Pandas

August 19, 2022

[1]:
```
#data Manipulation --->Pandas
#1D--> series
#2D--> dataframe
```

[2]:
```
import pandas as pd
import numpy as np
ar=np.array([2,3,4,5,6])
print(ar)
print(ar.shape)
p1=pd.Series(ar)    #-- to convert a data to series it enhances the␣
 ↪visualization aspect of data, provides op with "label index"
print(p1)
p1.index=['A','B','C','D','E']  #-- to change index labal for better␣
 ↪understanding
print(p1)
print(type(p1))
print(p1['D'])
print(p1[3])
```

```
[2 3 4 5 6]
(5,)
0    2
1    3
2    4
3    5
4    6
dtype: int64
A    2
B    3
C    4
D    5
E    6
dtype: int64
<class 'pandas.core.series.Series'>
5
5
```

[3]:
```
#create series data
s1=pd.Series([2,3,4,5],index=['A','B','C','D'])
print(s1)
```

```
A    2
B    3
C    4
D    5
dtype: int64
```

[4]:
```
# access elements
#iloc : acces using index
#loc : access using label
s1=pd.Series([2,3,4,5],index=['A','B','C','D'])
print(s1)
s2=s1.loc['A':'C'] # last value is inclusive
print(s2)
s3=s1.iloc[0:2] # last value is exclusive
print(s3)
```

```
A    2
B    3
C    4
D    5
dtype: int64
A    2
B    3
C    4
dtype: int64
A    2
B    3
dtype: int64
```

[5]:
```
# example for iloc and loc using numbers as label index
s1=pd.Series([2,3,4,5])
print(s1)
s2=s1.loc[0:2] # last value is inclusive
print(s2)
s3=s1.iloc[0:2] # last value is exclusive
print(s3)
```

```
0    2
1    3
2    4
3    5
dtype: int64
0    2
1    3
```

```
2    4
dtype: int64
0    2
1    3
dtype: int64
```

[6]:
```
#creating dataframe using dictionary
data=pd.DataFrame({'Name':['koh','se','kor'],'Age':[20,30,40],'Country':
 ↪['US','IND','Japan']})
print(data)
```

```
  Name  Age Country
0  koh   20      US
1   se   30     IND
2  kor   40   Japan
```

[7]:
```
#access data
data=pd.DataFrame({'Name':['koh','se','kor'],'Age':[20,30,40],'Country':
 ↪['US','IND','Japan']})
print(data)
y=data['Name'] #-- access single column[]
print(y)
print(type(y))
z= data[['Name','Country']]  #--- access multiple columns use [[]]
print(z)
print(type(z))
```

```
  Name  Age Country
0  koh   20      US
1   se   30     IND
2  kor   40   Japan
0    koh
1     se
2    kor
Name: Name, dtype: object
<class 'pandas.core.series.Series'>
  Name Country
0  koh      US
1   se     IND
2  kor   Japan
<class 'pandas.core.frame.DataFrame'>
```

[8]:
```
#loc method
data=pd.DataFrame({'Name':['koh','se','kor'],'Age':[20,30,40],'Country':
 ↪['US','IND','Japan']})
print(data)
y=data.loc[0:1,['Name','Country']]
```

```
print(y)
z=data.loc[[0,2],['Name','Country']]
print(z)
```

```
  Name  Age Country
0  koh   20      US
1   se   30     IND
2  kor   40   Japan
  Name Country
0  koh      US
1   se     IND
  Name Country
0  koh      US
2  kor   Japan
```

[11]:
```
# iloc
data=pd.DataFrame({'Name':['koh','se','kor'],'Age':[20,30,40],'Country':
 ↪['US','IND','Japan']})
print(data)
y = data.iloc[0:3,:]
print(y)
z= data.iloc[0:3,[0,2]]
print(z)
```

```
  Name  Age Country
0  koh   20      US
1   se   30     IND
2  kor   40   Japan
  Name  Age Country
0  koh   20      US
1   se   30     IND
2  kor   40   Japan
  Name Country
0  koh      US
1   se     IND
2  kor   Japan
```

[16]:
```
#add a column
data=pd.DataFrame({'Name':['koh','se','kor','Raina','MSD'],'Age':
 ↪[20,30,40,50,60],'Country':['US','IND','Japan','Ind','ind']})
print(data)
data['Gender']=['M','F','M','M','M'] #-- to add a column but we can add only␣
 ↪one column at a time
print(data)
data['Martial Status']=['Y','N','Y','N','N'] # this is series data so we add it␣
 ↪to dataframe
print(data)
```

```
x=data.loc[2:4,['Name','Age','Gender']]
print(x)
y=data.iloc[2:5,[0,1,3]]
print(y)
```

```
    Name  Age Country
0    koh   20      US
1     se   30     IND
2    kor   40   Japan
3  Raina   50     Ind
4    MSD   60     ind
    Name  Age Country Gender
0    koh   20      US      M
1     se   30     IND      F
2    kor   40   Japan      M
3  Raina   50     Ind      M
4    MSD   60     ind      M
    Name  Age Country Gender Martial Status
0    koh   20      US      M              Y
1     se   30     IND      F              N
2    kor   40   Japan      M              Y
3  Raina   50     Ind      M              N
4    MSD   60     ind      M              N
    Name  Age Gender
2    kor   40      M
3  Raina   50      M
4    MSD   60      M
    Name  Age Gender
2    kor   40      M
3  Raina   50      M
4    MSD   60      M
```