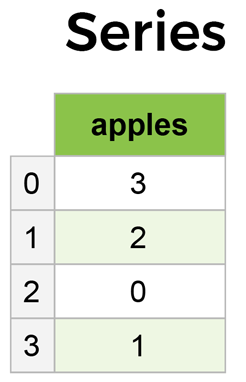
**Installing Pandas**

**if you have installed anaconda, pandas is already part of it. for other python installation use "pip install pandas" to install latest version**

**Section 5.1 Pandas Series**



In [2]:

*# must need to import pandas*

**import** **pandas** **as** **pd**

obj = pd.Series([4, 7, -5, 3])

print(obj)

0 4

1 7

2 -5

3 3

dtype: int64

**pandas Series has two things in one object**

**1 Values**

**2 Index**

In [3]:

sales = pd.Series([100, 200, 100, 400])

print(sales.values)

print(sales.index)

[100 200 100 400]

RangeIndex(start=0, stop=4, step=1)

In [5]:

sales = pd.Series([100, 200, 100, 400], index = ['Jan', 'Feb', 'Mar','Apr'])

print(sales)

print(sales.values)

print(sales.index)

Jan 100

Feb 200

Mar 100

Apr 400

dtype: int64

[100 200 100 400]

Index(['Jan', 'Feb', 'Mar', 'Apr'], dtype='object')

In [6]:

sales = pd.Series([100, 200, 100, 400], index = ['Jan', 'Feb', 'Mar','Apr'], name="4 month sales")

print(sales)

Jan 100

Feb 200

Mar 100

Apr 400

Name: 4 month sales, dtype: int64

**create a pandas series to store a canteen data to hold values of how many sandwithes are sold each days (one week)**

In [7]:

sw = pd.Series([20, 30, 20, 25, 30, 40, 0],

index = ['mon', 'tue', 'wed', 'thr','fri', 'sat', 'sun'])

print(sw)

mon 20

tue 30

wed 20

thr 25

fri 30

sat 40

sun 0

dtype: int64

In [8]:

print(sw[1])

print(sw["tue"])

print(sw["sun"])

30

30

0

In [9]:

print(sw[ [3,5] ])

print(sw[ [ "tue", "thr"] ])

thr 25

sat 40

dtype: int64

tue 30

thr 25

dtype: int64

**for providing multipe index for selecting element in an pd series, use array notation**

In [11]:

print(sw > 20)

print()

print(sw[sw > 20])

mon False

tue True

wed False

thr True

fri True

sat True

sun False

dtype: bool

tue 30

thr 25

fri 30

sat 40

dtype: int64

In [54]:

sw \* 2

sw = sw \* 2

print(sw)

mon 40

tue 60

wed 40

thr 50

fri 60

sat 80

sun 0

dtype: int64

In [55]:

print(sw)

mon 40

tue 60

wed 40

thr 50

fri 60

sat 80

sun 0

dtype: int64

In [56]:

sw = sw /2

print(sw)

mon 20.0

tue 30.0

wed 20.0

thr 25.0

fri 30.0

sat 40.0

sun 0.0

dtype: float64

In [57]:

'mon' **in** sw

Out[57]:

True

In [62]:

*#taking input from numpy arrays*

**import** **numpy** **as** **np**

ar = np.array([3,2,4,5,6])

ind = np.array( ['a', 'b', 'c', 'd', 'e'])

obj2 = pd.Series(ar, index = ind)

print(obj2)

a 3

b 2

c 4

d 5

e 6

dtype: int64

In [69]:

sdata = {"Sindh": 35000, "Panjab": 4500, "KPK": 3000, "Balochistan": 2000}

tax\_by\_state = pd.Series(sdata)

print(tax\_by\_state)

print(tax\_by\_state.index)

Sindh 35000

Panjab 4500

KPK 3000

Balochistan 2000

dtype: int64

Index(['Sindh', 'Panjab', 'KPK', 'Balochistan'], dtype='object')

In [79]:

sdata = {"Sindh": 35000, "Panjab": 45000, "KPK": 30000, "Balochistan": 20000}

tax\_by\_state = pd.Series(sdata, index = ["Panjab", "Sindh", "KPK", "Balochistan", "GB"])

print(tax\_by\_state)

print(pd.isnull(tax\_by\_state))

Panjab 45000.0

Sindh 35000.0

KPK 30000.0

Balochistan 20000.0

GB NaN

dtype: float64

Panjab False

Sindh False

KPK False

Balochistan False

GB True

dtype: bool

In [97]:

tax\_by\_state.name= "state tax paying capicity"

tax\_by\_state.index.name = "states name"

print(tax\_by\_state)

print(tax\_by\_state.index)

states name

Panjab 45000.0

Sindh 35000.0

KPK 30000.0

Balochistan 20000.0

GB NaN

Name: state tax paying capicity, dtype: float64

Index(['Panjab', 'Sindh', 'KPK', 'Balochistan', 'GB'], dtype='object', name='states name')

In [95]:

*# NaN values mean, values does not exist in pandas series*

pd.isnull(tax\_by\_state)

tax\_by\_state.isnull()

Out[95]:

states name

Panjab False

Sindh False

KPK False

Balochistan False

GB True

Name: state tax paying capicity, dtype: bool

In [99]:

sw = pd.Series([20, 30, 20, 25, 30, 40, 0],

index = ['mon', 'tue', 'wed', 'thr','fri', 'sat', 'sun'])

print(sw)

sw.index = ["m", "t", "w", "t", "f", "s", "s"]

print(sw)

mon 20

tue 30

wed 20

thr 25

fri 30

sat 40

sun 0

dtype: int64

m 20

t 30

w 20

t 25

f 30

s 40

s 0

dtype: int64

**Source for series data**

* direct data in the Series method
* from numpy array or list
* from dictonary