

▼ [Quiz] Series sr로 부터 인덱스 이름을 사용하여 print(sr[[1, 2]]) 와 동일한 결과를 출력하세요.

```
1 tup_data = ('영인', '2010-05-01', '여', True)
2 sr = pd.Series(tup_data, index=['이름', '생년월일', '성별', '학생여부'])
3 print(sr[[1, 2]])
4 print(sr[['생년월일', '성별']])
```



```
생년월일    2010-05-01
성별        여
dtype: object
생년월일    2010-05-01
성별        여
dtype: object
```

```
1 import pandas as pd
2
3 # 열이름을 key로 하고, 리스트를 value로 갖는 딕셔너리 정의.(2차원 배열)
4 dict_data = {'c0':[1,2,3], 'c1':[4,5,6], 'c2':[7,8,9], 'c3':[10,11,12], 'c4':[13
5
6 # 판다스 DataFrame() 함수로 딕셔너리를 데이터프레임으로 변환. 변수 df에 저장.
7 df = pd.DataFrame(dict_data)
8
9 # df의 자료형 출력
10 print(type(df))
11 print()
12 # 변수 df에 저장되어 있는 데이터프레임 객체를 출력
13 print(df)
14 print()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
   c0  c1  c2  c3  c4
0    1   4   7  10  13
1    2   5   8  11  14
2    3   6   9  12  15
```

```
1 data = [[15, '남', '덕영중'], [17, '여', '수리중']],
2 index=['준서', '예은'],
3 columns=['나이', '성별', '학교']
```

▼ [Quiz] DataFrame을 만드세요

[Quiz] df에서 행 인덱스, 열 이름을 변경하여 아래와 같이 출력 하세요.

```
연령  남녀    소속\
```

```
학생1 15 남 덕영중
```

```
학생2 17 여 수리중
```

[Quiz] 인덱스를 이용해서 덕영중을 찾으세요(2가지 방식)

```

1 data = [[15, '남', '덕영중'], [17, '여', '수리중']]
2 index = ['준서', '예은']
3 columns = ['나이', '성별', '학교']
4 df = pd.DataFrame(data, index=index, columns=columns, dtype=str)
5
6 df.rename(columns = {'나이' : '연령', '성별' : '남녀', '학교' : '소속'}, inplace=True)
7
8 print (df)
9
10 # print ( df['소속'].isin(['덕영중']) )
11 print ()
12
13 print ( df[df['소속'].isin(['덕영중'])] )
14 print ()
15
16 print (df[(df['소속']=='덕영중')])

```

```

☐➤      연령  남녀   소속
준서   15   남   덕영중
예은   17   여   수리중

```

```

      연령  남녀   소속
준서   15   남   덕영중

```

```

      연령  남녀   소속
준서   15   남   덕영중

```

▼ Q. 열 이름 중, '나이'를 '연령'으로, '성별'을 '남녀'로, '학교'를 '소속'으로 바꾸기

```

1 df.rename(columns = {'나이' : '연령', '성별' : '남녀', '학교' : '소속'}, inplace=True)
2
3 print (df)

```

```

☐➤      연령  남녀   소속
준서   15   남   덕영중
예은   17   여   수리중

```

▼ [Quiz] df의 행 인덱스 중에서, '준서'를 '학생1'로, '예은'을 '학생2'로 바꾸세요.

```

1 df.rename( index = { '준서' : '학생1', '예은' : '학생2'}, inplace = True)
2 print (df)

```

```

☐➤      연령  남녀   소속
학생1   15   남   덕영중
학생2   17   여   수리중

```

```

1 # 행 삭제
2
3 import pandas as pd
4
5 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
6 exam_data = {'스하' : [90, 80, 70], '연어' : [98, 89, 95]}

```

```

6 exam_data = { '수학' : [ 90, 80, 70 ], '영어' : [ 98, 89, 95 ],
7               '음악' : [ 85, 95, 100 ], '체육' : [ 100, 90, 90 ] }
8
9 df = pd.DataFrame(exam_data, index=[ '서준', '우현', '인아' ])
10 print(df)
11 print()
12 # 데이터프레임 df를 복제하여 변수 df2에 저장.
13 df2 = df[:]
14 print(df2)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

1 import warnings
2 warnings.filterwarnings('ignore')
3
4 # df2의 1개 행(row)을 삭제
5 df2.drop('우현', axis=0, inplace=True)
6 print(df2)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
인아   70   95  100    90

```

- ▼ [Quiz] 데이터프레임 df를 복제하여 변수 df3에 저장 df3의 2개 행(row) '우현', '인아'를 삭제하세요 (inplace=True를 사용하는 방법 및 사용하지 않는 방법 2개 모두)

```

1 df3 = df[:]
2 df3.drop('우현', axis=0, inplace=True)
3 df3.drop('인아', axis=0, inplace=True)
4 print (df3)
5 print ()
6 df3 = df[:]
7 df3.drop('우현', axis=0, inplace=False)
8 df3.drop('인아', axis=0, inplace=False)
9 print (df3)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100

```

```

   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90
   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

1 # 열 삭제
2
3 import pandas as pd
4
5 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
6 exam_data = {'수학' : [ 90, 80, 70], '영어' : [ 98, 89, 95],
7              '음악' : [ 85, 95, 100], '체육' : [ 100, 90, 90]}
8
9 df = pd.DataFrame(exam_data, index=['서준', '우현', '인아'])
10 print(df)
11 print()
12
13 # 데이터프레임 df를 복제하여 변수 df4에 저장. df4의 1개 열(column)을 삭제
14 df4 = df[: ]
15 print(df4)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

1 df4.drop('수학', axis=1, inplace=True)
2 print(df4)

```

```

↳
   영어  음악  체육
서준   98   85  100
우현   89   95   90
인아   95  100   90

```

▼ [Quiz] 데이터프레임 df를 복제하여 변수 df5에 저장. df5의 2개 열(column) '영어', '음악'을 삭제

```

1 df5 = df[: ]
2
3 print (df5)
4
5 df5.drop('영어', axis=1, inplace=True)
6 df5.drop('음악', axis=1, inplace=True)
7
8 print (df5)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90
   수학  체육
서준   90  100
우현   80   90
인아   70   90

```

```

1 # 행 선택

```

```

2
3 import pandas as pd
4
5 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
6 exam_data = {'수학' : [ 90, 80, 70], '영어' : [ 98, 89, 95],
7              '음악' : [ 85, 95, 100], '체육' : [ 100, 90, 90]}
8
9 df = pd.DataFrame(exam_data, index=['서준', '우현', '인아'])
10 print(df)          # 데이터프레임 출력
11 print('\n')

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

```

1 # 인덱스 이름을 기준으로 행을 선택할 때는 loc, 정수형 위치 인덱스를 사용할 때는 iloc를 이용.
2 # 행 인덱스를 사용하여 행 1개를 선택
3 label1 = df.loc['서준']      # loc 인덱스 활용
4 position1 = df.iloc[0]      # iloc 인덱스 활용
5 print(label1)
6 print()
7 print(position1)
8 print()

```

```

↳
   수학    90
   영어    98
   음악    85
   체육   100
   Name: 서준, dtype: int64

   수학    90
   영어    98
   음악    85
   체육   100
   Name: 서준, dtype: int64

```

```

1 # 행 인덱스를 사용하여 2개 이상의 행 선택
2 label2 = df.loc[['서준', '인아']] # 인덱스의 이름을 배열로 전달
3 position2 = df.iloc[[0, 2]] # 정수형 위치를 전달
4 print(label2)
5 print()
6 print(position2)

```

```

↳
   수학  영어  음악  체육
서준   90   98   85   100
인아   70   95  100    90

   수학  영어  음악  체육
서준   90   98   85   100
인아   70   95  100    90

```

▼ [Quiz] '서준', '우현' 학생을 행 인덱스의 범위(:)를 지정하여 선택하세요.

(인덱스 이름과 위치 인덱스 2가지 모두)

```
1 print (df.loc[['서준', '우현']])
2 print ()
3 print (df.iloc[0:2])
```

```
↳
```

	수학	영어	음악	체육
서준	90	98	85	100
우현	80	89	95	90

```
↳
```

	수학	영어	음악	체육
서준	90	98	85	100
우현	80	89	95	90

```
1 # 열 선택
2
3 exam_data = {'이름' : [ '서준', '우현', '인아' ],
4              '수학' : [ 90, 80, 70 ],
5              '영어' : [ 98, 89, 95 ],
6              '음악' : [ 85, 95, 100 ],
7              '체육' : [ 100, 90, 90 ]}
8
9 df = pd.DataFrame(exam_data)
10 print(df)
11 print(type(df))
```

```
↳
```

	이름	수학	영어	음악	체육
0	서준	90	98	85	100
1	우현	80	89	95	90
2	인아	70	95	100	90

```
<class 'pandas.core.frame.DataFrame'>
```

```
1 # 대괄호안에 열 이름을 입력하면 시리즈 객체로 추출
2 math1 = df['수학']
3 print(math1)
4 print(type(math1))
5 print()
6 # 도트 다음에 열 이름을 입력하면 시리즈 객체로 추출
7 english = df.영어
8 print(english)
9 print(type(english))
10 print()
11 # 대괄호안에 열 이름의 리스트로 입력하면 데이터프레임이 반환됨
12 music_gym = df[['음악', '체육']]
13 print(music_gym)
14 print(type(music_gym))
15 print()
16 # 2중 대괄호를 사용하면 열 이름 1개를 원소로 갖는 리스트를 사용하는 경우에도 데이터프레임을 반환
17 math2 = df[['수학']]
18 print(math2)
19 print(type(math2))
```

```
↳
```

```

0    90
1    80
2    70
Name: 수학, dtype: int64
<class 'pandas.core.series.Series'>

0    98
1    89
2    95
Name: 영어, dtype: int64
<class 'pandas.core.series.Series'>

      음악   체육
0    85   100
1    95    90
2   100    90
<class 'pandas.core.frame.DataFrame'>

      수학
0    90
1    80
2    70
<class 'pandas.core.frame.DataFrame'>

```

- ▼ [Quiz] df에서 '체육' 칼럼을 제외한 후 df_e 이름으로 새로운 데이터프레임을 생성, 출력하세요.

```

1 df_e = df[:]
2 df_e.drop('체육', axis=1, inplace=True)
3 print (df_e)

```

```

[>]      이름  수학  영어   음악
0  서준   90   98   85
1  우현   80   89   95
2  인아   70   95  100

```

```

1 # 전체행을 2행 간격으로 선택
2 print(df.iloc[::2])
3 print()
4 # 0행에서 2행까지 간격을 2로 선택
5 print(df.iloc[0:3:2])
6 print()
7 # 역순으로 인덱싱
8 print(df.iloc[::-1])

```

```

[>]      이름  수학  영어   음악  체육
0  서준   90   98   85  100
2  인아   70   95  100   90

```

```

      이름  수학  영어   음악  체육
0  서준   90   98   85  100
2  인아   70   95  100   90

```

```

      이름  수학  영어   음악  체육
2  인아   70   95  100   90
1  우현   80   89   95   90
0  서준   90   98   85  100

```

```

1 # 원소 선택
2 import pandas as pd
3
4 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
5 exam_data = {'이름' : [ '서준', '우현', '인아' ],
6              '수학' : [ 90, 80, 70 ],
7              '영어' : [ 98, 89, 95 ],
8              '음악' : [ 85, 95, 100 ],
9              '체육' : [ 100, 90, 90 ]}
10 df = pd.DataFrame(exam_data)
11
12 # '이름' 열을 새로운 인덱스로 지정하고, df 객체에 변경사항 반영
13 df.set_index('이름', inplace=True)
14 print(df)

```

☞

	수학	영어	음악	체육
이름				
서준	90	98	85	100
우현	80	89	95	90
인아	70	95	100	90

▼ iloc, loc를 이용해서 아래 요구사항들을 출력하세요.

- 데이터프레임 df의 특정 원소 1개 선택 ('서준'의 '음악' 점수 - 2가지 방법)
- 데이터프레임 df의 특정 원소 2개 이상 선택 ('서준'의 '음악', '체육' 점수 - 4가지 방법)
- df의 2개 이상의 행과 열로부터 원소 선택 ('서준', '우현'의 '음악', '체육' 점수 - 4가지 방법)
- 정수형 위치 인덱스로 열선택(0번째 열 선택 - Series와 Dataframe 형태)

```

1 print ( df )
2 print ()
3 print ( '서준의 음악 점수' )
4 print (df.loc[ '서준' ].loc[ '음악' ])
5 print (df.iloc[0].iloc[2])
6 print ()
7 print ( '서준의 음악, 체육 점수' )
8 print (df.loc[ '서준' ].iloc[2:4])
9 print (df.loc[ '서준' ].loc[ [ '음악', '체육' ] ])
10 print (df.iloc[0][ [ '음악', '체육' ] ])
11 print (df.loc[ '서준' ][ [2,3] ])
12 print ()
13 print ( '서준, 우현의 음악, 체육 점수' )
14 print (df.iloc[0:2, 2:4])
15 print (df.iloc[ [0,1], [2,3] ])
16 print (df.loc[ '서준': '우현', '음악': '체육' ])
17 print (df.loc[ [ '서준', '우현' ], [ '음악', '체육' ] ])
18 print (df.iloc[ [0,1], df.columns.get_indexer([ '음악', '체육' ] ) ])

```

☞

	수학	영어	음악	체육
이름				
서준	90	98	85	100
우현	80	89	95	90
인아	70	95	100	90

서준의 음악 점수
85
85

서준의 음악, 체육 점수
음악 85
체육 100
Name: 서준, dtype: int64
음악 85
체육 100
Name: 서준, dtype: int64
음악 85
체육 100
Name: 서준, dtype: int64
음악 85
체육 100
Name: 서준, dtype: int64
음악 85
체육 100
Name: 서준, dtype: int64

서준, 우현의 음악, 체육 점수
음악 체육

이름	음악	체육
서준	85	100
우현	95	90

음악 체육

이름	음악	체육
서준	85	100
우현	95	90

음악 체육

이름	음악	체육
서준	85	100
우현	95	90

음악 체육

이름	음악	체육
서준	85	100
우현	95	90

음악 체육

이름	음악	체육
서준	85	100
우현	95	90

```

1 # 열 추가
2
3 exam_data = {'이름' : [ '서준', '우현', '인아'],
4               '수학' : [ 90, 80, 70],
5               '영어' : [ 98, 89, 95],
6               '음악' : [ 85, 95, 100],
7               '체육' : [ 100, 90, 90]}
8 df = pd.DataFrame(exam_data)
9 print(df)
10 print()
11 df['국어'] = 80
12 # 모든 행에 동일한 값이 입력
13 print(df)

```

```
10 print(df)
```

```

이름  수학  영어  음악  체육
0  서준   90   98   85   100
1  우현   80   89   95    90
2  인아   70   95  100    90

```

```

이름  수학  영어  음악  체육  국어
0  서준   90   98   85  100   80
1  우현   80   89   95   90   80
2  인아   70   95  100   90   80

```

```

1 # 행 추가
2 import pandas as pd
3
4 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
5 exam_data = {'이름' : ['서준', '우현', '인아'],
6              '수학' : [ 90, 80, 70],
7              '영어' : [ 98, 89, 95],
8              '음악' : [ 85, 95, 100],
9              '체육' : [ 100, 90, 90]}
10 df = pd.DataFrame(exam_data)
11 print(df)
12 print()
13
14 # 새로운 행(row)을 추가 - 같은 원소 값을 입력
15 df.loc[3] = 0
16 print(df)
17 print()
18
19 # 새로운 행(row)을 추가 - 원소 값 여러 개의 배열 입력
20 df.loc[4] = ['동규', 90, 80, 70, 60]
21 print(df)
22 print()
23
24 # 새로운 행(row)을 추가 - 기존 행을 복사
25 df.loc['행5'] = df.loc[3]
26 print(df)

```

```


```

	이름	수학	영어	음악	체육
0	서준	90	98	85	100
1	우현	80	89	95	90
2	인아	70	95	100	90

	이름	수학	영어	음악	체육
0	서준	90	98	85	100
1	우현	80	89	95	90
2	인아	70	95	100	90
3	0	0	0	0	0

	이름	수학	영어	음악	체육
0	서준	90	98	85	100
1	우현	80	89	95	90
2	인아	70	95	100	90
3	0	0	0	0	0
4	동규	90	80	70	60

	이름	수학	영어	음악	체육
0	서준	90	98	85	100
1	우현	80	89	95	90
2	인아	70	95	100	90
3	0	0	0	0	0
4	동규	90	80	70	60
행5	0	0	0	0	0

```

1 # 원소 값 변경
2 import pandas as pd
3
4 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
5 exam_data = {'이름' : [ '서준', '우현', '인아' ],
6              '수학' : [ 90, 80, 70 ],
7              '영어' : [ 98, 89, 95 ],
8              '음악' : [ 85, 95, 100 ],
9              '체육' : [ 100, 90, 90 ]}
10 df = pd.DataFrame(exam_data)
11 print(df)
12 print()
13 df_idx = df.set_index('이름')
14 # df_idx = df.set_index('이름', inplace=True) # df_idx가 None이 됨
15 print(df_idx)

```

```

↳
    이름  수학  영어  음악  체육
0  서준   90   98   85   100
1  우현   80   89   95    90
2  인아   70   95  100    90

    수학  영어  음악  체육
이름
서준   90   98   85   100
우현   80   89   95    90
인아   70   95  100    90

```

▼ [Quiz] loc, iloc를 이용해서 아래 요구사항들을 출력하세요.

- 데이터프레임 df_idx에서 '서준'의 '체육' 점수를 3가지 방법으로 80, 90, 100점으로 변경(인덱스 이름 사

- 데이터프레임 df_idx에서 '서준'의 '음악', '체육' 점수를 100점, 50점으로 각각 변경(인덱스 이름 사용)
- 데이터프레임 df_idx에서 '서준'의 '체육' 점수를 90점으로 변경(정수 위치 인덱스 사용)
- 데이터프레임 df_idx에서 '서준'의 '음악', '체육' 점수를 80점, 70점으로 각각 변경(정수 위치 인덱스 사용)

```

1 print (df_idx)
2 print ()
3 df_idx.loc['서준'].loc['체육'] = 80
4 print (df_idx)
5 print ()
6 df_idx['체육']['서준'] = 90
7 print (df_idx)
8 print ()
9 df_idx.loc['서준']['체육'] =100
10 print (df_idx)

```

```

↳
      수학  영어   음악   체육
이름
서준    90   98    85   100
우현    80   89    95    90
인아    70   95   100    90

```

```

      수학  영어   음악   체육
이름
서준    90   98    85    80
우현    80   89    95    90
인아    70   95   100    90

```

```

      수학  영어   음악   체육
이름
서준    90   98    85    90
우현    80   89    95    90
인아    70   95   100    90

```

```

      수학  영어   음악   체육
이름
서준    90   98    85   100
우현    80   89    95    90
인아    70   95   100    90

```

```

1 # 행, 열의 위치 바꾸기
2
3 import pandas as pd
4
5 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
6 exam_data = {'이름' : [ '서준', '우현', '인아'],
7              '수학' : [ 90, 80, 70],
8              '영어' : [ 98, 89, 95],
9              '음악' : [ 85, 95, 100],
10             '체육' : [ 100, 90, 90]}
11 df = pd.DataFrame(exam_data)
12 print(df)
13 print()
14 # 데이터프레임 df를 전치하기 (메소드 활용)
15 df_t = df.transpose()
16 print(df_t)

```

```

↳
   이름  수학  영어  음악  체육
0  서준   90   98   85  100
1  우현   80   89   95   90
2  인아   70   95  100   90

   이름  수학  영어  음악  체육
0  서준   90   98   85  100
1  우현   80   89   95   90
2  인아   70   95  100   90

```

▼ [Quiz] df_t를 다시 전치해서 원래의 df를 출력하세요.

```

1 df_t = df.transpose().transpose()
2 print(df_t)

```

```

↳
   이름  수학  영어  음악  체육
0  서준   90   98   85  100
1  우현   80   89   95   90
2  인아   70   95  100   90

```

▼ 인덱스 활용

```

1 # 특정 열을 행 인덱스로 설정
2 import pandas as pd
3
4 # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장
5 exam_data = {'이름' : [ '서준', '우현', '인아'],
6              '수학' : [ 90, 80, 70],
7              '영어' : [ 98, 89, 95],
8              '음악' : [ 85, 95, 100],
9              '체육' : [ 100, 90, 90]}
10 df = pd.DataFrame(exam_data)
11 print(df)

```

```

↳
   이름  수학  영어  음악  체육
0  서준   90   98   85  100
1  우현   80   89   95   90
2  인아   70   95  100   90

```

▼ [Quiz] set_index() 함수를 이용, 아래 요구사항들을 출력하세요.

(기존 df는 변경하지 말고 별도의 이름으로 변경 내용 반영한 새로운 데이터프레임 생성)

- '이름' 칼럼을 인덱스로 설정
- '음악' 칼럼을 인덱스로 설정
- '수학','음악' 칼럼을 인덱스로 설정

```

1 # df = df.set_index('이름')
2 # print (df)

```

```
3 # print ()
4 # df = df.set_index('음악')
5 # print (df)
6 # print ()
7 df = df.set_index(['수학', '음악'])
8 print (df)
9 print ()
```

↗

		이름	영어	체육
	수학	음악		
90	85	서준	98	100
80	95	우현	89	90
70	100	인아	95	90