

IT와 비즈니스혁신

W4. 파이썬 기본 II



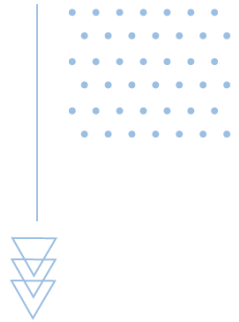
Contents

I . 주요 패키지 소개

- (심화) NumPy, SciPy
- Pandas, Matplotlib

II . 데이터 탐색





I. 주요 패키지 소개

- NumPy, SciPy
- Pandas, Matplotlib

NumPy (넘파이)

- 과학 계산을 위해 필요한 패키지로 다차원 배열을 위한 기능과 선형 대수 연산과 같은 고수준 수학 함수



SciPy (사이파이)

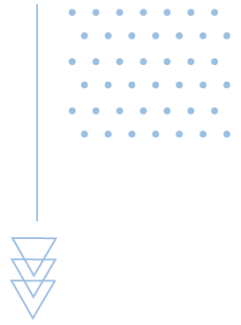
- 고성능 선형 대수, 함수 최적화, 신호 처리, 특수한 수학 함수와 통계 분포 등을 포함한 다양한 기능 제공

Pandas (판다스)

- 데이터 처리와 분석을 위한 패키지로 테이블 구조의 데이터 처리를 위한 다양한 기능 제공

Matplotlib (맷플롯립)

- 과학 계산용 그래프 패키지로 선 그래프, 히스토그램, 산점도 등을 지원



II. 데이터 탐색

■ 음식점의 주문 데이터를 분석하여 다음의 질문에 답해보자.

- 가장 인기 있는 메뉴 Top10 은?
- 메뉴별 주문량은?
- 메뉴별 가격은? 가장 비싼 메뉴는?

■ 데이터 탐색 과정

- 기술 통계 정보 살펴 보기 (총 주문 개수, 평균 주문량 등)
- 데이터 탐색과 시각화 (다양한 함수를 이용하여 주문 횟수, 주문량 등 계산, 막대 그래프로 시각화)
- 데이터 전처리 (문자 형식의 데이터를 숫자로 변환 등)