

Chapter 02 **PANDAS**

2 판다스가 뭐죠?

2020 Ver.

D.COM
DATA-SCIENCE

이번에 배울 것

1. Pandas는 왜 쓸까요?

2. Pandas 기초 사용법

1. Pandas는 왜 쓸까요?

Pandas는,

앞서 배운 Numpy와 유사하게 파이썬에서 데이터 저장 및 처리를 쉽고 빠르게 할 수 있는 파이썬 패키지입니다.

특히, 시계열(Series)이나 표(Table) 데이터를 구성하고 처리하는데 유용합니다.

마치 Excel을 파이썬으로 사용한다고 생각하시면 됩니다!

“Pandas를 쓰는 이유”

Numpy는 단일 산술 배열 데이터를 다룬다면
Pandas는 표 형식이나 다양한 형태의 데이터를 다루는데 특화되어 있습니다.

[https://medium.com/@5eo1ab/pandas-
%EC%93%B0%EB%8A%94-%EC%9D%B4%EC%9C%A0-
9063a90b0bd5](https://medium.com/@5eo1ab/pandas-%EC%93%B0%EB%8A%94-%EC%9D%B4%EC%9C%A0-9063a90b0bd5)

2. Pandas 기본 사용법

Numpy와 유사하므로
크게 어려울 것 없습니다!

학습자료

Pandas 패키지의 소개

<https://datascienceschool.net/view-notebook/ee0a5679dd574b94b55193690992f850/>

학습자료

Pandas 10분 완성

<https://numpy.org/doc/stable/reference/>

학습자료

Pandas Reference

<https://pandas.pydata.org/pandas-docs/stable/index.html>

CheckList

1. Pandas 패키지 불러오기
2. Series와 DataFrame이란?
3. 데이터 검색, 색인
4. NaN이란?
5. 재색인(reindex)
6. DataFrame row나 column 삭제
7. loc과 iloc
8. 산술 연산과 데이터 정렬
9. Pandas 유니버설 함수
10. Pandas 통계 메서드 (sum, max 등)
11. Pandas 데이터 로드와 저장