

교육 목표

1. Python의 개발 환경과 기본 문법 그리고 핵심 자료형을 파악하고 필수 라이브러리를 설치할 수 있다.
2. NumPy 라이브러리가 제공하는 다차원 배열과 행렬을 익히고 고수준의 수학 기능을 사용할 수 있다.
3. pandas 라이브러리의 데이터형과 필수 기능을 학습하고 다양한 수학과 통계 기능을 활용할 수 있다.
4. 다양한 형식으로 저장된 데이터를 읽거나 쓰고, 데이터를 변환하거나 결합하는 방법을 습득한다.
5. 주어진 데이터를 요구되는 형태로 시각화하여 효과적으로 표시하고 이를 파일에 저장할 수 있다.

개요

현대 사회는 유용한 데이터의 누적으로 이를 처리하고 분석하기 위한 요구가 증가하고 있다. Python은 간결한 문법과 풍부한 라이브러리를 제공함으로써 데이터를 분석하고 시각화하는 데 최적의 언어로 주목받고 있다.

이 과정에서는 Python의 기본 문법과 핵심 자료형을 토대로 데이터 분석과 시각화에 필요한 핵심 라이브러리의 사용법을 익힌다. 또한, 다양한 데이터 포맷을 처리하고 이를 분석하는 데 필요한 이론을 학습하고 실례를 다룬다. 수강자는 방대한 데이터를 분석하는 데 필요한 이론적 토대와 필수 도구를 익힘으로써 데이터 분석과 관련된 업무의 효율성을 높이고, 머신 러닝을 비롯한 최신 이슈를 학습하는 데 요구되는 기술적 기초를 다질 수 있다.

학습 대상

- Python 언어 초보자
- 소프트웨어 개발자
- 데이터 분석가

선행 수준

- 선수과정 : 없음
- 사전지식 : Python 기초

Python을 이용한 데이터 분석과 시각화



시간	월요일	화요일	수요일	목요일	금요일
오전				<p>[NumPy 활용]</p> <ul style="list-style-type: none"> - 다차원 배열 객체 - 배열 프로그래밍 - 배열 기반 파일 I/O - 선형 대수 기능 - 의사 난수 생성 	<p>[데이터 파일 처리]</p> <ul style="list-style-type: none"> - 텍스트 형식 데이터 - 바이너리 형식 데이터 - 웹 데이터 - DB 데이터
중식					
오후				<p>[pandas 활용]</p> <ul style="list-style-type: none"> - pandas 데이터 구조 - pandas 필수 기능 - 수학 기능 - 통계 기능 	<p>[데이터 시각화]</p> <ul style="list-style-type: none"> - matplotlib : 그림, 색상, 라벨 : 그림 파일 저장 - pandas와 seaborn : 선, 막대, 분산 : 히스토그램