

Chapter 05. 데이터 전처리와 탐색

- 데이터 전처리와 EDA란 무엇인가?



5-1. 코드실습 : 보험료 데이터

■ <https://colab.research.google.com/drive/1BAXhkTso7SH-Z7CNMjEmgEi3uTJXqSkC?usp=sharing>



학습정리 #5차

■ 데이터 전처리

- 분석 목적에 적합하도록 데이터를 사전에 가공 및 정리하는 과정
- 데이터의 결측치 및 이상치 확인 및 데이터의 형태 전환 등의 프로세스가 진행

■ EDA(Exploratory Data Analysis, 탐색적 데이터 분석)

- 데이터 분석 또는 인공지능 모델링에 전 단계로 데이터의 패턴을 발견하거나 특이성을 찾는 데이터 이해의 단계
- 통계적인 방법 또는 시각화 도구를 활용해서 데이터를 확인하는 과정

■ 예시: 보험료 데이터

- 데이터 전처리 : 데이터의 특징 확인(데이터 포맷, 자리 수 등), 결측치 및 이상치의 확인
- 탐색적 데이터 분석 : 데이터의 통계적 특성(평균, 표준편차, 중위 수 등), 각 변수(특성 값) 간의 관계 파악 등