

코멘토 직무 에센스 강의

# 데이터사이언티스트와 함께하는 인공지능 프로젝트 A to Z : 모델링부터 서빙까지



ylee 멘토

### 통계학 학사 전공

- 통계빅데이터 학과 전공
- 경영학과 복수 전공

### 前 스마트팩토리 스타트업 근무

- 센서 기반 이상탐지 모델링
- 헬스케어 포즈 탐지 데이터 정제 및 가공

### 前 CCTV 영상 분석 스타트업 근무

- CCTV 데이터 정제 및 가공
- CCTV내의 사람 이상행동 탐지 및 객체인식 모델링

### 現 디지털 교육 회사 데이터사이언티스트 근무

- 사용자 행동 데이터 기반 시각화 및 인사이트 도출
- 객체인식 및 분류 모델 개발 및 경량화
- 모델 배포와 서비스 모니터링
- 수집되는 데이터 정제 및 가공 관리

- 01 데이터사이언티스트 직무란 무엇인가요?
- 02 데이터사이언티스트 직무는 어떤 커리어패스가 있나요?
- 03 데이터사이언티스트에게 요구되는 역량은 무엇인가요?
- 04 데이터사이언티스트로 일한다는 것은?
- 05 Q&A

Chapter 01 |

# 데이터사이언티스트 직무란?



## 01. 데이터사이언티스트 직무란?

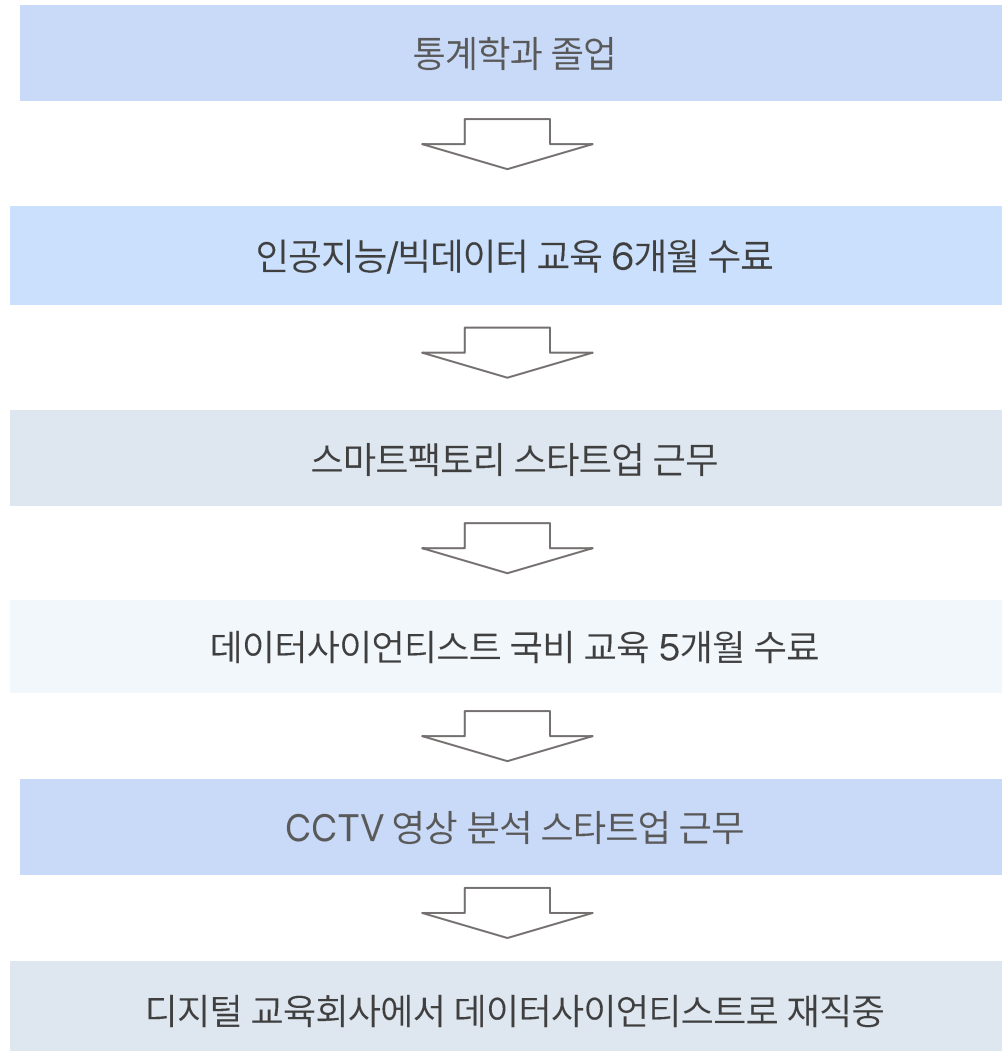
### 데이터사이언티스트란?

- 방대한 양의 데이터를 분석하여 비즈니스 인사이트를 도출하여 문제를 해결하거나 의사결정을 지원하는 전문가.
- 기술적 역량과 비즈니스 도메인 지식을 결합하여 데이터를 활용한 혁신적 솔루션을 제시하는 역할.

### 데이터사이언티스트 가치?

- 데이터를 활용해 기업의 문제를 해결하고, 새로운 기회를 창출.
- 비즈니스 최적화, 고객 예측, 신제품 개발을 통해 기업의 성장을 돕는 핵심 역할.

## 02. 데이터사이언티스트를 시작하게 된 에피소드





### 03. 데이터사이언티스트 하루 일과

08:00 ~ 10:00	<ul style="list-style-type: none"><li>▶ 주간/하루 계획 수립</li><li>▶ 프로젝트별 진행상황 확인 후 TODO 리스트업</li><li>▶ 전주 라벨링 데이터 검수 및 피드백</li></ul>
10:00 ~ 11:30	<ul style="list-style-type: none"><li>▶ 라벨링 현황 업무 보고</li><li>▶ 라벨링 데이터 분석 및 모델링 진행</li></ul>
점심 시간	
12:30 ~ 1:30	<ul style="list-style-type: none"><li>▶ 주간/프로젝트별 진행상황 파악용 팀 미팅</li></ul>
13:30 ~ 15:00	<ul style="list-style-type: none"><li>▶ 모델링 지속 및 모델 성능 확인</li><li>▶ 모델 배포 진행</li></ul>
15:00 ~ 17:00	<ul style="list-style-type: none"><li>▶ 서비스 모니터링 진행</li><li>▶ 배포된 모델로 테스트셋 검증</li></ul>

## 04. 데이터사이언티스트 직무 진입 경로

1

인공지능/데이터 관련 학사, 석사 졸업 후에 데이터사이언티스트로 취업

2

소프트웨어 엔지니어, 마케팅 분석가, 데이터 엔지니어 등 데이터를 다루는 직업에서 데이터사이언티스트로 변경

3

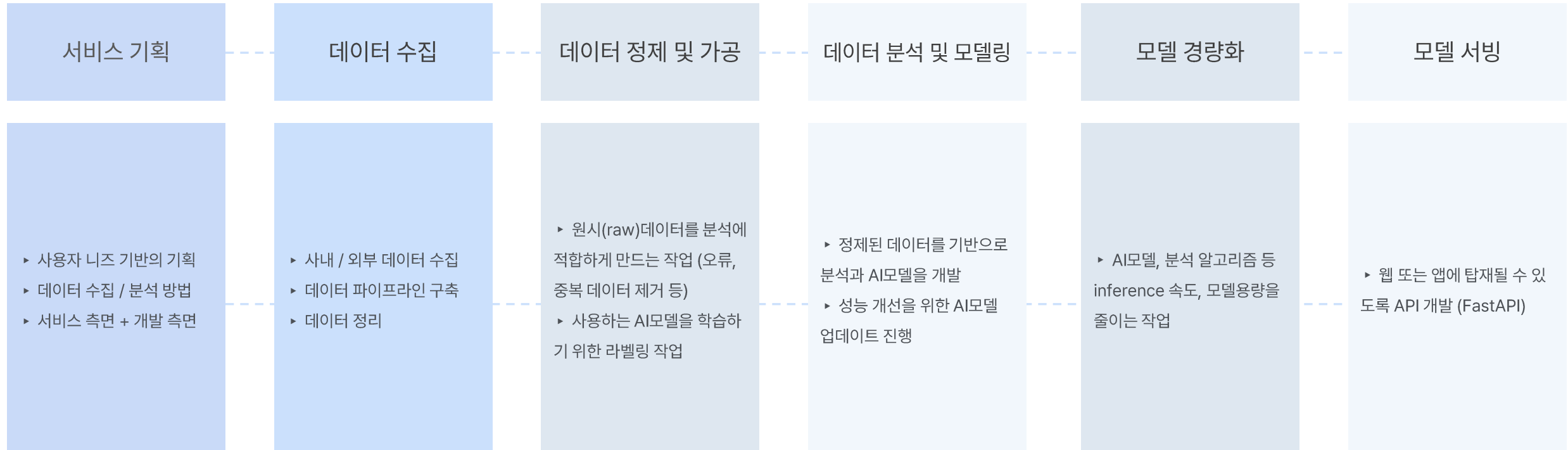
부트캠프 등 다양한 국비 교육을 듣고 데이터사이언티스트로 취업

4

유료 강의(코세라, 유데미 등) 수강 또는 대회(캐글, 데이콘) 참여로 다양한 프로젝트 경험을 쌓은 후, 데이터사이언티스트로 취업



## 05. 데이터사이언티스트가 하는일?



## 05. 데이터사이언티스트가 하는일?

### 서비스 기술 기획

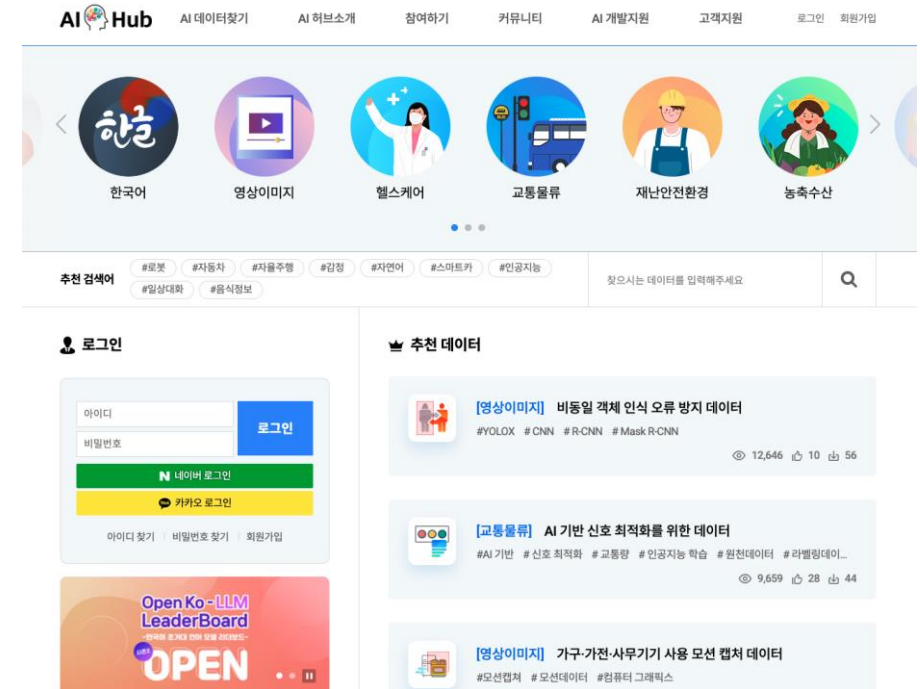
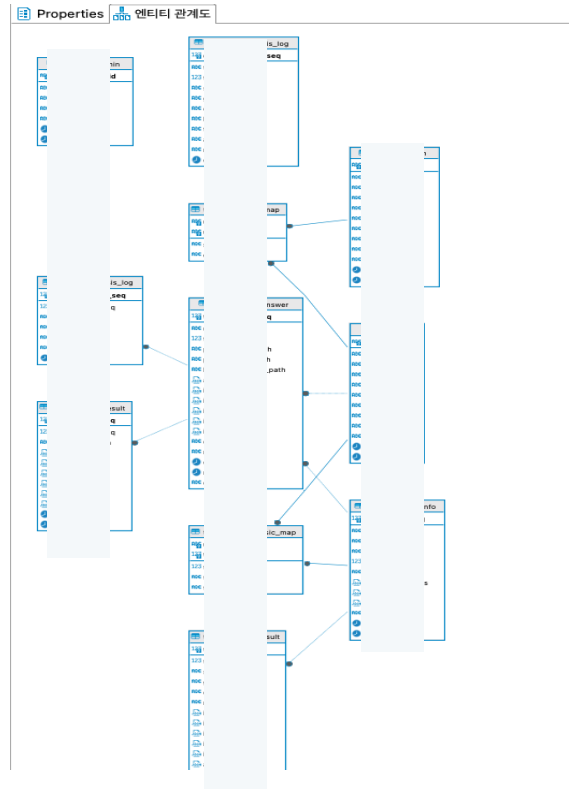
- ▶ 사용자의 니즈를 충족시키기 위해 필요한 데이터와 모델을 정의
- ▶ 서비스 기능 기획
- ▶ 데이터 수집과 정제 방법 정의, AI모델의 라이선스 확인 등
- ▶ 서비스 측면과 개발 측면 모두 고려되어야 하는 부분.



## 05. 데이터사이언티스트가 하는일?

### 데이터 수집

- ▶ 기획에서 개발 방향이 정해지면 데이터 수집 작업이 진행.
- ▶ 이미 데이터를 보유한 회사의 경우 사내 데이터를 활용할 수 있지만, 데이터가 부족하거나 서비스 초반인 경우에는 비슷한 외부 데이터를 사용.



## 05. 데이터사이언티스트가 하는일?

### 데이터 정제 및 가공

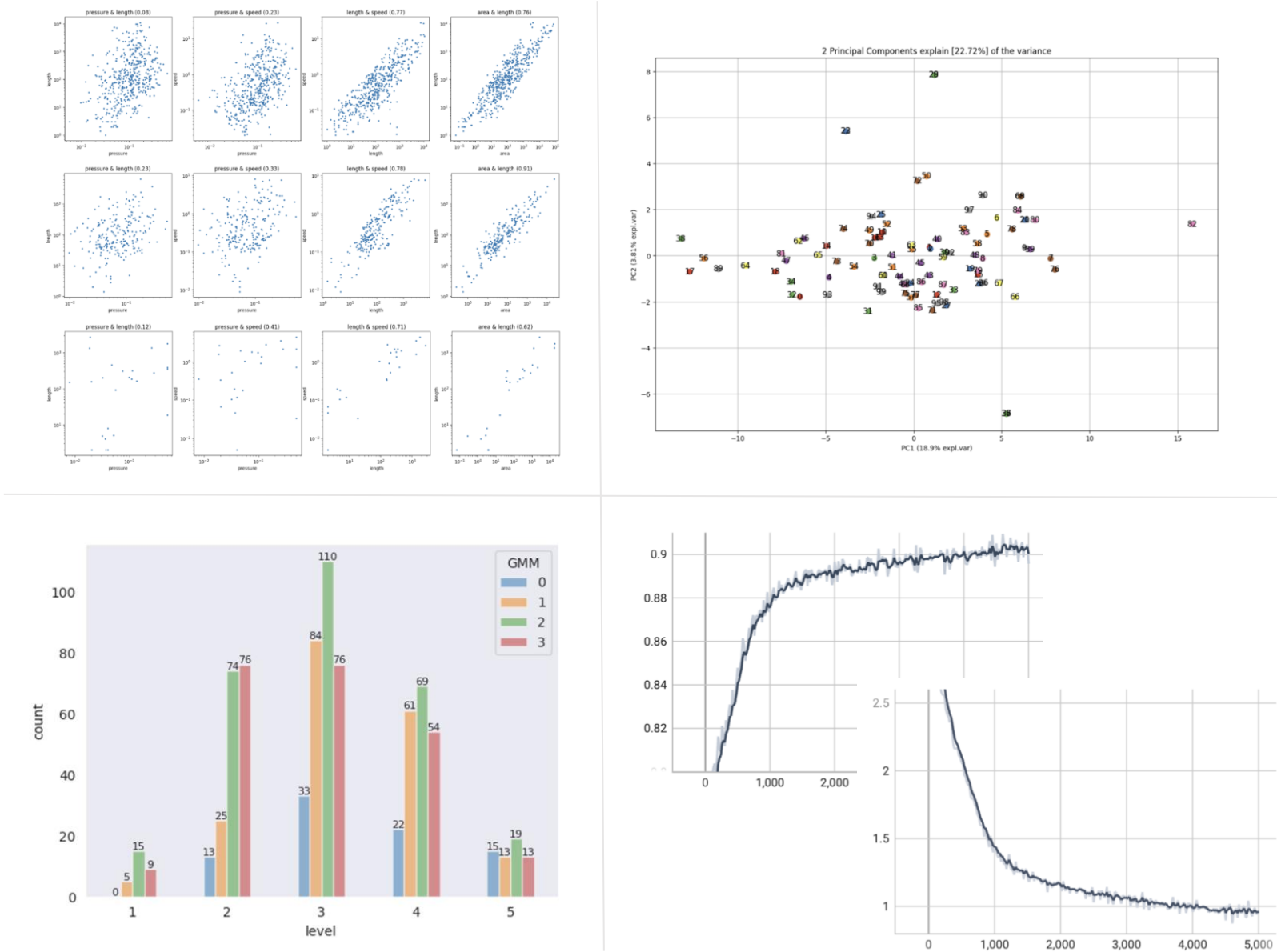
- ▶ 가장 시간이 많이 소요되는 작업
- ▶ 데이터사이언티스트가 직접 데이터를 정제하고 가공할 수는 있지만, 데이터 양이 많아지면 정제 작업에 시간이 많이 소요됨.
- ▶ AI모델 학습에 사용되지 않는 데이터를 걸러내고, 학습에 사용되는 데이터로 가공하는 작업



# 05. 데이터사이언티스트가 하는일?

데이터 분석  
및 모델링

▶ 정제된 데이터를 기반으로 AI모델을 계속 업데이트하며 서비스  
비스를 위한 성능 개선을 주기적으로 진행

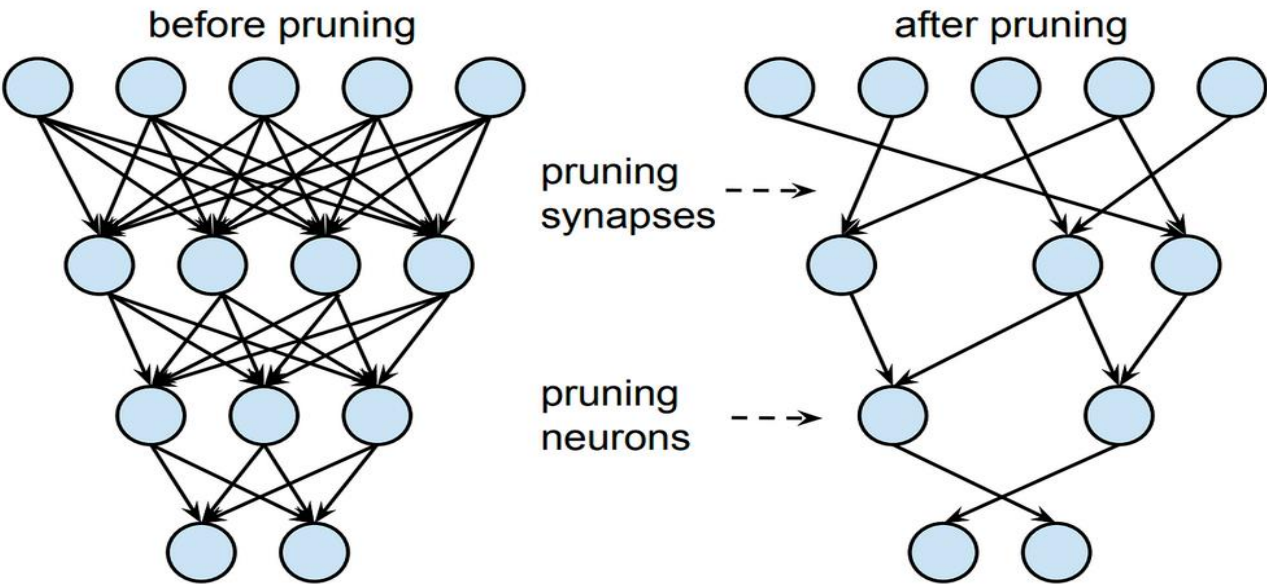




05. 데이터사이언티스트가 하는일?



▶ 딥러닝 모델은 용량이 크고, real time inference가 어려운 경우가 있다.  
이런 문제를 해결하기 위해 경량화 작업이 필요할 수도 있다.

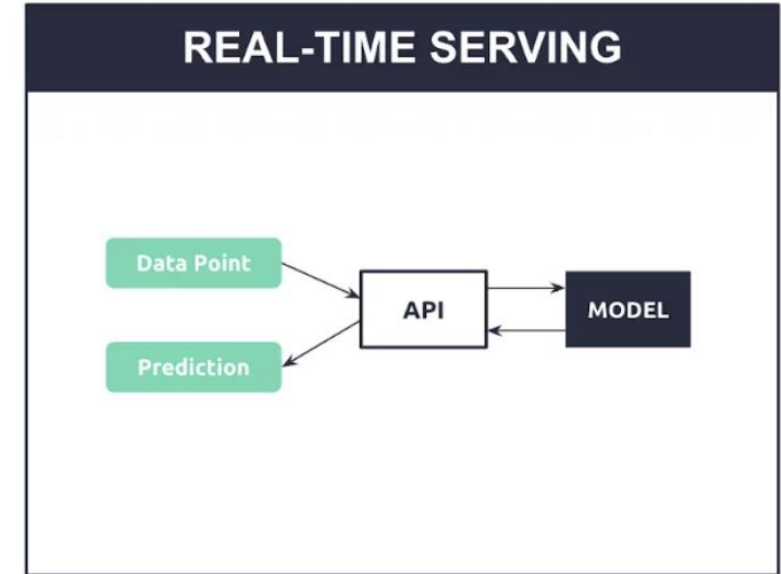
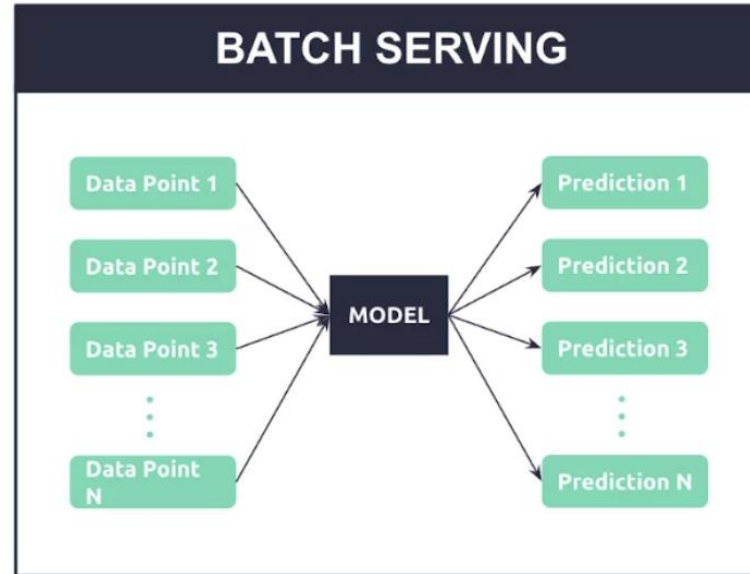


	기존 모델				양자화 경량화 모델				
모델 용량	512.2MB				↓ 1/4 감소	128.3MB			
모델 성능	precision    recall    f1-score					precision    recall    f1-score			
	back	1.00	0.75	0.86		back	1.00	0.62	0.77
	blank	1.00	0.20	0.33		blank	1.00	0.20	0.33
	front	0.94	0.96	0.95		front	0.92	0.99	0.95
	side	0.70	0.73	0.72		side	0.88	0.64	0.74
	accuracy			0.91		accuracy			0.91

## 05. 데이터사이언티스트가 하는일?

모델 서빙

- ▶ 모델이 만들어지면, 사용자가 실제로 서비스를 사용할 수 있도록 데이터 분석을 활용한 후처리 API개발 작업도 같이 진행.



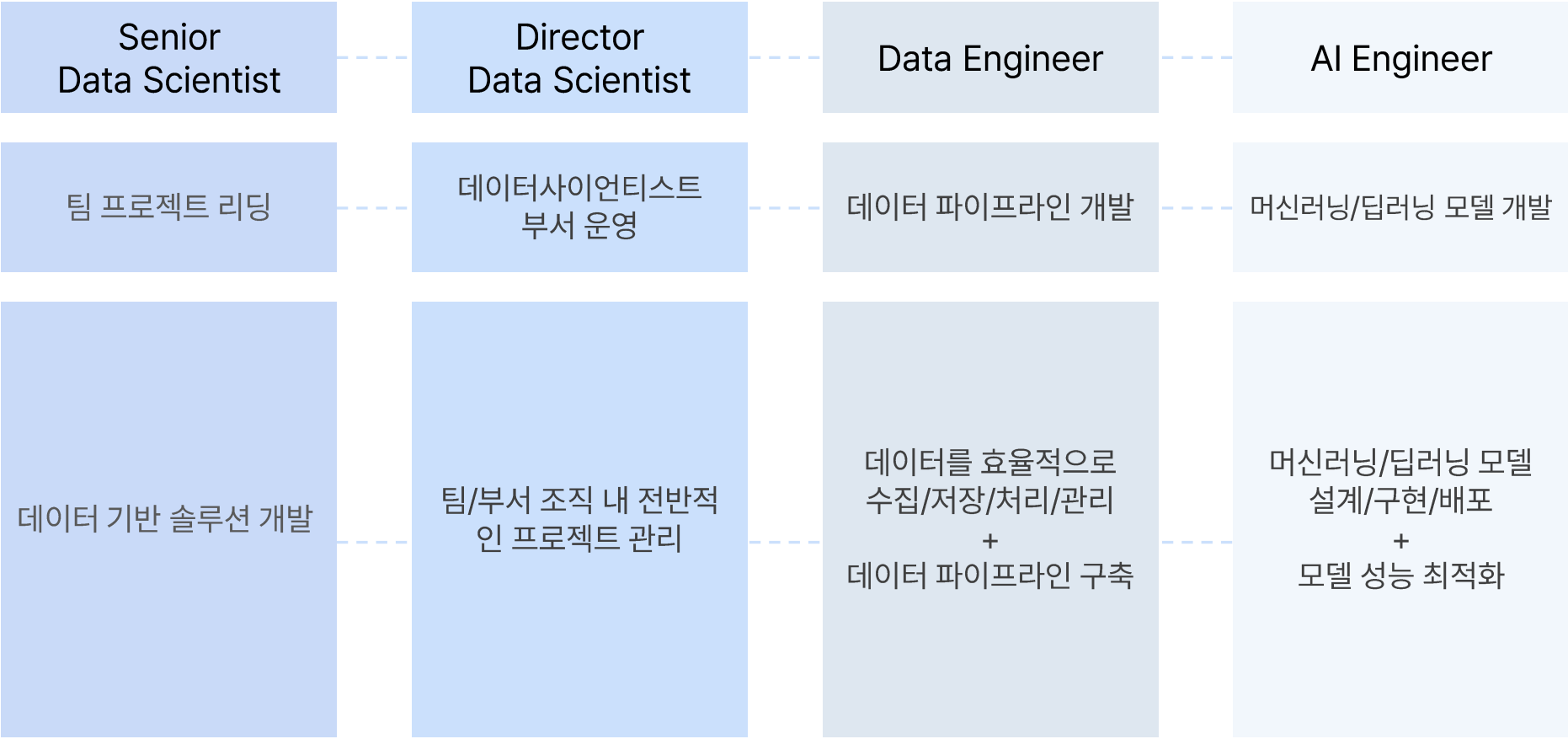
# 데이터사이언티스트 커리어 패스?

시작하는 사람들의 커리어 커뮤니티





01. 일반적인 데이터사이언티스트의 커리어 패스



데이터사이언티스트의 커리어는 종류와 상관없이 기술의 발전에 따라 지속적인 학습이 필수적입니다. 새로운 알고리즘, 기술 트렌드를 따라잡기 위해 온라인 강의, 워크샵, 컨퍼런스 참여 등 꾸준한 자기 개발이 필요한 직업입니다.

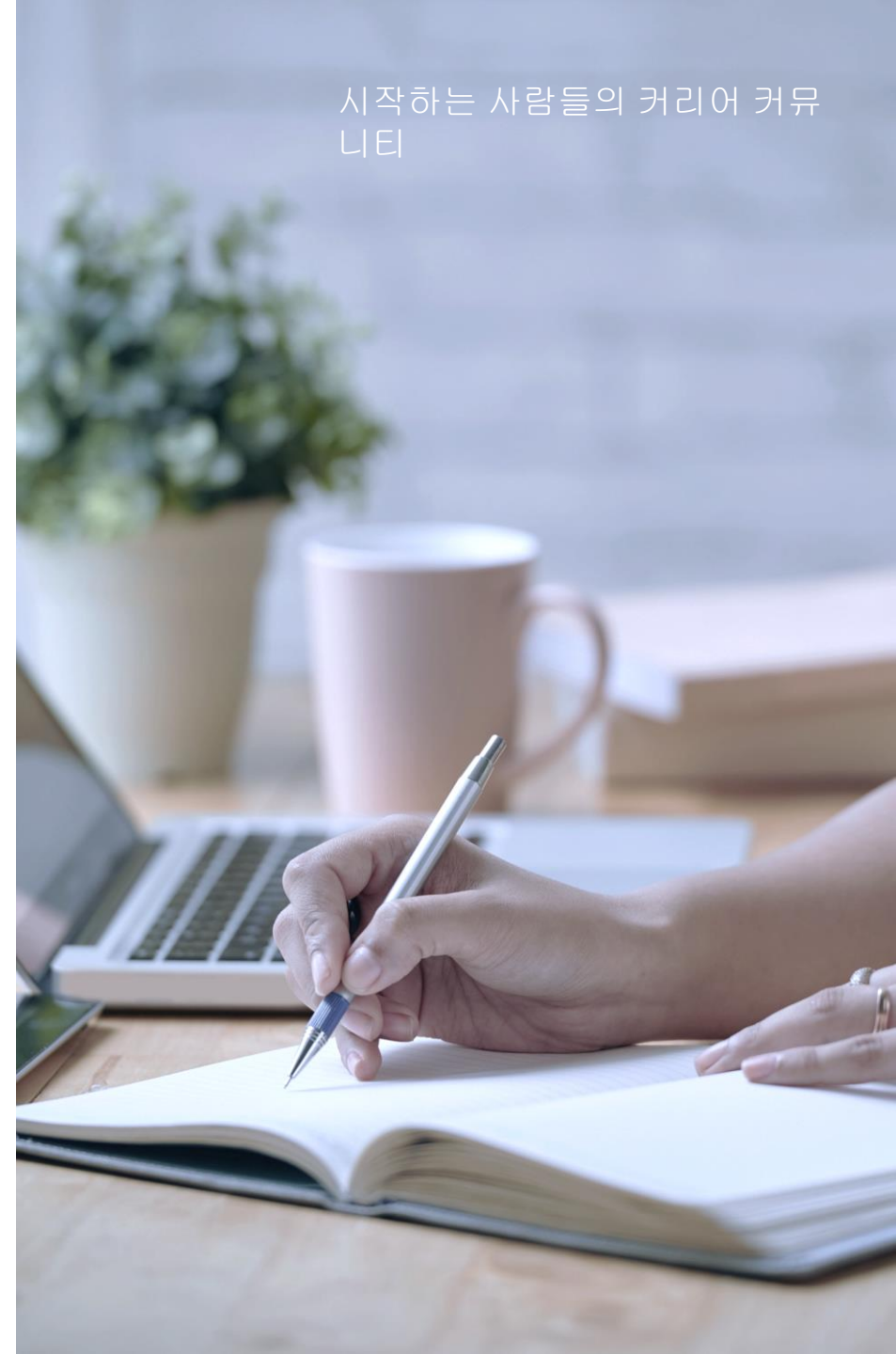
## 02. 제가 목표하는 커리어 패스는...

Director Data Scientist

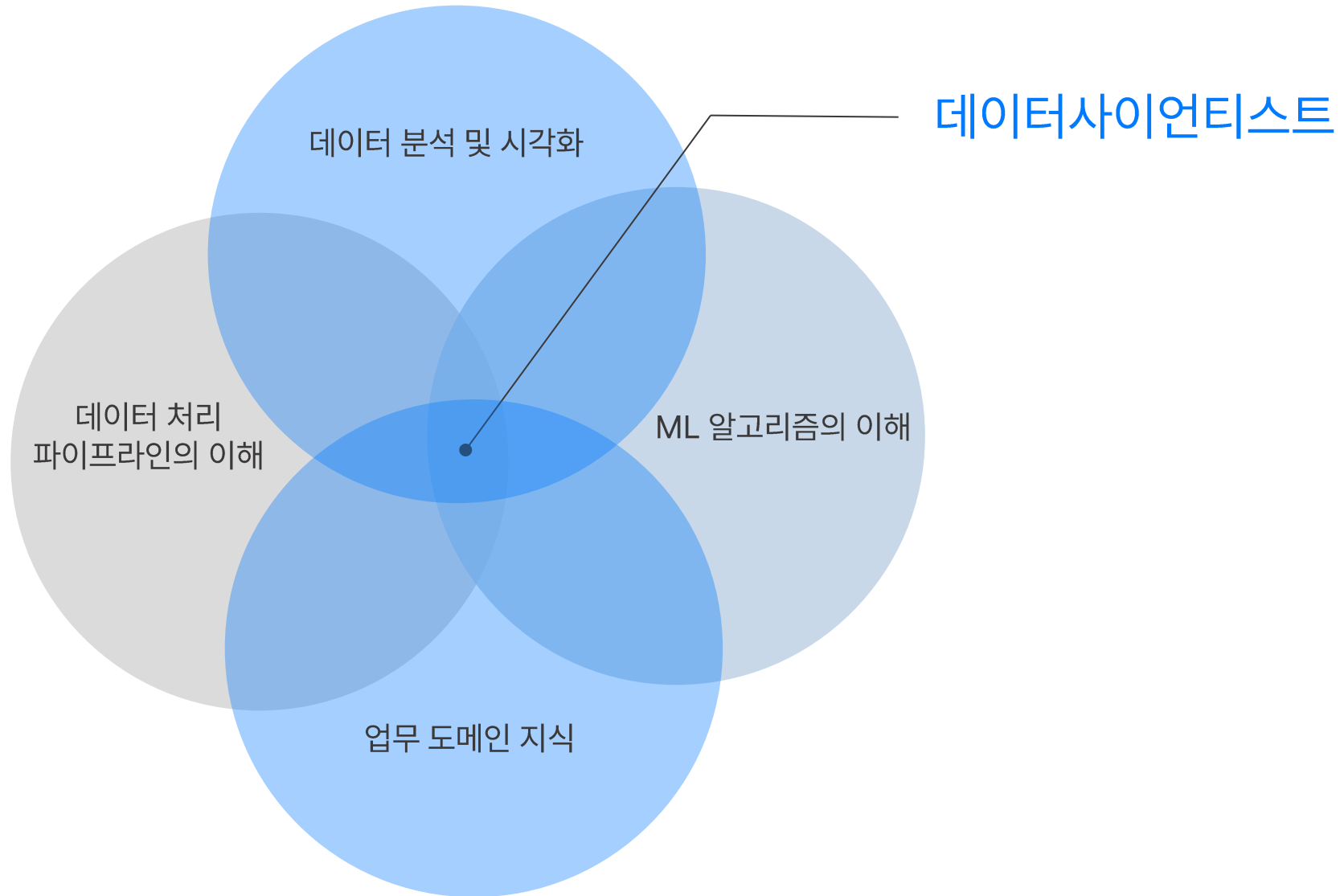
Chapter 03|

# 데이터사이언티스트 역량?

시작하는 사람들의 커리어 커뮤니티



## 01. 데이터사이언티스트 담당자에게 가장 중요한 역량



Chapter 04 |

# 데이터사이언티스트로 일한다는 것은?



## 01. 업무의 장점

### 다양한 산업에서의 적용

1

데이터 사이언티스트의 기술은 특정 산업에 국한되지 않고, 금융, 의료, 제조, 마케팅 등 다양한 산업에서 활용될 수 있습니다. 이로 인해 직업 선택의 폭이 넓어지고, 다양한 경험을 쌓을 수 있습니다.

### 커리어 성장의 기회

2

데이터 사이언스 분야는 빠르게 발전하고 있어, 끊임없이 새로운 기술과 도구를 배우고 적용할 수 있습니다. 이는 개인의 커리어 성장을 촉진하고, 최신 트렌드를 따라가면서 전문성을 높일 수 있는 기회를 제공합니다.

### 직업의 유연성

3

데이터 사이언티스트는 원격 근무가 가능한 직종 중 하나입니다. 데이터 분석은 주로 컴퓨터를 사용하여 수행되므로, 지리적 위치에 구애받지 않고 일할 수 있는 유연성이 높습니다.

### 높은 수요와 보상

4

데이터 사이언티스트는 현재와 미래의 디지털 경제에서 매우 수요가 높은 직업 중 하나입니다. 이로 인해, 많은 회사에서 경쟁력 있는 급여와 혜택을 제공합니다.

## 02. 업무의 단점

### 빠르게 변화하는 기술

1

데이터 사이언스 분야는 기술 발전이 매우 빠른 속도로 이루어지고 있습니다. 최신 기술과 도구를 따라잡기 위해 지속적인 학습이 필요하며, 이는 때로는 부담으로 작용할 수 있습니다.

### 직무의 모호성

2

데이터 사이언티스트의 역할은 회사마다 다를 수 있으며, 명확한 직무 정의가 없는 경우도 있습니다. 이는 업무 범위와 책임이 불분명해지는 문제를 일으킬 수 있으며, 때로는 데이터 분석 외에도 다른 업무를 요구받을 수 있습니다.

### 높은 진입 장벽

3

데이터 사이언티스트가 되기 위해서는 수학, 통계학, 컴퓨터 과학, 머신러닝 등에 대한 깊은 이해와 경험이 필요합니다. 이러한 기술적 요구 사항 때문에 진입 장벽이 높을 수 있으며, 초보자에게는 학습 곡선이 가파를 수 있습니다.

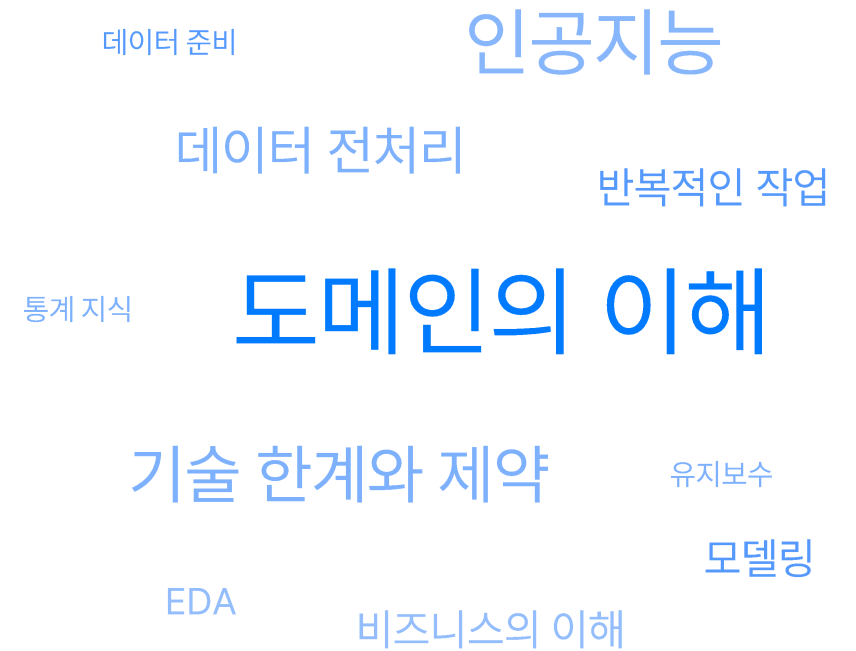
### 복잡한 데이터 처리

4

데이터 사이언티스트는 종종 정제되지 않은 데이터, 누락된 값, 불완전한 데이터셋 등을 다루어야 합니다. 이러한 데이터를 정리하고 전처리하는 데 많은 시간이 소요될 수 있으며, 이는 지루하고 반복적인 작업이 될 수 있습니다.

### 03. 데이터사이언티스트 역할

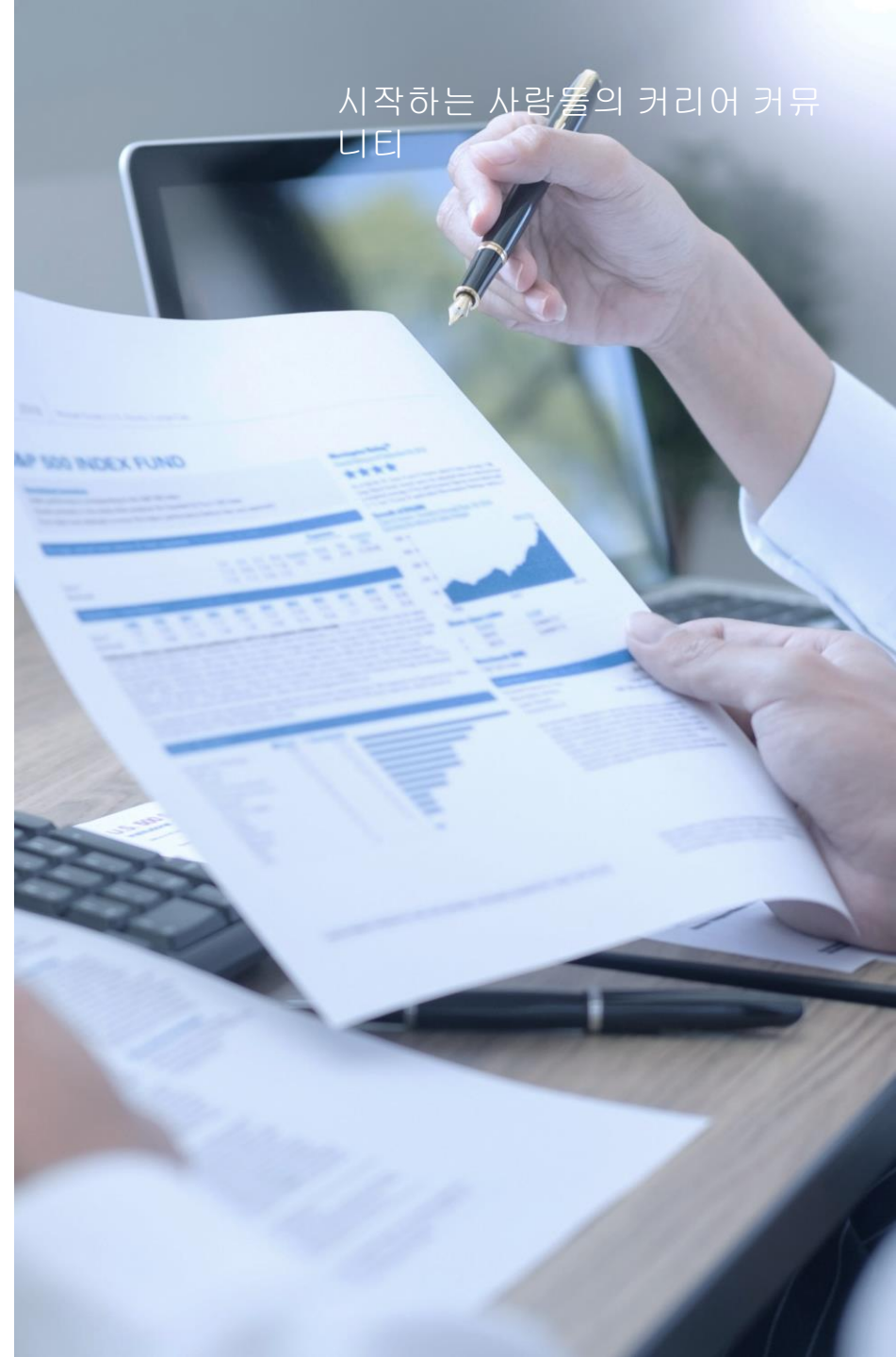
데이터 사이언티스트의 역할은 끊임없이 변화하고 있으며,  
이런 변화를 잘 받아들이고 적용하는 것이 성공적인 커리어를  
만들어가는데 중요한 요소가 된다.





# Q&A

시작하는 사람들의 커리어 커뮤니티



1. 석사를 꼭 졸업해야할까요?
2. 목표하는 회사의 도메인을 전혀 모르는데 괜찮을까요?
3. 데이터 사이언스 역량을 높이기 위해 주로 어떤 방법을 사용하시나요?
4. 프로젝트를 시작할 때 가장 먼저 고려하는 요소는 무엇인가요?