

01 구성과 특징

코들 수업자료는 각각의 단원이 하나의 노트북 파일(Ipython notebook)로 구성되어 있습니다.
각 장에는 핵심적인 이론 설명과 예제 그리고 연습문제들이 포함되어 있습니다.

1

핵심 이론과 자세한 소스 코드 해설이 담긴 실습예제로 효과적인 학습

[7] 데이터 인덱싱 - DataFrame편

[오늘 배울 것]

1. DataFrame의 열을 인덱싱할 수 있다.
2. DataFrame을 행을 인덱싱할 수 있다.
3. DataFrame의 부분을 2차원 인덱싱할 수 있다.

DataFrame은 데이터를 열과 행 두가지 차원으로 구조화하고 있습니다.
DataFrame을 인덱싱할 때에는 열과 행을 구분할 줄 알아야 합니다.
사용할 DataFrame을 하나 만들어봅시다.

```
1 # 예제 1: Pandas DataFrame 만들기
2
3 ShoppingList = pd.DataFrame(index=['apple', 'cup', 'pen', 'banana'],
4                               data={'price':[1000, 2000, 500, 4000],
5                                     'pcs':[5, 1, 3, 2],
6                                     'brand':['any', 'kakao', 'monami', 'delmont']})
7 ShoppingList
```

1. DataFrame 열 인덱싱

	Col0	Col1	Col2	Col3	Col4	Col5
Row0						
Row1						
Row2						
Row3						
Row4						

핵심 개념, 함수와 문법을 글과 도표를 통해 효과적으로 설명합니다.

실습 예제는 누구나 쉽게 따라하면서 학습할 수 있도록 코드를 짧은 단위로 나누어 실행해볼 수 있습니다.
매 코드 줄마다 해당 코드에 관한 자세한 설명을 주석으로 표시했습니다.

2

연습 문제를 통한 실력 점검

<연습문제>

판다스 기초 - 시리즈 만들기(1)

[시리즈 만들기]

문제

입력한 숫자들을 만들어진 리스트로 Pandas Series를 만들어봅시다.

입력

3개의 정수가 공백을 두고 입력됩니다.

출력

3개의 정수를 value로 하는 Pandas Series를 출력합니다.

주의! 주피터 내장 출력이 아닌 print 함수를 사용해서 출력하세요.

예시 입력

0 1 999

예시 출력

```
0    1
1    9
2   999
dtype: int64
```

```
1 import pandas as pd
2
3 a = map(int, input().split())
4 print(pd.Series(a))
```



자료 중간 중간에 배치된 연습문제를 통해 학생들의 참여도를 높이고, 배운 내용을 점검합니다.
연습문제는 작성한 코드를 자동으로 채점해주며, 선생님께서는 제출보드를 통해 학생들의 풀이 결과를 한 눈에 확인할 수 있습니다.

01 구성과 특징

코들 수업자료는 각각의 단원이 하나의 노트북 파일(Ipython notebook)로 구성되어 있습니다.
각 장에는 핵심적인 이론 설명과 예제 그리고 연습문제들이 포함되어 있습니다.

3 데이터 분석 문제와 추천 문제를 통한 실력 향상!

[7] 데이터 인덱싱 - DataFrame편

[오늘 배울 것]

1. DataFrame의 열을 인덱싱할 수 있다.
2. DataFrame을 행을 인덱싱할 수 있다.
3. DataFrame의 부분을 2차원 인덱싱할 수 있다.

DataFrame은 데이터를 열과 행 두가지 차원으로 구조화하고 있습니다.
DataFrame을 인덱싱할 때에는 열과 행을 구분할 줄 알아야 합니다.
사용할 DataFrame을 하나 만들어봅시다.

```
1 # 예제 1: Pandas DataFrame 만들기
2
3 ShoppingList = pd.DataFrame(index=['apple', 'cup', 'pen', 'banana'],
4                               data={'price':[1000, 2000, 500, 4000],
5                                     'pcs':[5, 1, 3, 2],
5                                     'brand':['any', 'kakao', 'monami', 'delmont']})
6
7 ShoppingList
```

1. DataFrame 열 인덱싱

	Col0	Col1	Col2	Col3	Col4	Col5
Row0						
Row1						
Row2						
Row3						
Row4						

인공지능과 데이터 분석과 연관되어있는 온라인 저지 문제도 탑재하였습니다.
문제를 일찍 푼 학생들을 위해 각 자료들은 맨 마지막에 추가 문제를 추천합니다.
여기서 이 단원과 연관된, 더 다양한 문제를 풀어볼 수 있습니다.

4 프로젝트를 통한 총정리

프로젝트 : 데이터 분석을 활용해 공약 설정하기

정치인이 된 영희는 지역구 선거를 나가려고 합니다. 영희의 핵심 공약은 "공공 와이파이 설치하기"입니다.

영희가 "공공 와이파이 설치하기" 공약으로 당선될 확률을 높이기 위해서는 공공 와이파이기가 적은 지역에 기기를 설치해야 합니다.

따라서 영희는 공공 와이파이기가 가장 적은 지역구에 출마하여 자신의 공약을 효과적으로 홍보하고 당선 확률을 올리고자 합니다.

영희가 자신이 출마할 지역구를 고를 수 있도록 도와줍니다.



수업자료의 중간, 마지막에는 프로젝트 수업으로 구성되어 있습니다.
학생들은 프로젝트를 통해 자신만의 결과물을 도출하고, 앞서 배운 내용들을 총정리합니다

02 목차

커리큘럼은 총 34차시로 구성되어 있습니다. 각 학교의 사정에 따라 데이터 분석 프로젝트, 데이터 종합 프로젝트의 분량을 조절하여 시수를 맞출 수 있습니다.

Pandas 라이브러리를 이용한 데이터 분석	데이터의 단위	1	데이터 분석 시작하기 : 시리즈와 데이터프레임
		2	시리즈와 데이터프레임을 만들어보아요
		3	데이터프레임을 읽어와 정보를 확인해보요
	데이터를 조작하고 사용하기	4	전부 사용하기에는 너무 많아! 시리즈 일부분 가져오기
		5	전부 사용하기에는 너무 많아! 데이터프레임 일부분 가져오기
		6	데이터로 계산하고 싶어요.
		7	크기가 다른 데이터는 어떻게 계산되나요?
		8	데이터에 함수들을 사용하고 싶어요.
		9	내가 원하는 데이터만 선택하고 싶어요.
	데이터를 가공하는 방법	10	잘못된 데이터를 수정하고 싶어요.
		11	새로운 행, 열을 추가하고 없애고 싶어요
		12	여러 데이터를 내 맘대로 합치고 나눠봐요
	데이터의 의미를 이해하는 방법	13	데이터의 결측값은 어떻게 처리하나요?
		14	데이터의 통계값을 확인하고 싶어요.
		15	그룹별을 나누어 계산해보아요.
데이터 분석 프로젝트	데이터 분석 프로젝트	16	누가누가 잘했나? 수상자를 찾아주세요! :그룹별 통계 프로젝트
		17	빅세일! 새로운 가격표가 필요해요. :데이터 연산 프로젝트
		18	타이타닉호에서 누가 살아남았나?! :데이터 가공 프로젝트
matplotlib 라이브러리를 이용한 데이터 시각화	데이터의 값 그리기	19	꺾은선 그래프를 그려보아요.
		20	여러 개의 꺾은선 그래프를 한번에 그려보아요.
		21	산점도를 그려보아요.
		22	여러 개의 산점도를 겹쳐서 그려보아요.
	데이터의 분포 그리기	23	막대그래프를 그려보아요.
		24	원그래프를 그려보아요.
		25	히스토그램을 그려보아요.
		26	2차원 히스토그램을 그려보아요.
	그래프 꾸미기	27	그래프의 배경을 꾸며보아요.
		28	그래프를 꾸며보아요.
		29	그래프를 여러 개 나열해보아요.
데이터 종합 프로젝트	데이터 종합 프로젝트	30	자유 프로젝트:데이터 선정하기
		31	자유 프로젝트:데이터 가공하기
		32	자유 프로젝트:데이터 분석하기
		33	자유 프로젝트:데이터 시각화하기
		34	자유 프로젝트:데이터 분석 결과 공유하기

03 단원설명

파이썬으로 데이터를 분석할 때 가장 널리 쓰이는 라이브러리 중 하나인 Pandas를 이용해 데이터를 분석하는 법을 배웁니다.
그 다음에는 Matplotlib를 사용해 데이터를 다양한 그래프로 시각화 하는 방법을 배웁니다.
마지막으로, 배운 내용을 바탕으로 데이터를 직접 선택하고 가공하는 프로젝트를 진행합니다.

a 데이터의 단위

Pandas의 series, dataframe에 대해 배우고 직접 series와 dataframe을 만들어봅니다.
series, dataframe의 형태를 확인하고 일부 데이터를 확인하는 방법을 배웁니다.

사용함수: `pandas.Series()`, `pandas.DataFrame()`, `describe()`, `info()`, `head()`, `tail()`

b 데이터를 조작하고 사용하기

series, dataframe을 인덱싱(특정 데이터 가져오기)하는 방법을 배웁니다.

Series, dataframe으로 산술연산, 비교연산, 논리연산에 대해 배웁니다.

브로드캐스팅, 불 인덱싱에 대해 배웁니다.

여러가지 연산을 위한 함수를 배웁니다.

사용함수: `iloc[]`, `loc[]`, `abs()`, `round()`, `rank()`, `isin()`

c 데이터를 가공하는 방법

조건에 맞는 series, dataframe의 데이터 값을 한번에 변경하는 방법에 대해서 배웁니다.

여러 개의 series와 dataframe을 합치는 방법에 대해서 배웁니다.

사용함수: `replace()`, `drop()`, `concat()`

d 데이터의 의미를 이해하는 방법

데이터를 분석하기 위해 결측치를 처리하는 방법에 대해 배웁니다.

데이터의 통계량을 구하고 이를 통해 데이터의 의미를 이해합니다.

데이터 그룹화를 통해 데이터를 분석하는 방법에 대해 배웁니다.

사용함수: `sum()`, `count()`, `idxmax()`, `idxmin()`, `fillna()`, `dropna()`, `isna()`
`max()`, `min()`, `mean()`, `median()`, `mode()`, `std()`, `groupby()`, `unique()`, `get_values()`

e 데이터 분석 프로젝트

중간고사 이전의 수행평가 및 복습에 사용할 수 있는 3개의 프로젝트형 수업이 준비되어 있습니다.

‘누가누가 잘했나? 수상자를 찾아주세요!’에서는 데이터를 여러 기준으로 그룹화하고

가장 점수가 높은 사람, 총 점수가 높은 반 등을 찾으며 그룹화를 연습합니다.

‘빅세일! 새로운 가격표가 필요해요.’에서는 상품별로 다른 가격 할인 방법을 적용해보며

연산 방법을 연습합니다. ‘타이타닉호에서 누가 살아남았나?!’에서는 데이터를 분석하기 이전에

결측치 처리, 값 수정이 필요하여 데이터 가공 방법을 연습할 수 있습니다.

선생님께서 상황에 맞추어 적절히 사용하실 수 있습니다.

03 단원설명

파이썬으로 데이터를 분석할 때 가장 널리 쓰이는 라이브러리 중 하나인 Pandas를 이용해 데이터를 분석하는 법을 배웁니다.
그 다음에는 Matplotlib를 사용해 데이터를 다양한 그래프로 시각화 하는 방법을 배웁니다.
마지막으로, 배운 내용을 바탕으로 데이터를 직접 선택하고 가공하는 프로젝트를 진행합니다.

f 데이터의 값 그리기

산점도와 꺾은선 그래프를 그리는 방법에 대해서 배웁니다.
한번에 여러개의 꺾은선 그래프, 산점도를 그리는 방법에 대해서 배웁니다.
사용함수: `scatter()`, `plot()`

g 데이터의 분포 그리기

데이터의 분포를 나타낼 수 있는 막대그래프, 히스토그램, 원그래프, 히트맵을 그리는 방법에 대해서 배웁니다.
사용함수: `bar()`, `hist()`, `pie()`, `hist2d()`, `pcolor()`

h 그래프 꾸미기

그래프의 제목, 축, 범례, 눈금, 그리드를 설정하는 방법을 배웁니다.
그래프의 색, 스타일을 바꾸는 방법에 대해서 배웁니다.
사용함수: `title()`, `xlabel()`, `ylabel()`, `legend()`, `xticks()`, `yticks()`, `grid()`

i 그래프 나열하기

여러개의 그래프를 한번에 그리는 방법에 대해서 배웁니다.
그래프에 표를 추가하는 방법에 대해 배웁니다.
사용함수: `subplots()`, `subplot()`, `set_title()`, `set_xlabel()`, `set_ylabel()`, `set_xticks()`, `set_yticks()`, `set_facecolor()`

j 데이터 종합 프로젝트

한 학기 동안 배운 내용을 모두 사용하여 학생들이 진행하는 프로젝트 입니다.
코드에서 제공하는 다양한데이터를 사용해 각자 프로젝트 결과물을 만들어 내는 프로젝트입니다.
단계별 가이드라인만 제공하기 때문에 학생들이 직접 데이터를 선정하고 분석할 수 있습니다.
데이터 선정, 가공, 분석, 시각화의 단계로 구성되어 있습니다.
선생님께서 시수에 맞추어 적절히 사용하실 수 있습니다.