Data Science Essential learnings 1

August 13, 2022

[10]:
```
import pandas as pd
import numpy as np

from pandas import Series, DataFrame
```

0.0.1 Selecting and retreving data

you can write an index value in two form

+label index or

+interger index

[16]:
```
series_obj = Series(np.arange(6), index=['row 1','row 2','row 3','row 4','row␣
↪5','row 6'])
```

[17]:
```
series_obj
```

[17]:
```
row 1    0
row 2    1
row 3    2
row 4    3
row 5    4
row 6    5
dtype: int64
```

[18]:
```
series_obj['row 6']
```

[18]:
```
5
```

[20]:
```
series_obj['row 5']
```

[20]:
```
4
```

[28]:
```
series_obj[[0, 5]]
```

[28]:
```
row 1    0
row 6    5
```

```
dtype: int64
```

[29]:
```
series_obj[[0,5]]
```

[29]:
```
row 1    0
row 6    5
dtype: int64
```

[30]:
```
series_obj[[4]]
```

[30]:
```
row 5    4
dtype: int64
```

0.0.2 chreating data frame obj

[51]:
```
np.random.seed(25)
```

[65]:
```
DF_obj =  DataFrame(np.random.rand(36).reshape((6,6)),
index=['row 1','row 2','row 3','row 4','row 5','row 6'],
columns=['colums 1','column 2','columns 3','column 4','column 5', 'column 6'])
DF_obj
```

[65]:

	colums 1	column 2	columns 3	column 4	column 5	column 6
row 1	0.590885	0.163652	0.836928	0.775203	0.169041	0.766994
row 2	0.335366	0.472398	0.215064	0.912094	0.759208	0.676561
row 3	0.021376	0.660874	0.094440	0.831163	0.112749	0.566830
row 4	0.174626	0.790607	0.033683	0.795971	0.689437	0.491846
row 5	0.088554	0.937550	0.084362	0.469394	0.805613	0.085646
row 6	0.244380	0.892806	0.478611	0.190401	0.253044	0.762339

[68]:
```
DF_obj.loc[['row 2','row 5'],['column 5','column 2']] # to get valuve of row 2␣
↪row 5
#colu, 5 and 2
```

[68]:

	column 5	column 2
row 2	0.759208	0.472398
row 5	0.805613	0.937550

0.0.3 Data Slicing

[69]:
```
series_obj['row 2':'row 5']
```

[69]:
```
row 2    1
row 3    2
row 4    3
row 5    4
dtype: int64
```

0.0.4 comparing with scaler

[73]:
```
series_obj[series_obj > 4]
```

[73]:
```
row 6    5
dtype: int64
```

0.0.5 setting values with scalars

[79]:
```
series_obj['row 1','row 5','row 6'] = 8
series_obj
```

[79]:
```
row 1    8
row 2    1
row 3    2
row 4    3
row 5    8
row 6    8
dtype: int64
```