[Python] Pandas 사용법 - 다양한 인덱스 함수(reset_index,set_index,sort_index) — 나무늘보의 개발 블로그

노트북: 첫 번째 노트북

만든 날짜: 2020-11-06 오후 5:24

URL: https://continuous-development.tistory.com/134?category=736681

Python

[Python] Pandas 사용법 - 다양한 인덱스 함 수(reset_index,set_index,sort_index)

2020. 10. 16. 00:45 수정 삭제 공개

reset_index()

기존 행 인덱스를 제거하고 인덱스를 데이터 열 추가

DataFrame.reset_index(inplace=True,drop=True)
=> 원본의 인덱스를 대체한다.
inplace는 생략 가능하다.

DataFrame.reset_index(inplace=False,drop=False)
=> 원본의 인덱스를 컬럼으로 만들고 인덱스를 생성한다.

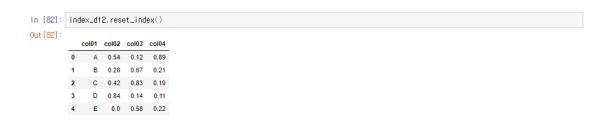
새로운 인덱스를 할당하고, 기존 인덱스는 인덱스라는 새로운 칼럼명으로 추가

n [141]:	ti	titanic.head()														
ut [141] :		survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
	3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	С	Southampton	yes	False
	4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True
	5	0	3	male	NaN	0	0	8.4583	Q	Third	man	True	NaN	Queenstown	no	True
	6	0	1	male	54.0	0	0	51.8625	S	First	man	True	Е	Southampton	no	True
	7	0	3	male	2.0	3	1	21.0750	S	Third	child	False	NaN	Southampton	no	False

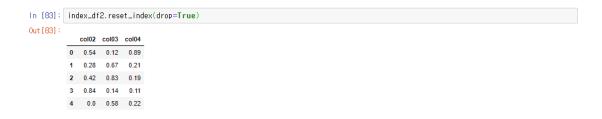


이 같은식으로 쓴다 inplace가 False 기 때문에 받는 변수를 지정해줘야한다.

만약 True로 한다면 원본데이터 변경되고 따로 지정은 안 해줘도 된다.



인덱스가 새로 생성 된 걸 볼 수 있다. 여기서 drop 옵션을 주면



기존의 인덱스였던 것을 drop 하고 새로 만들어진다.



리스트로 결과를 볼 수 있고

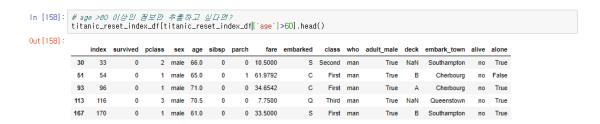
```
In [151]: | titanic_reset_index_df[titanic_reset_index_df['pclass'] == 3].head()
                                            fare embarked class who adult_male deck embark_town alive alone
          index survived pclass
                            sex age sibsp parch
        1 4 0 3 male 35.0 0 0 8.0500 S Third man True NaN Southampton no True
                        3 male NaN 0 0 8.4583
                                                      Q Third man
                                                                      True NaN Queenstown
                                                                    False NaN Southampton no False
        4 7 0 3 male 2.0 3 1 21.0750 S Third child
            8
                  1
                        3 female 27.0
                                     0
                                          2 11.1333
                                                      S Third woman
        7 10 1 3 female 4.0 1 1 16.7000 S Third child False G Southampton yes False
```

boolean indexing을 써서 해당하는 값을 가져올 수 도 있다.



다양한 방법으로 추출 할 수 있다.

ex)



```
In [159]: # age > 80 이상인 pc/ass, survived, who 만추출하고 싶다면?
titanic_reset_index_df[titanic_reset_index_df['age']>60][['pc/ass', 'survived', 'who']].head()

Out[159]:

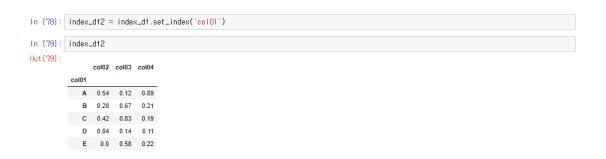
| pc/ass | survived | who | |
| 30 | 2 | 0 | man |
| 51 | 1 | 0 | man |
| 93 | 1 | 0 | man |
| 113 | 3 | 0 | man |
| 167 | 1 | 0 | man |
```

-set_index

```
DataFrame.set_index('컬럼')
```

기존 행 인덱스를 제거하고 데이터 열 중 하나를 인덱스로 설정

위와 같은 프레임이 있었을때 col01을 인덱스 값으로 변경 하기 위해서 는



데이터 프레임.set_index('컬럼명') 을 통해 변경할 수 있다.

```
In [81]: index_df3 = index_df2.set_index('co102')

Out[81]: 

| col03 | col04 |
| col02 |
| 0.54 | 0.12 | 0.89 |
| 0.28 | 0.67 | 0.21 |
| 0.42 | 0.83 | 0.19 |
| 0.84 | 0.14 | 0.11 |
| 0.0 | 0.58 | 0.22 |
```

이렇게 된다.

sort_index

index를 정렬한다.

정렬을 하기 전에 일단 랜덤으로 값을 뽑아서 가져온다.

```
In [185]: sort_df.columns = ['A','B','C','D']
sort_df.index = pd.date_range('20201014',periods=6) # pariods量 医胡 世來量 하나씩 늘起다.

Out [185]:

A B C D
2020-10-14 8 8 8 3 7
2020-10-16 5 7 0 4 2
2020-10-16 5 2 2 2
2020-10-17 1 0 8 4
2020-10-18 0 9 6 2
2020-10-19 4 1 5 3
```

여기까지 랜덤으로 데이터프레임을 만들었다.

sort_index를 통해 해당 인덱스를 sort 한다. 여기서는 axis =1은 col 즉 칼럼을 sort 하고 ascending=False로 하면 내림차순이 된다.

이렇게 axis=0으로 할 경우에는 행에 대한 sort 를 해준다.

단 하나의 컬럼을 sort 하기 위해서는

웨이와 같이 by라고 하고 해당 칼럼을 지정해준다.

만약 같은 값에 대한 우선순위를 정하고 싶다면

이런 식으로 값을 두 개를 넣는다.

'Python' 카테고리의 다른 글□

[Python] Pandas 사용법 - 그룹화 및 그룹 함수 (groupby, qcut, cut, transfrom ...

[Python] Pandas 사용법 - 두가지의 DataFrame 합치기 (merge, join) 🗆

[Python] Pandas 사용법 - 다양한 인덱스 함수(reset_index,set_index,sort_in...

[Python] Pandas 사용법 - 인덱싱 접근,데이터 조작, 인덱스조작(loc,iloc)□

[Python] Pandas 사용법 - 다양한 함수 사용(데이터 입출력, 대소문자변환, 공백...

[Python] Pandas 사용법 - DataFrame 생성, 추가 , 수정, 삭제, indexing 🗆

reset_index

set_index

sort_index



나무늘보스

혼자 끄적끄적하는 블로그 입니다.