

데이터 과학 단계 >> 퀴즈 풀어보기



# 3. 데이터 과학 단계

데이터 과학 단계 >> 퀴즈 풀어보기

### 데이터 과학 단계

문제 정의 및 목표 설정

데이터 획득 및 수집

데이터 준비 데이터 탐색 및 분석 데이터 표현 및 시각화 모델 개발 및 학습 모델 구축 및 자동화

**Suanlab** • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

1

데이터 과학 단계

>> 퀴즈 풀어보기

#### 문제 정의 및 목표 설정

#### 문제 정의

Problem Definition

- 요구사항 분석(Requirements Specification)
- 자료 수집 및 분석
- 문제 정의서 / 요구사항 분석서

#### 목표 설정

Goal Setting

- 계획 수립(Planning)
- 실행 계획(Action Plan)
- 작업 기술서(Statements of Work)

Suanlab ● 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

## 데이터 획득 및 수집(Data Acquisition and Collection)



**> Suanlab** • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

### 데이터 준비(Data Preparation)



데이터 정제

**Data Cleaning** 

누락 데이터나 잡음, 모순된 데이터 등을 정합성이 맞도록 교정하는 작업



데이터 통합

Data Integration

여러 개의 데이터베이스, 데이터집합 또는 파일을 통합하는 작업



**Data Reduction** 

샘플링, 차원축소, 특징 선택 및 추출을 통 해 데이터 크기를 줄이는 작업



데이터 변환

**Data Transformation** 

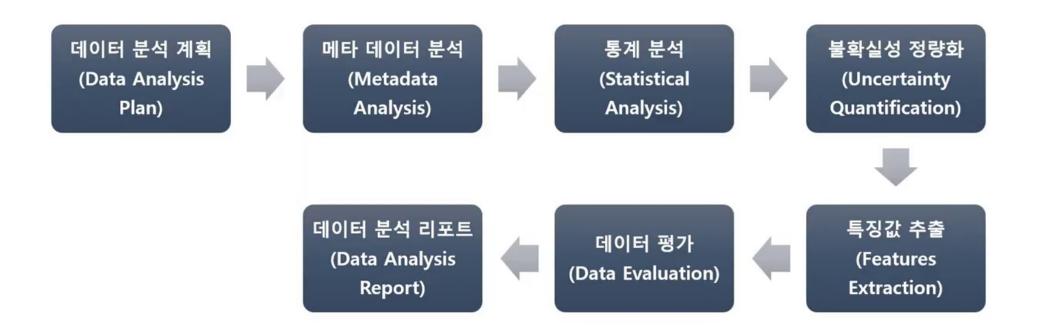
데이터를 정규화, 이산화 또는 집계를 통 해 변환하는 작업



Sugnigb • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

데이터 과학 단계

### 데이터 탐색 및 분석(Data Exploration and Analysis)



Sugnigb • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

### 데이터 표현 및 시각화(Data Representation and Visualization)

- 메타데이터
- 데이터 차원 분석
- 데이터 관계 분석

데이터 분석

스토리텔링

- 스토리 계획
- 유의미한 정보 정의

• 텍스트, 시간, 분포, 관계, 비교, 공간 시각화

시각화 기법 선택

디자인

- 이해와 공감
- 심플하게
- 형상화 & 색상

Suanlab • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

### 모델 개발 및 학습(Model Development and Learning)

#### 데이터 준비

- •특징 추출 (Feature Extraction)
- •데이터 분리 (Training / Test / Validation)
- •데이터 검증 (Data Verification)



#### 모델 생성

- •모델 설계 (Model Design)
- •특징 엔지니어링 (Feature Engineering)
- •모델 구조 정의
- •목표 변수(목적 함수) 정의



#### 모델 학습

- 학습 알고리즘 (Learning Algorithm)
- 하이퍼파라미터 (Hyperparameter)
- 학습 반복(Learning Iteration)
- 최적화(Optimization)



#### 모델 평가

- 평가 지표 (Evaluation Metrics)
- •성능 평가 (Performance **Evaluation**)
- 비교 분석 (Comparative Analysis)

Sugnigb • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

### 모델 구축 및 자동화(Model Deployment and Automation)



https://blog.inedo.com/best-deployment-automation-tools

Suaniab • 데이터 과학(Data Science) - 02 데이터 과학이란 - 3. 데이터 과학 단계

12

데이터 과학 단계



데이터 과학 단계 >> 퀴즈 풀어보기