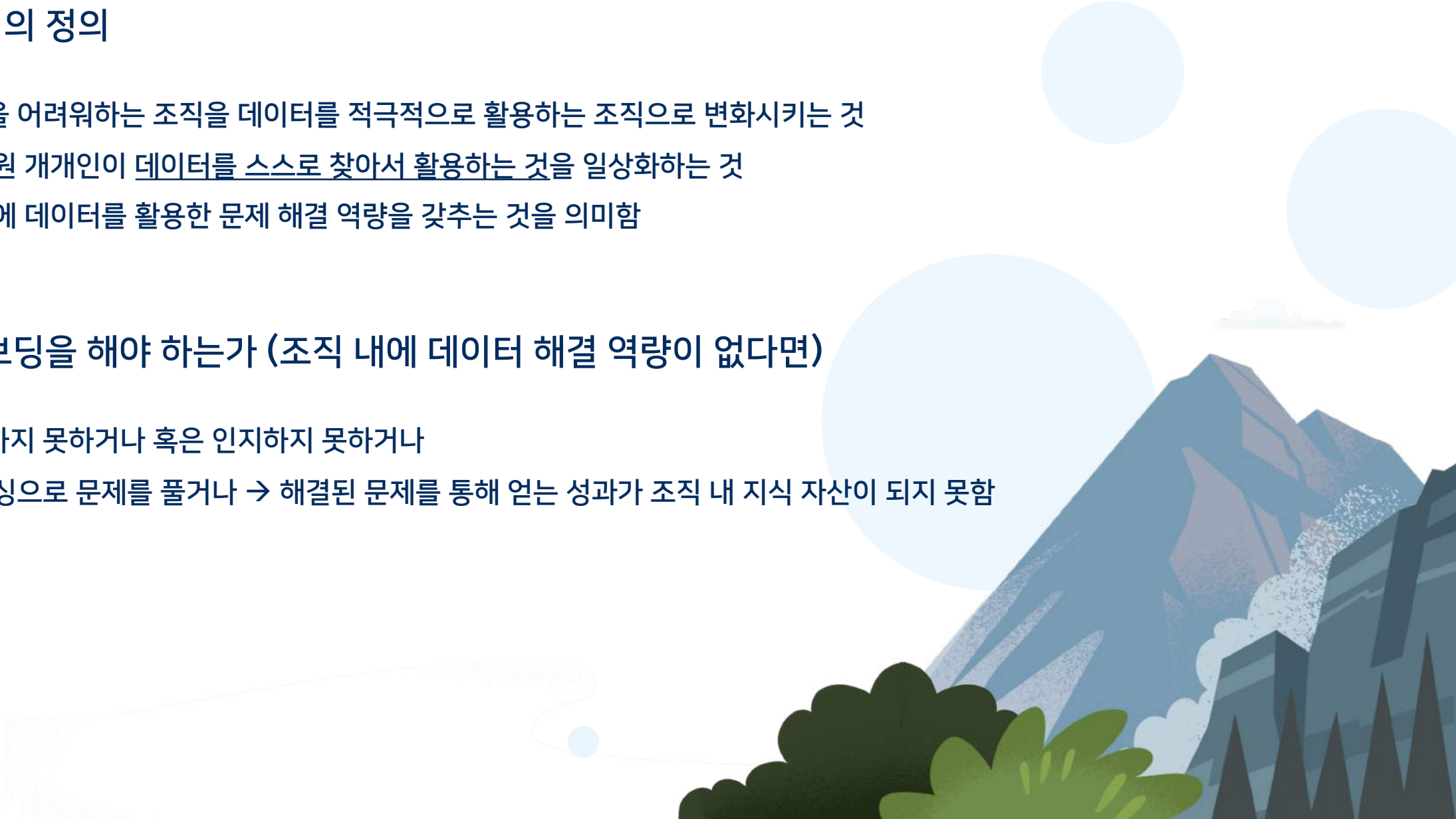
Data On-Boarding

데이터 온보딩의 정의

- 데이터 활용을 어려워하는 조직을 데이터를 적극적으로 활용하는 조직으로 변화시키는 것
- 조직 내 임직원 개개인이 데이터를 스스로 찾아서 활용하는 것을 일상화하는 것
- 결국 조직 내에 데이터를 활용한 문제 해결 역량을 갖추는 것을 의미함

왜 데이터 온보딩을 해야 하는가 (조직 내에 데이터 해결 역량이 없다면)

- 문제를 해결하지 못하거나 혹은 인지하지 못하거나
- 외부 아웃 소싱으로 문제를 풀거나 → 해결된 문제를 통해 얻는 성과가 조직 내 지식 자산이 되지 못함



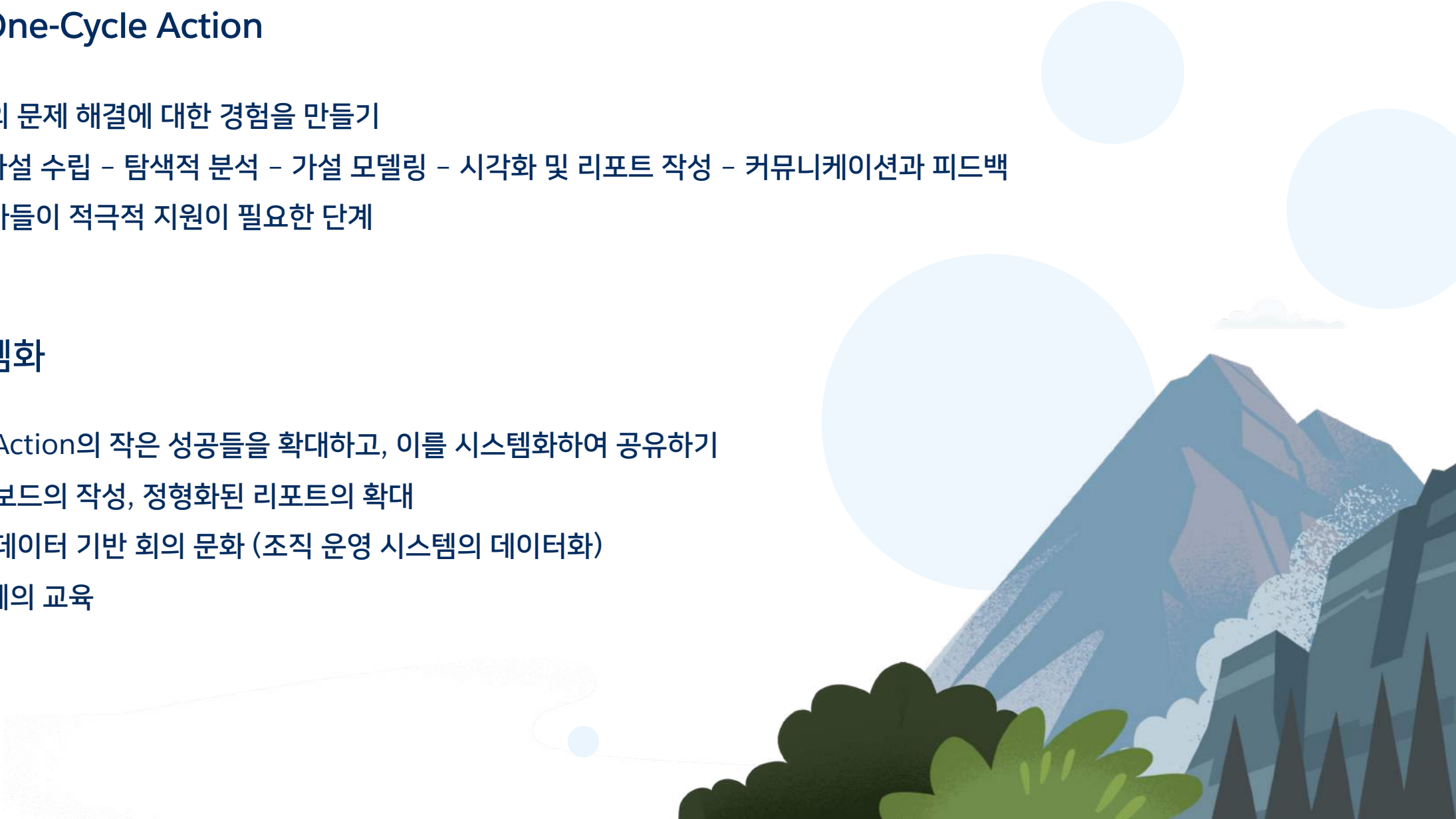
Data On-Boarding : How ?

Initiation : One-Cycle Action

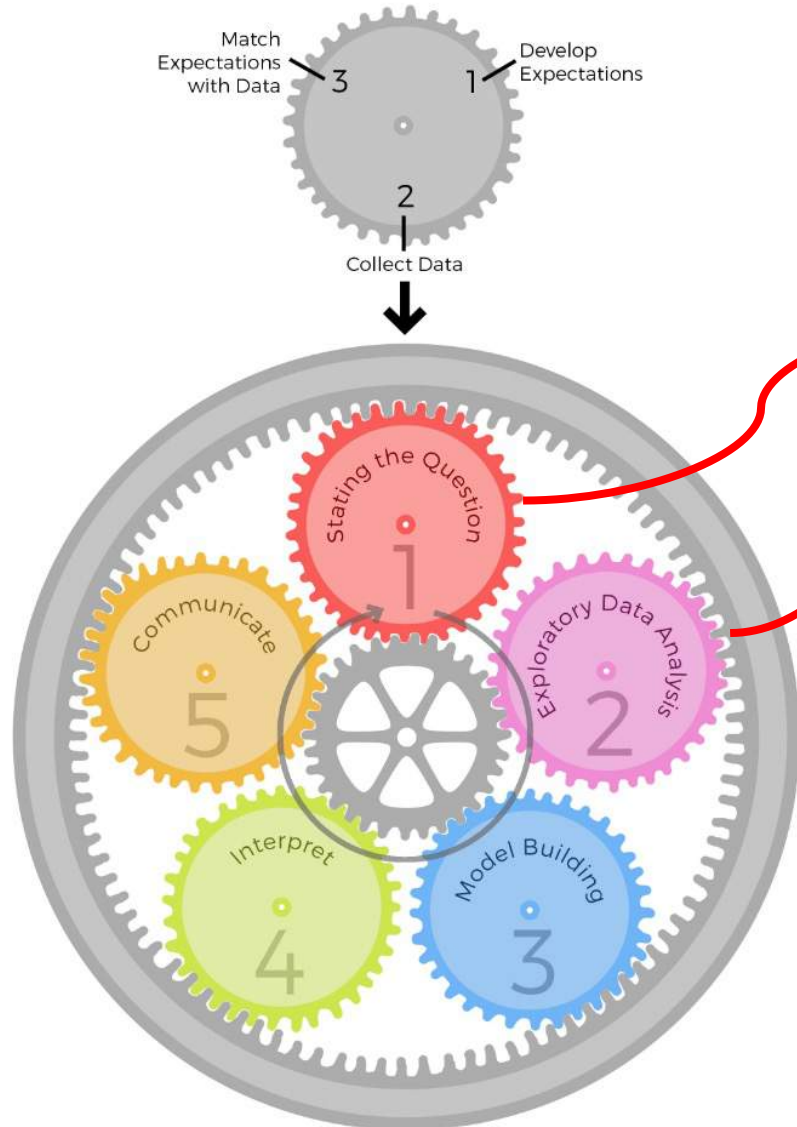
- 데이터 기반의 문제 해결에 대한 경험을 만들기
- 문제 정의 - 가설 수립 - 탐색적 분석 - 가설 모델링 - 시각화 및 리포트 작성 - 커뮤니케이션과 피드백
- 데이터 전문가들이 적극적 지원이 필요한 단계

확산 및 시스템화

- One-Cycle Action의 작은 성공들을 확대하고, 이를 시스템화하여 공유하기
- 조직 내 대시보드의 작성, 정형화된 리포트의 확대
- 리더 중심의 데이터 기반 회의 문화 (조직 운영 시스템의 데이터화)
- 정형화된 사례의 교육



One-Cycle Actions



주목할 부분

- (1) 좋은 질문을 생각하기
- (2) 데이터를 찾아보고 가설을 세우고, 분석하기
- (3) 기타 등등 - 전문가들에게 맡길 부분들



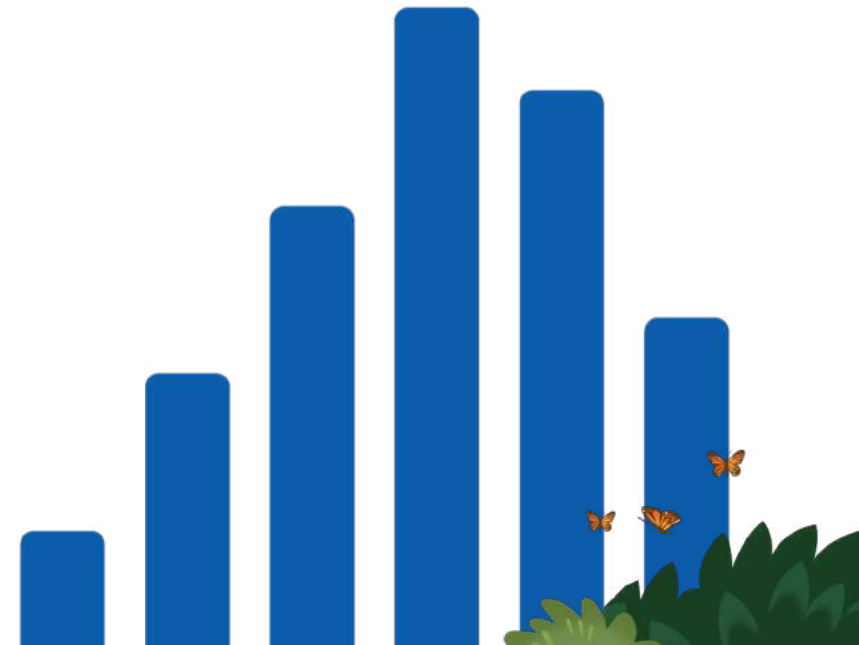
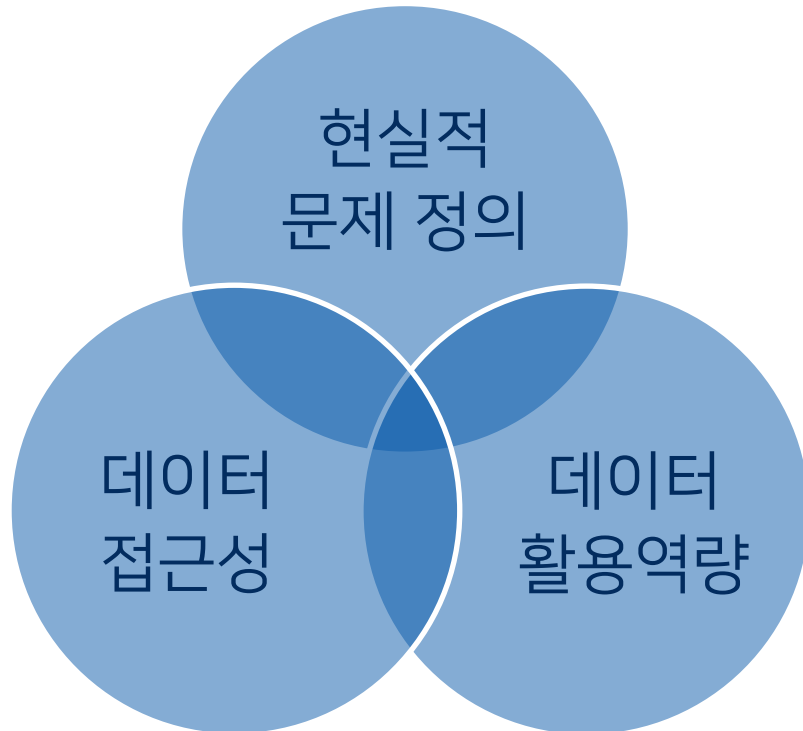
데이터 활용에 대한 조직적 변화 관리와
데이터 접근성의 확대/강화가 필요함

CIO/CTO님들이 집중할 문제

Q1. 현장의 비즈니스 문제를 데이터를 활용하면 '더 잘 해결할 수 있다'는 조직적 공감대를 어떻게 만들까?

Q2. 데이터 전문가가 아니더라도 직접 데이터를 추출하고 활용할 수 있는 환경을 어떻게 만들까?

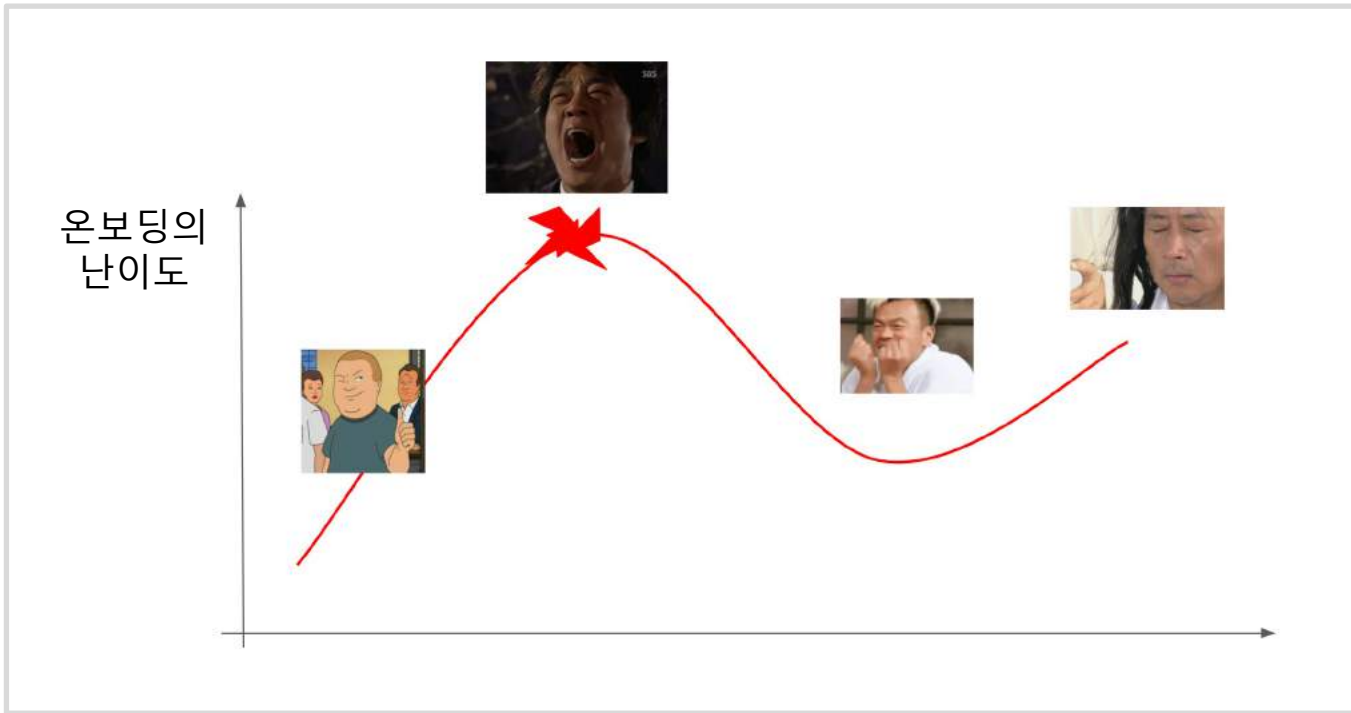
Q3. 위의 2가지를 어떻게 현장의 인력들이 직접 실행하게 만들까? (CoE 조직의 한계를 극복해야 함)



Data On-Boarding 수준과 단계의 고민

무엇을 지원해줄 것인가 : 데이터 도구? 교육? 분석가 직접 지원?

어디까지 지원해줄 것인가



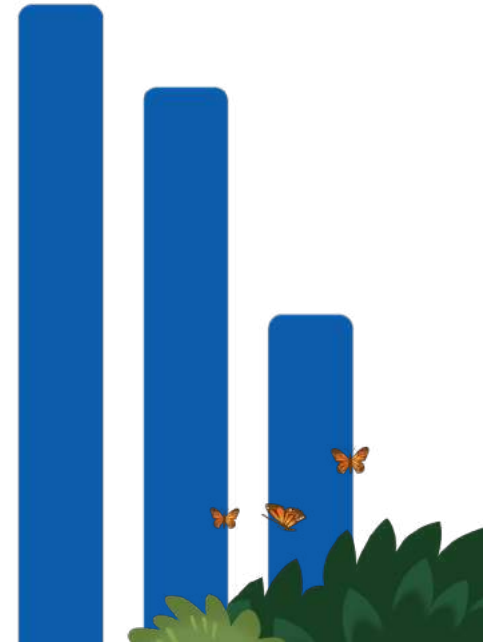
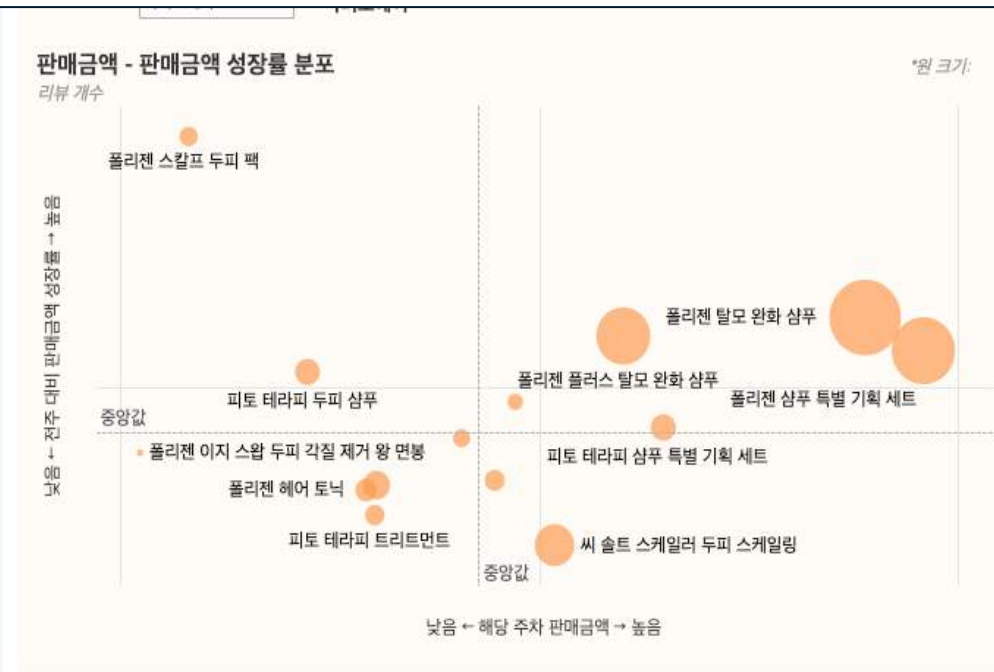
현실적 문제 정의하기 : story telling

쉽지만 사업에 임팩트있는 지표와 데이터 스토리를 정의하기

제품의 시장 점유율
고객의 재방문율, 재구매율

복수 개의 지표를 연결해 보기, 타겟간 비교해 보기

매출과 성장율
고객수와 재구매율
모든 시표를 시계열로 연계해보기



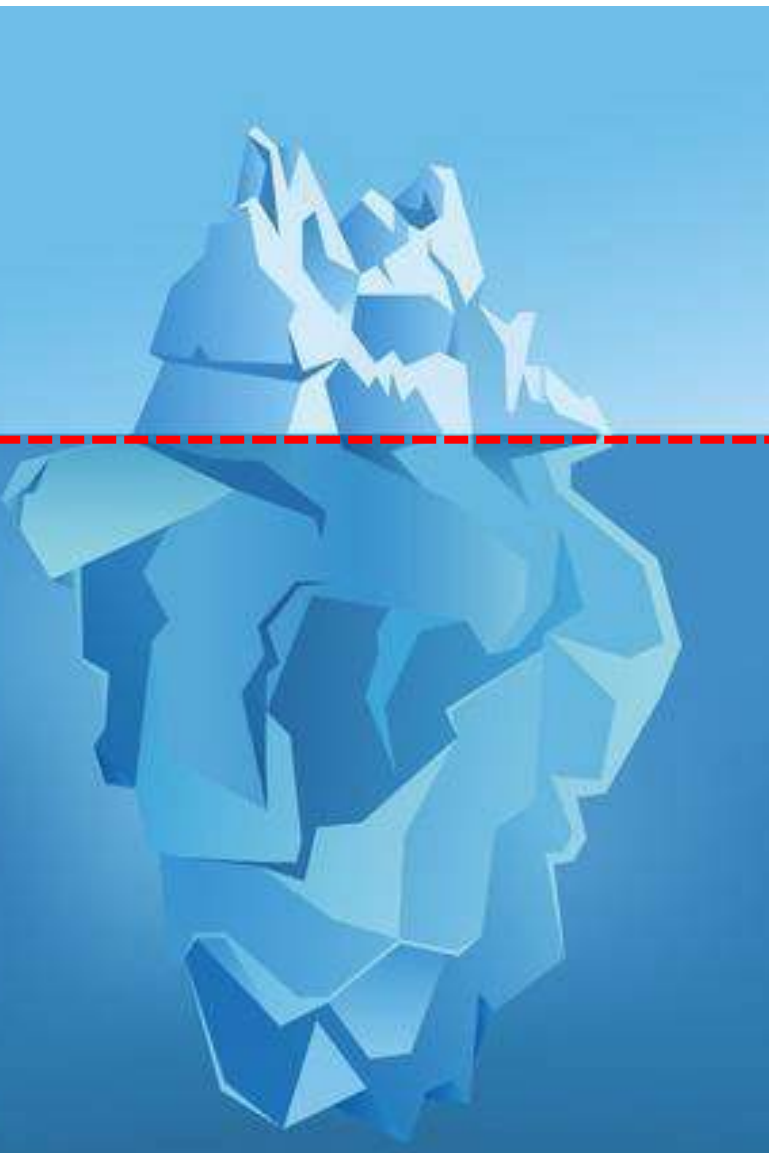
현실적 문제 정의하기 : 데이터 활동의 수준 가이드하기

데이터 장인들의 영역 → Special Skills, Governance, 데이터 기획 필요



전문가의 영역도 있지만,
일반인도 할 수 있는 (해야 하는) 영역

현실적 문제 정의하기 : 분석의 수준을 가이드하기



The diagram features a large iceberg floating in a blue ocean. A horizontal red dashed line divides the iceberg into two parts. The top part, which is above the water, is divided into two sections corresponding to 'Descriptive Analysis' and 'Exploratory Analysis'. The bottom part, which is submerged, is divided into two sections corresponding to 'Inferential Analysis' and 'Predictive/Causal Analysis'. The iceberg is rendered in various shades of blue, with the submerged part appearing darker.

Descriptive Analysis

데이터 요약/집계 후 결과 도출
Report & Visualize

Exploratory Analysis

트렌드, 패턴, 관계를 찾는 것
그래프/표를 통한 직관적 정보 탐색 중심

Inferential Analysis

실험을 통해 얻은 표본 집단의 특징이
일반적으로 적용되는 사실인지 검증하는 분석

Predictive/Causal Analysis

일어나지 않은 현상에 대한 기술적 예측
수십 개 이상 변수들간의 상관관계 해석

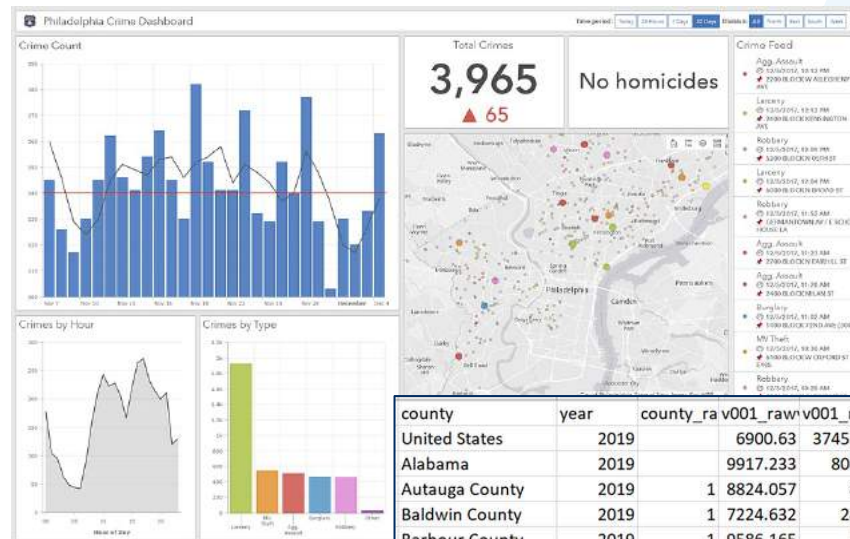


데이터 접근성 확보하기 : Data Preparation

Dashboard? Or Raw Data?

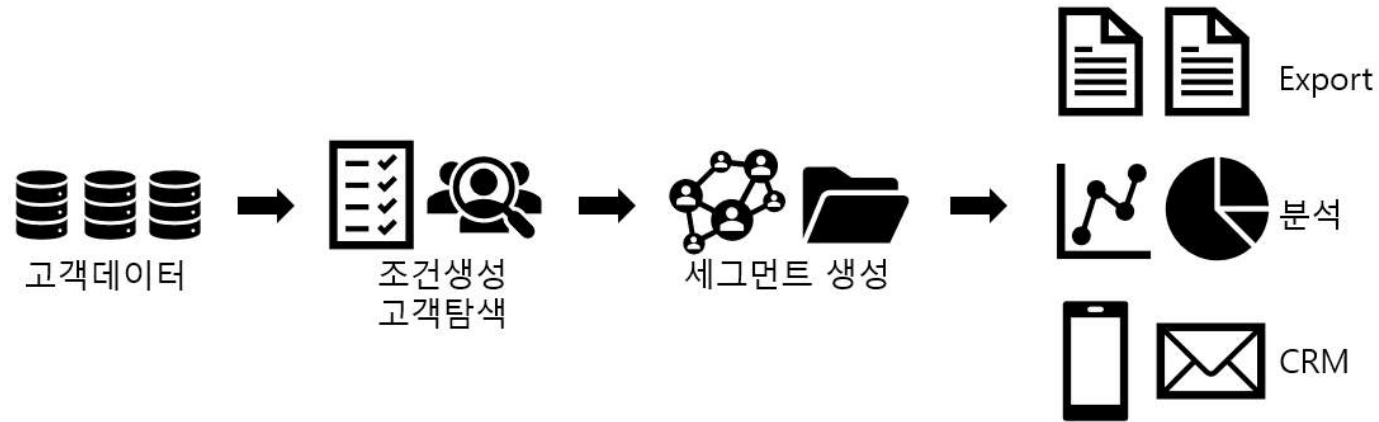


현장의 Needs가 먼저냐?
데이터 준비가 먼저냐?

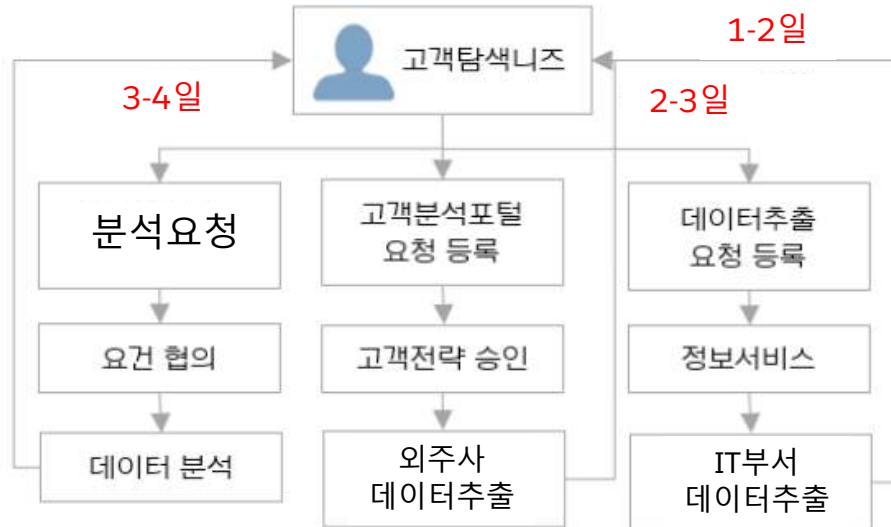


county	year	county_ra	v001_raw	v001_nun	v001_denv	v001_cilov	v001_cihig	v001_race	v001_race	v001_ra
United States	2019		6900.63	3745538	9.08E+08	6890.349	6910.912			
Alabama	2019		9917.233	80440	13636816	9815.191	10019.27			
Autauga County	2019	1	8824.057	815	156132	7935.326	9712.788	10471.25		8706.6
Baldwin County	2019	1	7224.632	2827	576496	6794.128	7655.136	10042.47	3086.606	7277.7
Barbour County	2019	1	9586.165	451	72222	8200.118	10972.21	11332.56		7309.6
Bibb County	2019	1	11783.54	445	63653	10159.98	13407.11	14812.54		11327.5
Blount County	2019	1	10908.1	1050	161107	9895.583	11920.62		5619.645	11336.6
Bullock County	2019	1	12066.91	205	29266	9519.704	14614.12	14336.54		6638.6
Butler County	2019	1	14018.61	393	54804	11976.43	16060.78	15793.58		12797.7
Calhoun County	2019	1	12217.76	2333	321406	11484.59	12950.93	12955.02		12447.3
Chambers County	2019	1	11273.17	691	93770	9948.759	12597.58	10774.82		11794.6
Cherokee County	2019	1	11294.41	575	71014	9767.023	12821.8	16395.25		10965.3
Chilton County	2019	1	10831.92	850	123617	9733.842	11929.99	12000.44		11110.6
Choctaw County	2019	1	12047.19	273	35489	9740.13	14354.24	12215.47		12163.6
Clarke County	2019	1	9688.374	412	67037	8207.626	11169.12	12447.32		7394.5
Clay County	2019	1	9666.041	271	36916	7821.948	11510.13	12628.16		9302.0
Cleburne County	2019	1	11985.56	334	41389	9955.606	14015.5			
Coffee County	2019	1	7933.13	743	143899	7085.258	8781.002	10333.51	6033.742	7652.6
Colbert County	2019	1	10332.48	1016	149595	9321.505	11343.46	11604.94		10386.6
Conecuh County	2019	1	10635.37	272	34170	8546.834	12723.9	11100.61		10319.3

데이터 접근성 : Easy Access & Direct Analysis (CRM case)



AS-IS

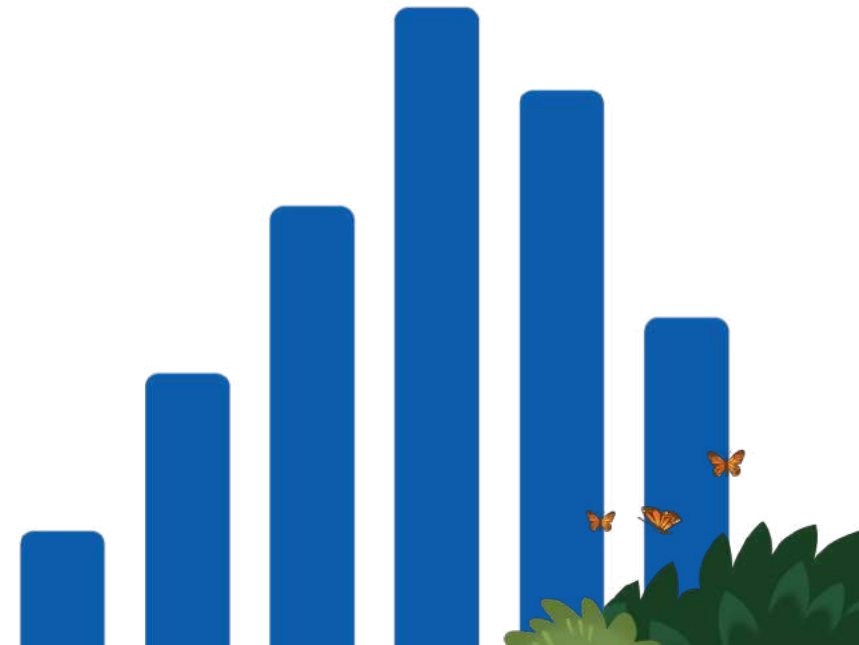


TO-BE



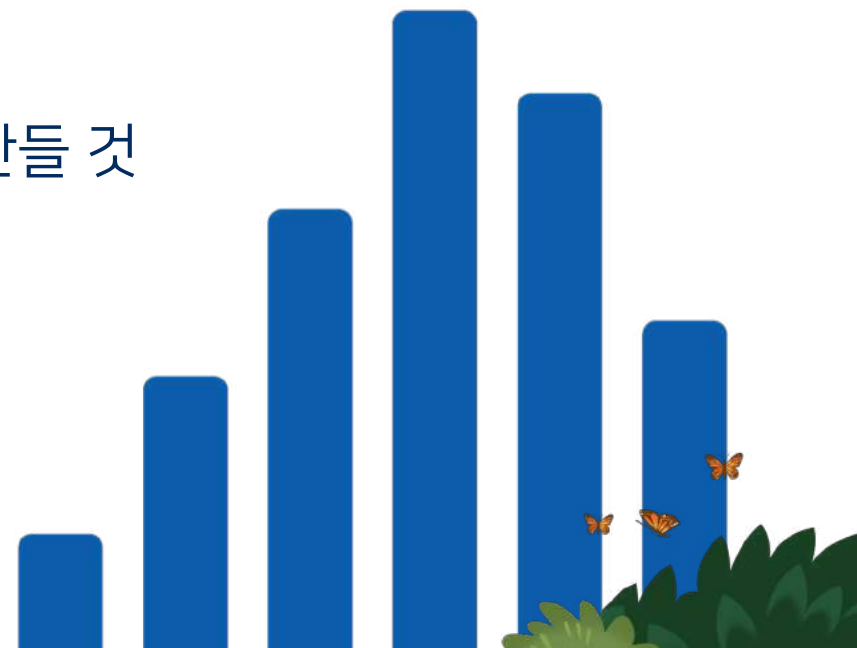
일반 임직원 대상의 데이터 활용 역량의 강화

- 그들의 업무와 연관된 데이터 세트와 마트를 준비하기
- 대시보드와 보고서를 같이 만들어 보기
- 지속적인 활용 교육의 실시 (오프라인, 개인 대상)
- 현장 특화된 소규모 프로젝트 해보기
- (Optional) 사내 데이터 활용 경진대회 해보기



Summary

- 데이터 플랫폼 구축이 끝나면 바로 '활용'을 생각할 것
데이터 활용의 핵심은 단위 조직이 '스스로' 하는 것만이 의미가 있음
- 데이터의 직접 활용을 위하여,
데이터 전문 조직은 데이터 접근성, 데이터 활용 역량을 집중 지원할 것
- 데이터 활용의 조직 변화 관리를 위하여,
CIO/CTO는 각 부서의 리더와 현실적 문제 정의를 위한 구조를 만들 것



Thank you

