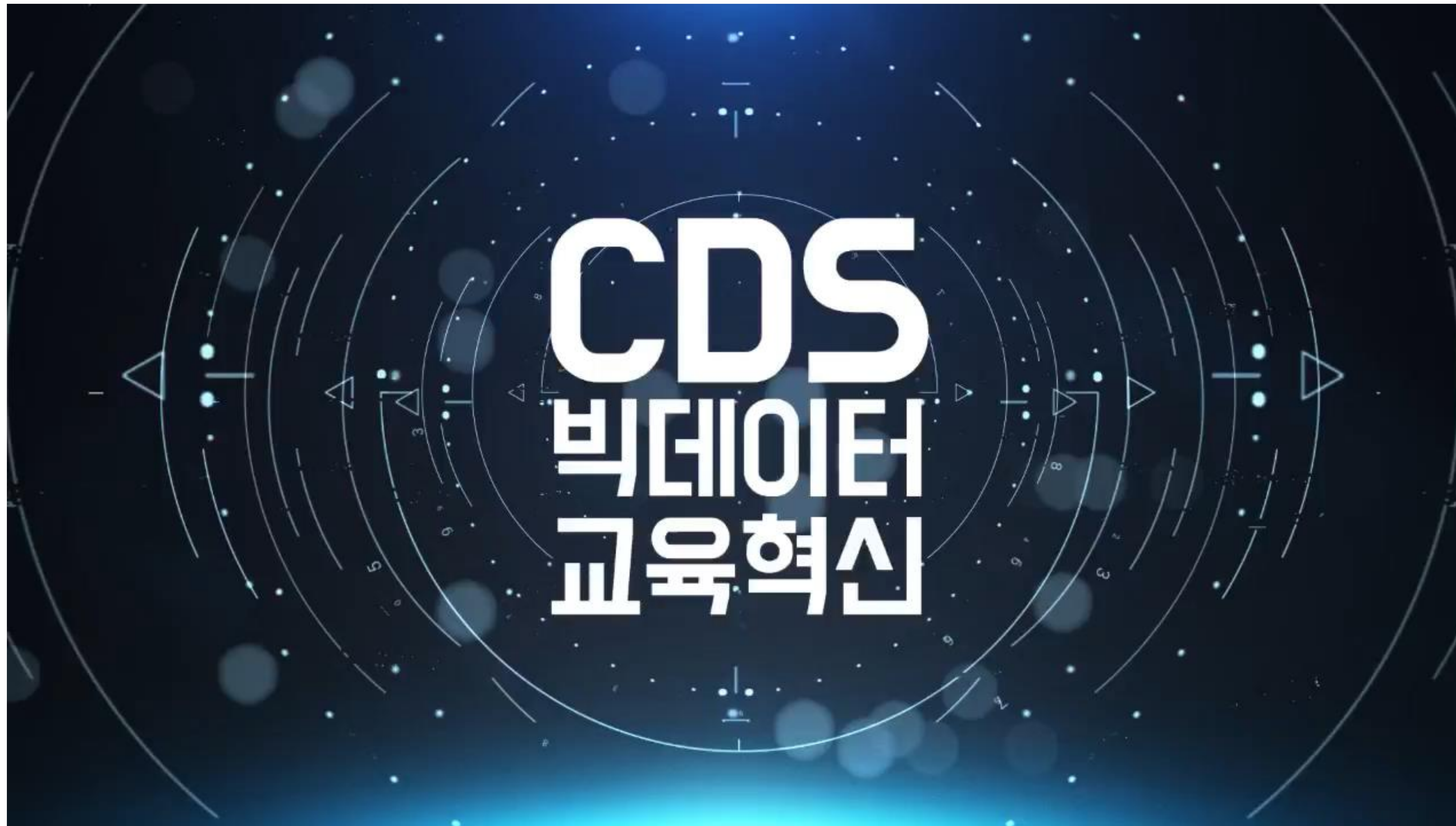


CDS

대학연합을 통한 교육혁신

- 비대면 환경 교육혁신사례

시티즌 데이터 사이언티스트란? (Citizen Data Scientist : 이하 CDS)



CDS 개념의 도입

- 1 추진 배경
- 2 추진 목적
- 3 CDS 육성

CDS 교육혁신의 시작

- 4 CDS 교육혁신
- 5 CDS 빅데이터 교육
- 6 CDS 빅데이터 경진대회

CDS 개념의 도입

Industry 4.0시대는 ICT발달로 인한 자동화/지능화 혁명, 뒤 바뀐 시대의 중심에는 빅데이터가 존재함

4차산업혁명

4차산업혁명이란 “인공지능(AI)/사물인터넷(IoT)/빅데이터(Big-data) /모바일 등 첨단 정보통신기술이 경제/사회 전반에 융합되어 혁신적인 변화가 나타나는 차세대 산업혁명” (한국정보통신기술협회/용어사전)

Industry 4.0의 개념

Industry 4.0의구체적인개념은 “데이터화 → 정보화 → 지능화 → 스마트화”
(4차산업혁명 국가로드맵 정책연구, 2018.05)

Industry 4.0형 교육

Industry 4.0형인력양성을 위한 교육은 빅데이터 관련 교육을 중심으로 확산되고 있는데 이는 4차산업개념인 데이터화 → 정보화 → 지능화의 단계가 바로 빅데이터 생성 → 저장 → 분석의 단계이기 때문임.

경제 · 산업계의 변화의 방향

경제 · 산업계의 변화의 방향은 기술 융합과 진화된 아이디어가 신사업, 신제품 형태로 출시하는 방향과 빅데이터 분석과 인공지능이 결합하여 경영, 마케팅, 생산 분야에서 혁신을 이루고자 하는 방향이 2가지 축으로 나뉘어 진행됨.

빅데이터 인력양성을 통한 창업 활성화 및 새로운 아이디어를 찾기 위해
“**빅데이터 활용 창업 아이디어 경진 대회**” 가 2017년 부터 광범위하게 진행되고 있는 추세



그러나 빅데이터 분석 전문가 과정이 기업/대학/협회/학원 등 광범위하게 진행되고 있음에도 일반인들의 접근이 어려움. (현 교육과정의 99.9%는 오픈소스 R, Python 등을 활용한 코딩 교육, 기본 5개월 이상 긴 시간 투자 필요)

한양대학교와 함께하는 빅데이터 분석 전문가 교육생 모집

4차 산업혁명 선도인력 양성

프로젝트형 IoT를 활용한 빅데이터 분석 전문가 양성과정

교육기간 | 2019년 12월 9일(월)~2020년 5월 8일(금) (총 100일)
 정원 | 20명
 교육내용 | 이론/실습(400시간)
 파이썬 언어, 파이썬 알고리즘, 라즈베리파이 활용, 빅데이터 수집, 클라우드 구축, 빅데이터 처리/분석, 데이터 시각화, Azure 머신러닝
 프로젝트(360시간)
 모바일 IoT 프로젝트, 미니프로젝트1, 2, 산학프로젝트
 직업기초능력(40시간)
 모집마감 | (1차) ~ 2019년 11월 26일(화)
 (2차) ~ 2019년 12월 3일(화)

프로젝트형 머신러닝 기반 빅데이터 분석 전문가 양성과정

교육기간 | 2019년 12월 23일(월)~2020년 5월 22일(금) (총 100일)
 정원 | 20명
 교육내용 | 이론/실습(400시간)
 파이썬 언어, 파이썬 알고리즘, 빅데이터 수집/운영/처리/분석, 데이터 시각화, Azure 머신러닝, 클라우드 빅데이터 컴퓨팅, 빅데이터 머신러닝
 프로젝트(360시간)
 미니프로젝트1, 2, 산학프로젝트
 직업기초능력(40시간)
 모집마감 | (1차) ~ 2019년 12월 10일(화)
 (2차) ~ 2019년 12월 17일(화)

교육대상 | 대한민국 국적 청년구직자 (전공 무관)
 2020년 9월 1일 이전 대학교 졸업(예정)자 (※ 기술업자 포함)
 만 34세 미만자 (1985년 1월 1일 이후 출생자)

교육시간 | 1일 8시간(09:00~18:00) 총 800시간

교육비용 | 교육비 전액 무료 (※ 매월 116천원 후원장려금 지급)

전형절차 | (1차) 서류심사, (2차) 면접심사 (※ 수시 면접)
 ※ 서류심사 통과자는 수시 면접시를 진행교를 방문하여 면접을 볼 수 있습니다. 각양별 지원할 수직을 영우 요망사항이 반영되어도 조가 마감됩니다.

교육특징 | 산학프로젝트 총시간 45% 편성
 한양대 교수진 및 현장 전문가 등 교강사, 보조강사 학습 지원
 담당강사 및 기술(기업) 원투 투입
 MPP(Microsoft Professional Program) 온라인 교육 제공
 22시까지 자율학습 보장
 한양대학교 산학협력단 수수료
 취업상담 상시 지원

교육장소 | 한양대학교 퓨전테크센터(FTC) 11층 1104호
 서울지하철 2호선 한양대학 하차 도보거리 5분 내외
 (한양대 건물번호 208)

문의처 | 한양대학교 SW융합교육센터
 ☎ 02-220-4416, 4417 ● hyswec@hanyang.ac.kr
 ● cafe.naver.com/hyswec

한양대학교 | 고용노동부

교육기간 2019년 12월 9일(월)~2020년 5월 8일(금) (총 100일)	교육기간 2019년 12월 23일(월)~2020년 5월 22일(금) (총 100일)
정원 20명	정원 20명
교육내용 이론/실습(400시간) 파이썬 언어, 파이썬 알고리즘, 라즈베리파이 활용, 빅데이터 수집, 클라우드 구축, 빅데이터 처리/분석, 데이터 시각화, Azure 머신러닝	교육내용 이론/실습(400시간) 파이썬 언어, 파이썬 알고리즘, 빅데이터 수집/운영/처리/분석, 데이터 시각화, Azure 머신러닝, 클라우드 빅데이터 컴퓨팅, 빅데이터 머신러닝
프로젝트(360시간) 모바일 IoT 프로젝트, 미니프로젝트1, 2, 산학프로젝트	프로젝트(360시간) 미니프로젝트1, 2, 산학프로젝트
직업기초능력(40시간)	직업기초능력(40시간)
모집마감 (1차) ~ 2019년 11월 26일(화) (2차) ~ 2019년 12월 3일(화)	모집마감 (1차) ~ 2019년 12월 10일(화) (2차) ~ 2019년 12월 17일(화)

파이썬 언어, 파이썬 알고리즘, 라즈베리파이 활용, Azure머신러닝 등

*일반적인 빅데이터 분석 전문가 교육생 모집 요강

- 결국 빅데이터 전문가는 IT 전문가로 이루어지고, 이는 일반기업 빅데이터 분석을 외부 전문가에게 의뢰해야 하는 상황을 초래.
- **현장 데이터를 가장 잘 알고 있는 현업 구성원을 빅데이터 분석과 활용에서 배제시키는 악순환 발생.** (현재 0.1%를 위한 빅데이터 분석 전문가 과정 → 99.9%의 비전문가를 소외계층으로 구분)

오픈소스 R

```
> PCAout<-prcomp(X[,-7],center=T,scale.=T); PCAout; summary(PCAout)
standard deviations (1, .., p=6):
[1] 1.8159122 1.2207001 0.6771557 0.6162153 0.5612595 0.2430651
```

Rotation (n x k) = (6 x 6):

	PC1	PC2	PC3	PC4	PC5	PC6
population	-0.09760366	0.72204022	-0.48477084	-0.4153580	0.24064887	-0.06078283
white	0.52647569	0.04188256	-0.16805956	0.2973748	0.06695656	-0.77453872
black	-0.49556254	-0.00818906	0.48613708	-0.3980097	-0.03529646	-0.59863499
foreignBorn	-0.46822583	0.03925322	-0.44338277	0.3286252	-0.67199691	-0.15185881
hhIncome	0.46194190	-0.13807073	-0.06093533	-0.6311170	-0.60405975	0.02522163
poverty	-0.18732030	-0.67544976	-0.54783442	-0.2726371	0.34625291	-0.11972521

Importance of components:

	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	1.8159	1.2207	0.67716	0.61622	0.5613	0.24307
Proportion of Variance	0.5496	0.2483	0.07642	0.06329	0.0525	0.00985
Cumulative Proportion	0.5496	0.7979	0.87436	0.93765	0.9901	1.00000

파이썬

```
Python 3.5.3rc1 (v3.5.3rc1:de530d7f21c0, Jan 2 2017, 06:41:25) [MSC v.1900 64 b
it (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> 1 + 2
3
>>> 3 / 2.4
1.25
>>> 3 * 9
27
>>> a=1
>>> b=2
>>> a+b
3
>>> a="Python"
>>> print(a)
Python
>>> a=2+3j
>>> b=3
>>> a*b
(6+9j)
>>> a=3
>>> if a>1
File "<stdin>", line 1
  if a>1
  ^
SyntaxError: invalid syntax
>>> a=3
>>> if a>1: print("a is greater than 1")
File "<stdin>", line 1
  if a>1: print("a is greater than 1")
  ^
SyntaxError: invalid syntax
>>> a=3
>>> if a>1: print("a is greater than 1")
File "<stdin>", line 1
  if a>1: print("a is greater than 1")
  ^
SyntaxError: invalid syntax
>>> a=3
>>> if a>1:
...     print("a is greater than 1")
...
a is greater than 1
```


분명 Industry 4.0 관련 신사업/R&D 등은 빅데이터 분석 전문가 또는 IT 전문가가 활동해야 하는 영역임.

(예를 들어 데이터베이스 전문가, 프로그래밍 언어(C, C++, JAVA, JAVA Script, R, Python) 전문가 등)

빅데이터 기업 및 솔루션

Log Data Apps: splunk, loggly, sumologic

Vertical Apps: PREDICTIVE POLICING, bloomreach

Business Intelligence: ORACLE Hyperion, SAP Business Objects, Microsoft Business Intelligence, IBM COGNOS, SAS MicroStrategy, GoodData

Analytics and Visualization: tableau, Palantir, METAMARKETS, TERADATA ASTER, visual.ly, EMC GREENPLUM, platforma, ClearStory

Data Providers: Gnip, DATASIFT, INRIX

Analytics Infrastructure: Hortonworks, cloudera, MAPR, VERICA

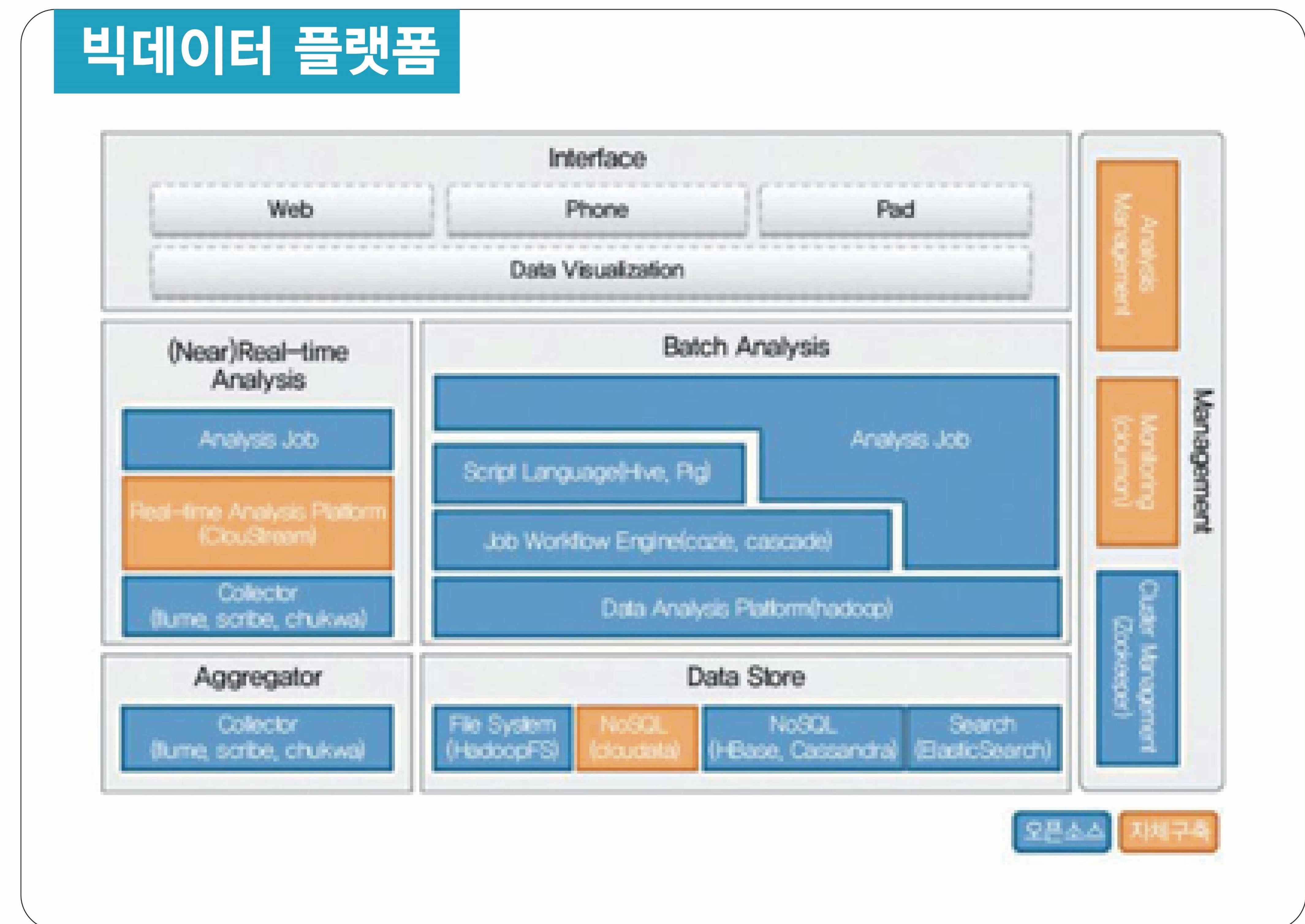
Operational Infrastructure: Couchbase, TERADATA, 10gen, HADAPT

Infrastructure As A Service: amazon, infochimps, Windows Azure

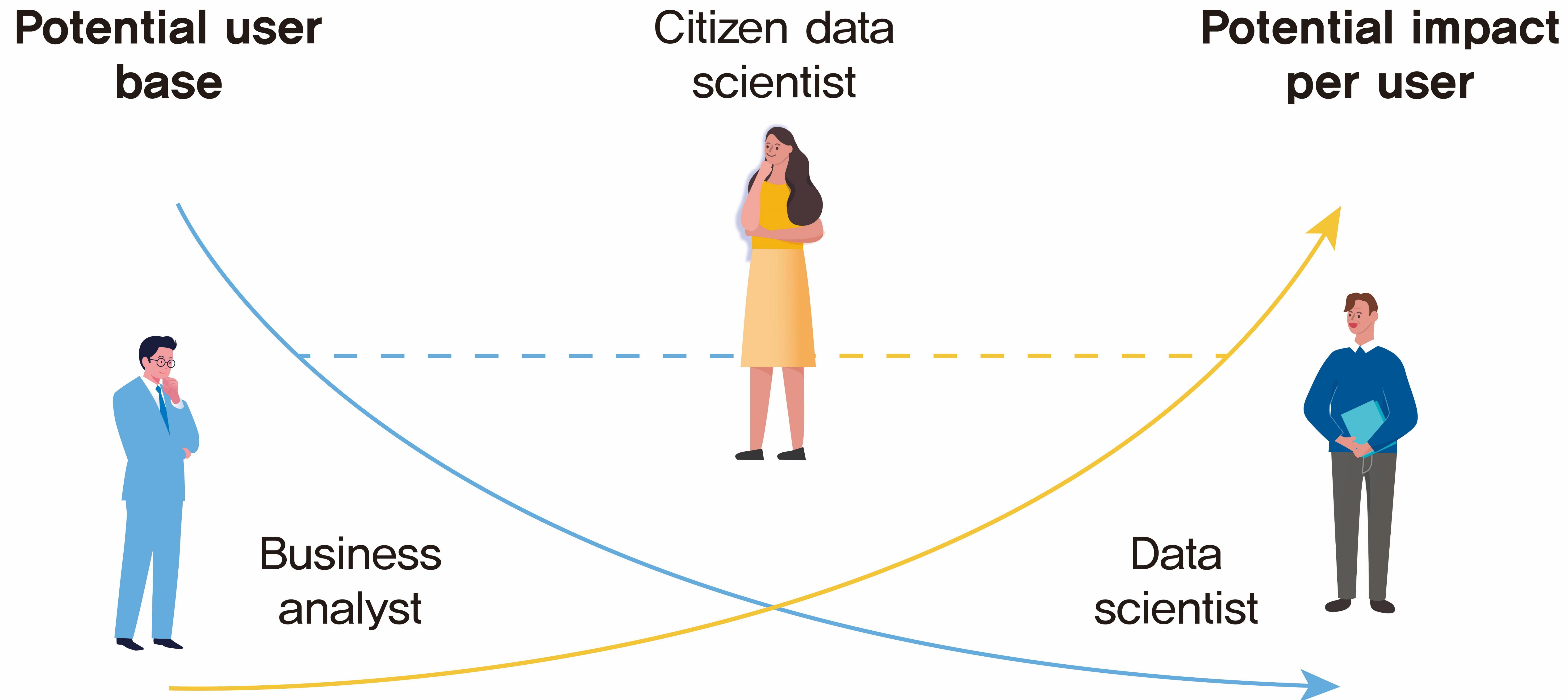
Structured Databases: ORACLE, MySQL, SQL Server, PostgreSQL

hadoop, MapR, HBASE, Cassandra

Copyright © 2012 Dave Feinleb | dave@hcdave.com | <http://blogs.forbes.com/davefeinleb/>



기업의 경영/마케팅/생산 등 폭넓은 영역에서는 실제적으로 현장에서 발생하는 빅데이터를 직접 분석 활용하여 일상적인 혁신 활동 영역으로 진행할 수 있는 현업 구성원인 CDS(Citizen Data Scientist) 육성 절실히 필요.



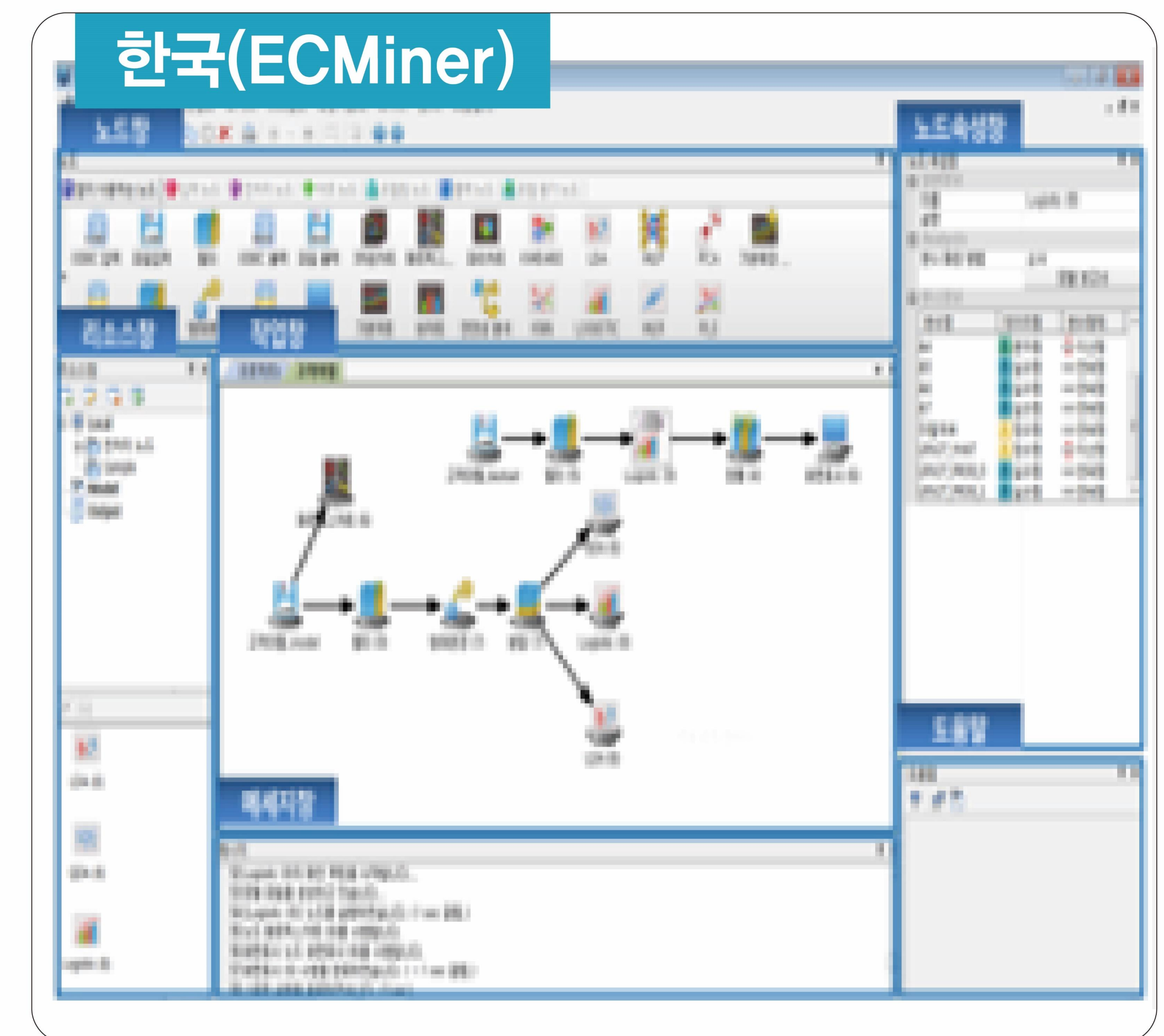
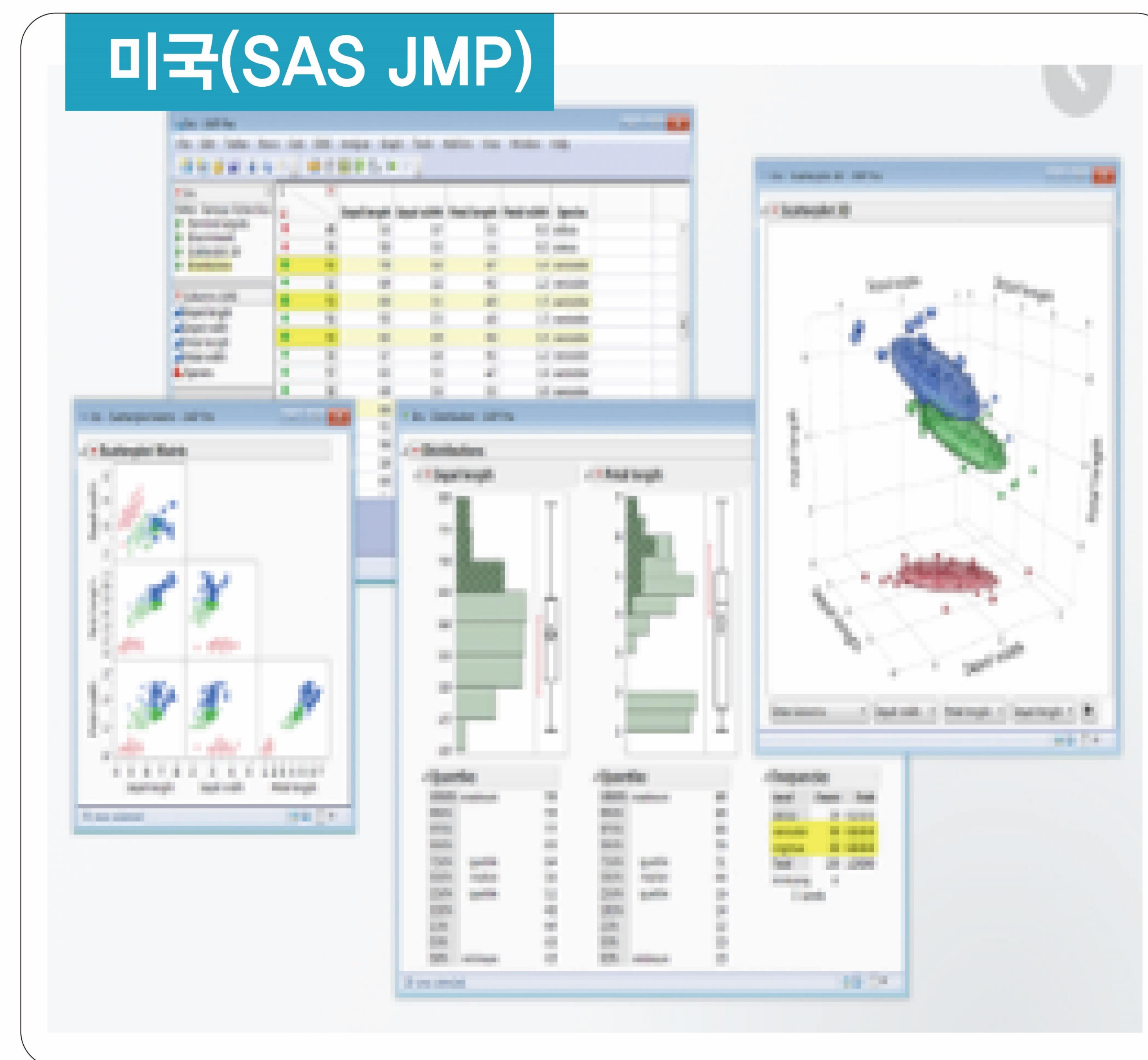
CDS(Citizen Data Scientist)란

“고급 진단 분석 또는 예측 규범적 기능을 사용하는 모델을 만들거나 생성할 수 있으나, 주 업무는 통계 및 분석 분야가 아닌 사람”

이들은 데이터 과학자를 대체하지 않지만, 가설을 세우고 검증하는데 핵심적인 파트너로서 역할을 수행. (가트너 정의)

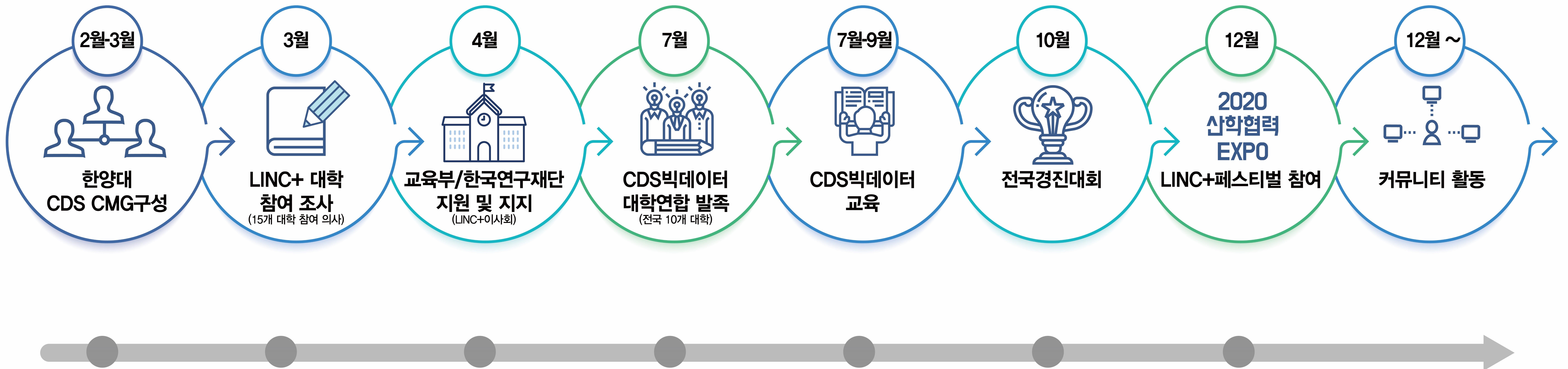
- 빅데이터 도구 및 기술의 도움으로 데이터를 분석하고 회사 데이터 및 비즈니스 모델을 만드는 역할을 하는 조직의 구성원.
- 회사내 현업 구성원으로 비즈니스 및 업계 통찰력을 활용하여 고급 데이터 및 분석 기능을 사용할 수 있도록 함.
- 반드시 데이터 과학 또는 비즈니스 인텔리전스 전문가일 필요는 없음. 그러나 전문가 데이터를 대체하지 않음.
- 시민 데이터 과학자는 중간 정도의 데이터 분석 작업을 수행 할 수 있는 소프트웨어 고급 사용자.
- 드래그 앤 드롭 도구, 사전 구축된 모델 및 데이터 파이프 라인과 같은 소프트웨어 기능을 사용하여 코드없이 모델링함.
- 빅데이터 전문가와 상호 보완시 효과 가장 큼.

CDS는 IT 담당자나 분석가에게 정보 분석을 요청하지도 않고, SQL을 몰라도 쉽게 데이터를 추출하고, 코딩 없이 쉽게 리포트를 만들고 인사이트를 도출할 수 있는 역량, 다시말해 드래그 앤 드롭 방식 분석 솔루션을 다룰 수 있는 역량이 필요.



CDS 교육혁신의 시작

CDS(Citizen Data Scientist)교육혁신 2020년 추진사업

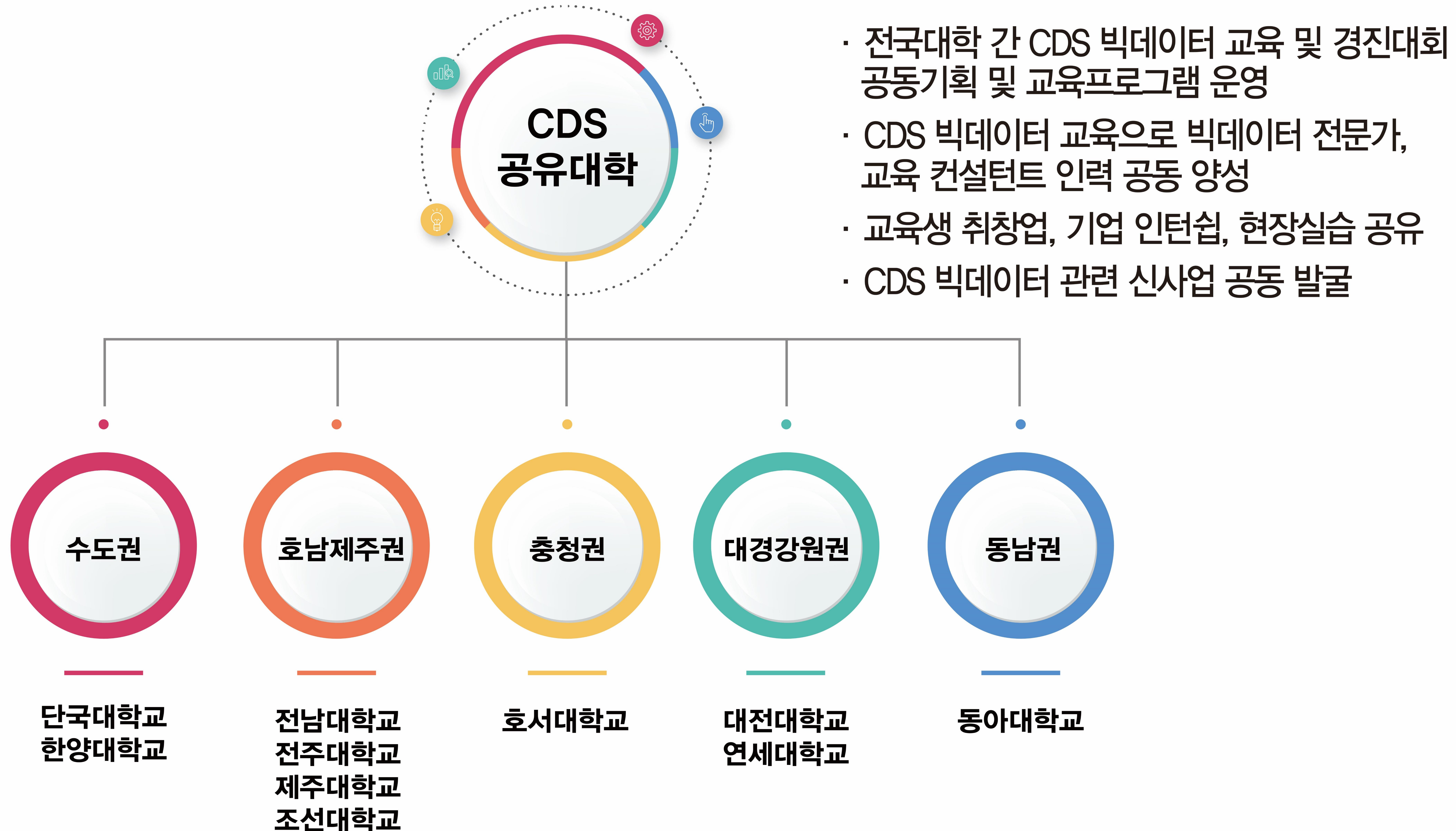


CDS 빅데이터 교육 및 경진대회 추진 발표

- 권역간/대학간 경쟁을 벗어나 협업을 통한 전국 규모의 LINC+사업 대학간 공동프로그램 추진
- LINC+참여대학/교육부/한국연구재단/IPA 및 후원기업의 적극적 지원 및 참여 의사 표명



CDS 빅데이터 대학연합 - 공유대학 구축



CDS 빅데이터 대학연합-공유대학 구축(2020년 7월)



대학연합구축기사

전자신문

한양대 등 10개 대학, 빅데이터 대학연합 구축... 빅데이터
육성

전지연 입력 2020.07.08. 14:54



전병훈 CDS 빅데이터 대학연합 초대
회장

한양대를 포함한 전국 10개 대학이 최근 빅데이터 인재양성을 위해 'CDS(Citizen Data Sc
데이터 대학연합'을 구축했다.

CDS 빅데이터 대학연합은 누구나 쉽게 접근해서 분석·활용할 수 있는 빅데이터 교육을
이 요구하는 빅데이터 인재 양성을 위해 출범했다. CDS는 자신의 전문 분야가 따로 있으
터 가공 능력을 기본 소양으로 갖추고 있는 사람들을 의미한다. △단국대 △대전대 △동
서대(원주) △전남대 △전주대 △제주대 △조선대 △한양대 △호서대가 연합 구축을 위

CDS 빅데이터 대학연합은 7월부터 두 달간 대학생 대상 빅데이터 교육을 실시한다. 10월
진대회를 개최할 예정이다. 내년에는 기업 실무자와 초·중·고 학생 대상으로 교육을 확대

또 4개의 권역별 거점대학을 통해 지역사회와 산업특성에 맞는 빅데이터 기반 플랫폼 사
하고, 궁극적으로 신규 비즈니스 모델을 통한 대학생 취창업 지원 사업 등도 추진한다.

CDS 빅데이터 대학연합 초대 회장인 전병훈 한양대 LINC+사업단장(자원환경공학과 교
이더 교육과 인재양성은 하나의 대학이 주도적으로 할 수는 없다고 생각한다"며 "CDS 부
학연합은 앞으로 우리나라 발전에 주도적인 역할을 할 것"이라고 말했다.

전지연기자 now21@etnews.com

CDS 빅데이터 교육(7~10월)



한양대



대전대



동아대



호서대



단국대



전주대



조선대



제주대

- 교육 참가 대학 : CDS 빅데이터 대학연합(8개 대학)
- 인원 : 대학별 30명 내외(총 223명 교육 수료)
- 시간 : 80hr
- S/W : SAS jmp, ECMiner
- 교육기관 : DNA경영(한양대 CMG 구성원), 생산성본부 등

CDS 빅데이터 교육 특징

01

사용하고 다루기 쉬운 분석솔루션
활용(SAS jmp, ECMiner)

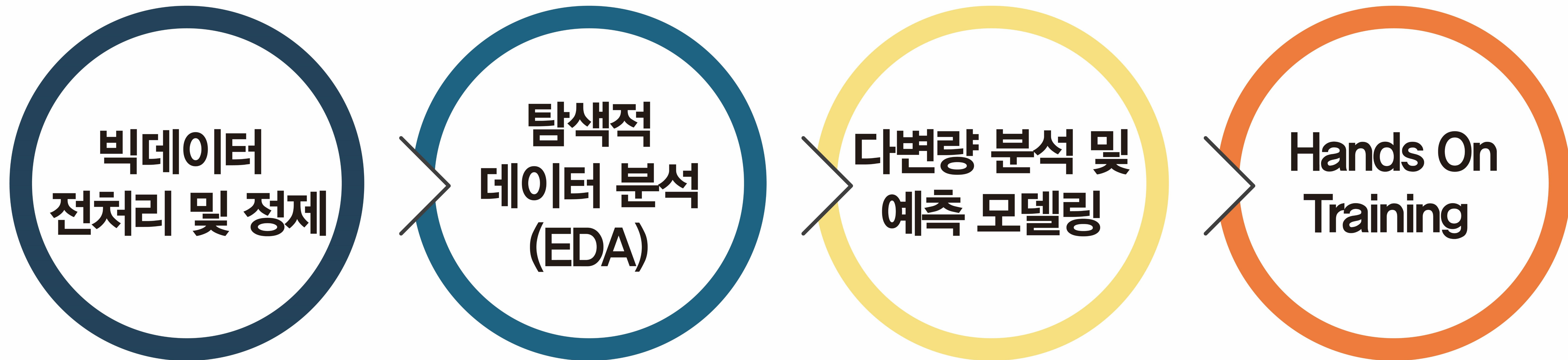
산업 실제 데이터(일부가공)를
활용한 실무적인 교육

02

03

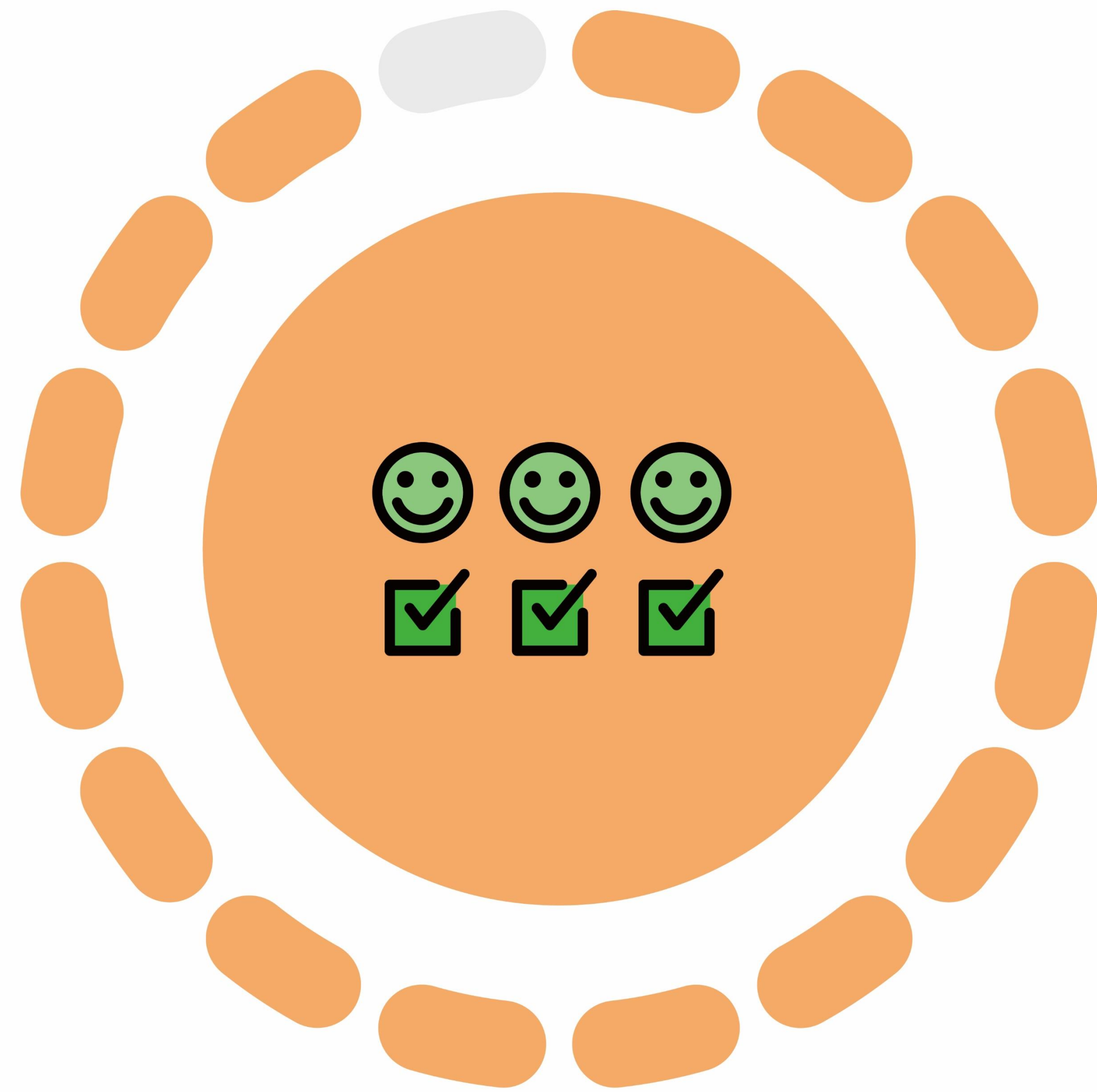
다양한 사례 실습을 통한 분석
알고리즘의 적용 방법 이해

CDS 빅데이터 교육 내용



CDS 빅데이터 교육 설문조사

CDS 교육에 대한 전반적인 만족도는?



교육만족도 | 96%

CDS 교육 내용의 도움되는 수준은?



교육수준 | 96%

CDS 교육을 다른 친구에게 추천한다면?



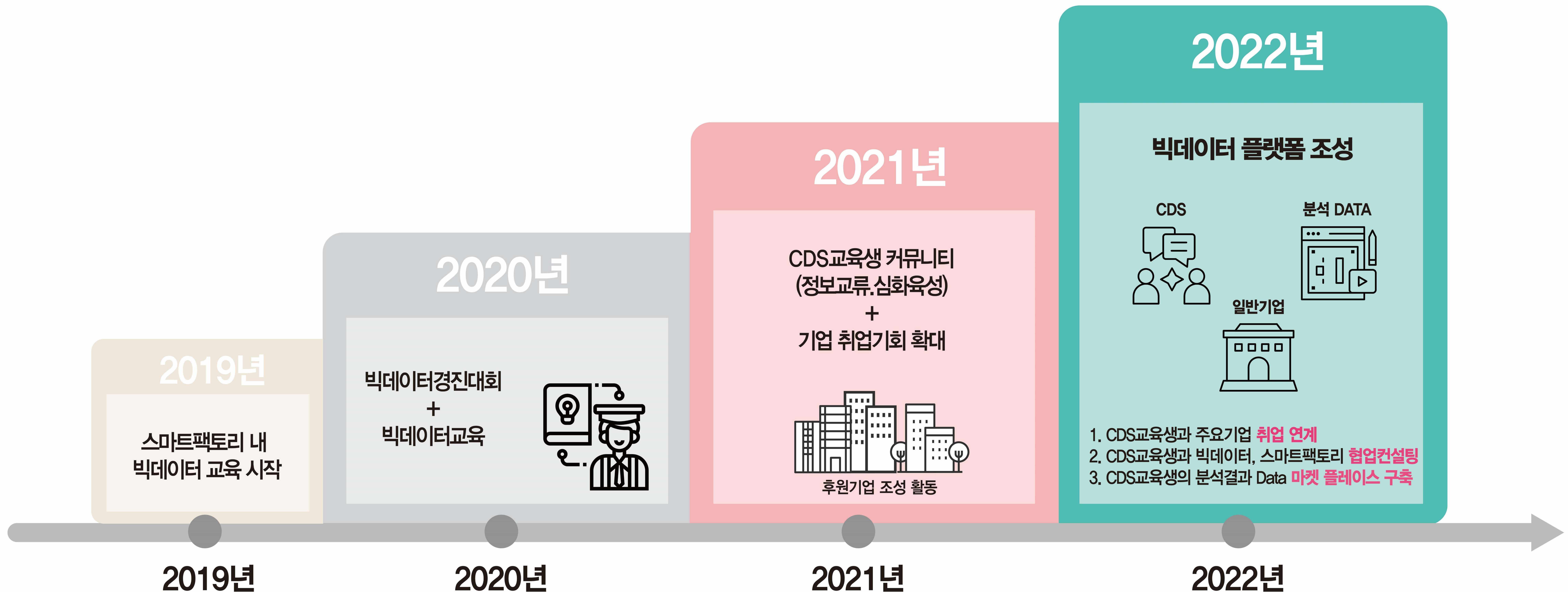
교육추천 | 96%

CDS 빅데이터 경진대회 및 시상



일 시 | 2020년 10월 31일 (토) 08:30~17:00
장 소 | 온라인 경진대회로 각 대학의 강의실에서 진행
참여대학 | 대전대, 동아대, 연세대, 전주대, 제주대, 조선대, 한양대, 호서대
참여인원 | 총 66명

LINC+사업단은 전국대학생 대상 CDS 인재 배출, 정보교류와 심화육성을 위한 커뮤니티 운영, 빅데이터 플랫폼 조성을 통한 주요 기업 취업 연계, 협업 컨설팅, 분석Data 마켓 플레이스 구축으로 단계적 추진.



CDS 빅데이터 경진대회 영상 스케치

