

Learning Data Analysis and Visualization Easily with Pandas

<https://product.kyobobook.co.kr/detail/S000212266036>

파이썬 데이터 분석 시작 : 데이터 분석 자동화, 다양한 라이브러리 활용

주피터 노트북 설치 : 개발 환경 구축, 코드 실행 및 시각화 편리

주피터 노트북 사용 : 코드 작성, 실행, 시각화, 문서화 가능

넘파이(Numpy) 기초 : 배열 기반 연산, 리스트보다 효율적

Numpy와 리스트 비교 : 메모리 효율, 브로드캐스팅 지원

Numpy 속성 : shape, dtype, ndim, size

Numpy.arange() : 지정 범위 배열 생성

Numpy 통계 함수 : mean, sum, std, min, max 등

인덱싱, 슬라이싱, 반복 : 배열 요소 접근, 부분 선택, 반복 처리

논리적 인덱싱 : 조건에 맞는 요소 선택

2차원 배열 인덱싱/슬라이싱 : 행/열 단위 선택

배열 결합과 분리 : concatenate, transpose, split 활용

특별한 행렬과 벡터 : zeros, ones, full, eye, random

판다스 시리즈 : 1차원 데이터 구조, 인덱싱 지원

시리즈 생성 : list, dict, ndarray → Series

인덱싱/슬라이싱 : 값 선택, 부분 데이터 접근

변경/추가 : series 값 수정, 새 항목 추가

index와 values : 시리즈 구성 확인, index 재설정 가능

정렬 : sort_index(), sort_values()

주요 메서드 : head(), tail(), unique(), nunique(), value_counts()

데이터프레임(DataFrame) : 2차원 데이터 구조, 행/열 접근 용이

생성과 이름 변경 : dict, list, ndarray → DataFrame, 컬럼 이름 지정

행/열 접근 : loc, iloc, 열 이름, 인덱스 이용

복사 : DataFrame.copy()

연산 : 산술 연산, 브로드캐스팅

행/열 삭제 : drop()

논리적 인덱싱 : 조건 기반 필터링

정렬 : sort_values(), sort_index()

데이터 타입 : category, datetime 처리

CSV 파일로 DataFrame 생성 : pd.read_csv()

데이터 요약 : describe(), info()

결측치 처리 : isna(), dropna(), fillna(), 단일/다중 대체

문자열 처리 : 기본 파이썬 함수, str 액세서, 정규 표현식

str.contains(), str.startswith(), str.endswith() : 문자열 조건 필터링

apply(), agg(), map(), applymap() : 함수 적용 방법 비교

데이터 결합/병합 : concat(), merge()

그룹 집계와 멀티인덱스 : groupby(), swaplevel(), droplevel(), transform(), unstack(), stack(), crosstab(), pivot(), pivot_table()

데이터 시각화 : 데이터 이해, 분석 결과 표현

pandas plot() : line, box, bar, pie, hist, kde, scatter 그래프

matplotlib : Figure 생성, subplot, 다양한 그래프 그리기

Seaborn : 고급 시각화, 예제 데이터 활용, countplot, barplot, histplot, boxplot, violinplot, scatterplot, relplot