

In [7]:

```
import pandas as pd
df = pd.read_csv('./data/scores.csv')
df.index = 'i' + df.index.astype('str')
df.head()
```

Out[7]:

	name	kor	eng	math
i0	Aiden	100.0	90.0	95.0
i1	Charles	90.0	80.0	75.0
i2	Danial	95.0	100.0	100.0
i3	Evan	100.0	100.0	100.0
i4	Henry	NaN	35.0	60.0

인덱스



- 인덱스 : 행을 구분하기 위한 구분자, 각 행에 붙여진 이름
- 행번호 : 각 행의 위치
- 따라서 인덱스 != 행번호



숫자로 이루어질 수도 있고, 정수로 이루어질 수도 있다.

인덱스로 행 추출하기

```
df.loc['인덱스명']
df.loc[] = 인덱스 리스트]
```



1.1 시리즈 형태로 추출하기

In [10]:

```
# 인덱스가 i3인 행 추출하여 s1에 담기
s1 = df.loc['i3']
```

In [11]:

```
# s1 type
type(s1)
```

Out[11]:

pandas.core.series.Series

In [13]:

```
# s1 index
s1.index
```

```
Index(['name', 'kor', 'eng', 'math'], dtype='object')
```

```
# s1 value
s1.values
```

```
array(['Evan', 100.0, 100.0, 100.0], dtype=object)
```

In [18]:

```
# 인덱스가 i1, i3, i5인 행 추출하기
df.loc[['i1', 'i3', 'i5']]
```

	name	kor	eng	math
i1	Charles	90.0	80.0	75.0
i3	Evan	100.0	100.0	100.0
i5	Ian	90.0	100.0	90.0

```
# 인덱스가 i3인 행을 데이터프레임 형태로 추출하기
df.loc[['i3']]
```

	name	kor	eng	math
i3	Evan	100.0	100.0	100.0

```
# 인덱스에 없는 값을 사용하면 error
df.loc['i-1']
```

The above exception was the direct cause of the following exception:

[illegible]

```
Cell In [22], line 2
      1 # 인덱스에 없는 값을 사용하면 error
----> 2 df.loc['i-1']
```

```
File c:\Users\com\anaconda3\envs\vscode\lib\site-packages\pandas\core\indexing.py:967, in
_LocationIndexer.__getitem__(self, key)
    964 axis = self.axis or 0
    966 maybe_callable = com.apply_if_callable(key, self.obj)
--> 967 return self._getitem_axis(maybe_callable, axis=axis)
```

```
File c:\Users\com\anaconda3\envs\vscode\lib\site-packages\pandas\core\indexing.py:1205, i
n _LocIndexer._getitem_axis(self, key, axis)
    1203 # fall thru to straight lookup
    1204 self._validate_key(key, axis)
-> 1205 return self._get_label(key, axis=axis)
```

```
File c:\Users\com\anaconda3\envs\vscode\lib\site-packages\pandas\core\indexing.py:1153, i
n _LocIndexer._get_label(self, label, axis)
    1151 def _get_label(self, label, axis: int):
    1152     # GH#5667 this will fail if the label is not present in the axis.
-> 1153     return self.obj.xs(label, axis=axis)
```

```
File c:\Users\com\anaconda3\envs\vscode\lib\site-packages\pandas\core\generic.py:3864, in
NDFrame.xs(self, key, axis, level, drop_level)
    3862         new_index = index[loc]
    3863     else:
-> 3864         loc = index.get_loc(key)
    3866         if isinstance(loc, np.ndarray):
    3867             if loc.dtype == np.bool_:
```

```
File c:\Users\com\anaconda3\envs\vscode\lib\site-packages\pandas\core\indexes\base.py:363
1, in Index.get_loc(self, key, method, tolerance)
    3629     return self._engine.get_loc(casted_key)
    3630 except KeyError as err:
-> 3631     raise KeyError(key) from err
    3632 except TypeError:
    3633     # If we have a listlike key, _check_indexing_error will raise
    3634     # InvalidIndexError. Otherwise we fall through and re-raise
    3635     # the TypeError.
    3636     self._check_indexing_error(key)
```

KeyError: 'i-1'



2. 인덱스로 행, 열 추출하기

2.1 인덱스, 컬럼에 해당하는 한 개의 데이터 뽑아오기

```
loc[인덱스, 컬럼명]
```

```
In [23]:
```

```
# 인덱스 i1의 kor 점수
df.loc['i1', 'kor']
```

```
Out[23]:
```

```
90.0
```

```
In [24]:
```

```
#type
type(df.loc['i1', 'kor'])
```

```
Out[24]:
```

```
numpy.float64
```

2.2 특정 인덱스의 여러 컬럼 데이터 추출하기

```
df.loc['인덱스', 컬럼명 리스트]
```

In [25]:

```
# 인덱스 i1의 [name, kor] 컬럼  
df.loc['i1', ['name', 'kor']]
```

Out[25]:

```
name    Charles  
kor      90.0  
Name: i1, dtype: object
```

In [26]:

```
# type  
type(df.loc['i1', ['name', 'kor']])
```

Out[26]:

```
pandas.core.series.Series
```

2.3 여러 인덱스의 특정 컬럼 데이터 추출하기

```
df.loc[인덱스 리스트, 컬럼 명]
```

In [27]:

```
# 인덱스 i1, i3, i5의 name 컬럼  
df.loc[['i1', 'i3', 'i5'], 'name']
```

Out[27]:

```
i1    Charles  
i3      Evan  
i5      Ian  
Name: name, dtype: object
```

In [29]:

```
# type  
type(df.loc[['i1', 'i3', 'i5'], 'name']) # Series
```

Out[29]:

```
pandas.core.series.Series
```

2.4 여러 인덱스의 여러 컬럼 가져오기

```
df.loc[인덱스 리스트, 컬럼명 리스트]
```

In [30]:

```
# 인덱스 i1, i3, i5의 name, kor  
df.loc[['i1', 'i3', 'i5'], ['name', 'kor']]
```

Out[30]:

	name	kor
i1	Charles	90.0
i3	Evan	100.0
i5	Ian	90.0

In [31]:

```
# type
type(df.loc[['i1', 'i3', 'i5'], ['name', 'kor']]) # dataframe
```

Out[31]:

pandas.core.frame.DataFrame

2.5 모든 행에서 특정 열 가져오기

```
df.loc[:, 컬럼명]
df.loc[:, 컬럼 리스트]
```



In [33]:

```
# 모든 행에서 'name' 가져오기
df.loc[:, 'name'].head()
```

Out[33]:

```
i0      Aiden
i1    Charles
i2    Danial
i3      Evan
i4    Henry
Name: name, dtype: object
```

In [34]:

```
# 모든 행에서 ['name', 'math'] 가져오기
df.loc[:, ['name', 'math']].head()
```

Out[34]:

	name	math
i0	Aiden	95.0
i1	Charles	75.0
i2	Danial	100.0
i3	Evan	100.0
i4	Henry	60.0

In [36]:

```
# 모든 행에서 'name' DataFrame으로 가져오기
df.loc[:, ['name']].head()
```

Out[36]:

	name
i0	Aiden
i1	Charles
i2	Danial
i3	Evan
i4	Henry

In [37]:

```
# type
type(df.loc[:, ['name']]) # DataFrame
```

Out[37]:

```
Out[57]:
```

```
pandas.core.frame.DataFrame
```