Q1 - Pandas Series

1. How to create Series with nd array

2. How to create Series with Mutable index

3. Creating a series from a Dictionary

4. Print all the values of the Series by multiplying them by 2.

5. Print Square of all the values of the series.

6. Print all the values of the Series that are greater than2

7. Addition of two series

8. Print the first and last 5 elements of a series

9. Print the values from index 0 to 5

10. Selection Using loc, iloc index label

11. Retrieve subsets of data using slicing

Code and output

1.

```
import pandas as pd

import numpy as np
arr=np.array([10,15,18,22])
s = pd.Series(arr)
print(s)
```

Output

```
0    10
1    15
2    18
3    22
dtype: int64
```

CodeText

2.

```
import pandas as pd
s = pd.Series(50,index =[0,1,2,3,4])
print(s)
```

Output

```
0    50
1    50
2    50
3    50
4    50
dtype: int64
```

CodeText

3

```
import pandas as pd
d = {'Name':'Hardik','Iplteam':'MI','Runs':'1500'}
s = pd.Series(d)
print(s)
```

Output

```
Name       Hardik
Iplteam        MI
Runs         1500
dtype: object
```

4

```
import pandas as pd
s=pd.Series([1,2,3,4,5])
print("To multiply all the values in the series by 2")
print('****************************************')
print(s*2)
```

Output

To multiply all the values in the series by 2

```
****************************************
0    2
1    4
2    6
3    8
4   10
dtype: int64
```

5

```
import pandas as pd
s=pd.Series([1,2,3,4,5])
print("To find square all the values in the series")
print('****************************************')
print(s**2)
print("To print all the values in the series by greater than 2")
print('****************************************')
print(s[s>2])
```

Output

```
To find square all the values in the series
****************************************
0    1
1    4
2    9
3   16
4   25
```

6

```
import pandas as pd
s=pd.Series([1,2,3,4,5])
```

```
print("To print all the values in the series by greater than 2")
print('****************************************')
print(s[s>2])
```

Output

```
dtype: int64
To print all the values in the series by greater than 2
****************************************
2  3
3  4
4  5
dtype: int64
```

7

```
import pandas as pd
s1=pd.Series([1,2,3,4,5],index=['a','b','c','d','e'])
s2=pd.Series([10,20,30,40,50],index=['a','b','c','d','e'])
s3=pd.Series([5,14,23,32],index=['a','b','c','d'])
print('Add series1 & series2')
print(s1+s2)
print('Add series2 & series3')
print(s2+s3)
print('to add series2 & series3 and filled non matching index with 0')
print(s2.add(s3,fill_value=0))
```

Output

```
Add series1 & series2
a  11
b  22
c  33
```

d 44

e 55

dtype: int64

Add series2 & series3

a 15.0

b 34.0

c 53.0

d 72.0

e NaN

dtype: float64

to add series2 & series3 and filled non matching index with 0

a 15.0

b 34.0

c 53.0

d 72.0

e 50.0

dtype: float64

8

a)

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77,42,48,97])
s=pd.Series(arr)
print(s.head())
print(s.head(5))
```

Output

0 10

1 15

2 18

3 22

4 55

dtype: int64

b)

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77,42,48,97])
s=pd.Series(arr)
print(s.tail())
print(s.tail(5))
```

Output

4 55

5 77

6 42

7 48

8 97

dtype: int64

9

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77,88,99])
s=pd.Series(arr)
print(s)
print(s.loc[0:5])
```

Output

0 10

1 15

2 18

3 22

4 55

5 77

6 88

7 99

dtype: int64

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

10

a)

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77])
s=pd.Series(arr)
print(s)
print(s.loc[:2])
print(s.loc[3:4])
s.loc[2:3]
```

Output

0 10

```
1    15
2    18
3    22
4    55
5    77
6    88
7    99
dtype: int64
0    10
1    15
2    18
3    22
4    55
5    77
dtype: int64
```

b)

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77])
s=pd.Series(arr)
print(s)
print(s.iloc[:2])
print(s.iloc[3:4])
s.iloc[2:3]
```

Output

```
0    10
1    15
2    18
```

```
3    22
4    55
5    77
dtype: int64
0    10
1    15
dtype: int64
3    22
dtype: int64
2    18
dtype: int64
```

c)

```
import pandas as pd
import numpy as np
arr=np.array(['a','b','c','d'])
s=pd.Series(arr,index=['first','second','third','fourth'])
print(s)
print('\n index in the series are::')
print(s.index)
```

Output

```
first     a
second    b
third     c
fourth    d
dtype: object

 index in the series are::
Index(['first', 'second', 'third', 'fourth'], dtype='object')
```

CodeText

11

```
import pandas as pd
import numpy as np
arr=np.array([10,15,18,22,55,77])
s=pd.Series(arr,index=['a','b','c','d','e','f'])
print(s)
print(s[1:5:2])
print(s[0:6:2])
```

Output

```
a    10
b    15
c    18
d    22
e    55
f    77
dtype: int64
b    15
d    22
dtype: int64
a    10
c    18
e    55
dtype: int64
```

Q2 Dataframe

1. create Dataframe From Series

2. DataFrame from List of Dictionaries

3. Display the first 5 rows of data frame

4. Select the last two columns of the data frame

5. Add two data frames

6. Demonstrate deletion, and renaming of columns

7. Demonstrate concat, Merge operations in data frame

8. Write a Pandas program to join the two given dataframes along rows and assign all data

Test Data:

student_data1:

student_id name marks

0 S1 Danniella Fenton 200

1 S2 Ryder Storey 210

2 S3 Bryce Jensen 190

3 S4 Ed Bernal 222

4 S5 Kwame Morin 199

student_data2:

student_id name marks

0 S4 Scarlette Fisher 201

1 S5 Carla Williamson 200

2 S6 Dante Morse 198

3 S7 Kaiser William 219

4 S8 Madeeha Preston 201

1.

```
import pandas as pd
s = pd.Series(['a','b','c','d'])
```

```
df=pd.DataFrame(s)
print(df)
```

Output

```
  0
0 a
1 b
2 c
3 d
```

CodeText

2.

```
import pandas as pd
l=[{'Name':'Sachin','SirName':'Bhardwaj'},
   {'Name':'Vinod','SirName':'Verma'},
   {'Name':'Rajesh','SirName':'Mishra'}]
df1=pd.DataFrame(l)
print(df1)
```

Output

```
     Name   SirName
0  Sachin  Bhardwaj
1   Vinod     Verma
2  Rajesh    Mishra
```

3.

```
import pandas as pd
l=[{'Name':'Sachin','SirName':'Bhardwaj'},
   {'Name':'Vinod','Sirname':'verma'},
   {'Name':'Sankar','SirName':'Mahadev'},
   {'Name':'Meth','SirName':'Singh'},
   {'Name':'Surbi','SirName':'Chandra'} ]
df1=pd.DataFrame(l)
print(df1)
for(row_index,row_value) in df1.iterrows():
   print('\n Row index is::',row_index)
   print('Row value is::')
   print(row_value)
```

Output

```
     Name   SirName Sirname
0  Sachin  Bhardwaj     NaN
1   Vinod       NaN   verma
2  Sankar   Mahadev     NaN
3    Meth     Singh     NaN
4   Surbi   Chandra     NaN
```

```
 Row index is:: 0
Row value is::
Name          Sachin
SirName    Bhardwaj
Sirname          NaN
Name: 0, dtype: object

 Row index is:: 1
Row value is::
Name       Vinod
SirName      NaN
Sirname    verma
Name: 1, dtype: object

 Row index is:: 2
Row value is::
Name         Sankar
SirName    Mahadev
Sirname          NaN
Name: 2, dtype: object

 Row index is:: 3
Row value is::
Name         Meth
SirName    Singh
Sirname      NaN
Name: 3, dtype: object

 Row index is:: 4
Row value is::
Name          Surbi
SirName    Chandra
Sirname          NaN
Name: 4, dtype: object
```

4

```
import pandas as pd
l=[{'Name':'Sachin','SirName':'Bhardwaj'},
   {'Name':'Vinod','Sirname':'verma'}]
df1=pd.DataFrame(l)
print(df1)
for(col_name,col_value) in df1.iteritems():
  print('\n')
  print('column name:',col_name)
  print('column values:')
  print(col_value)
```

Output

```
  Name   SirName Sirname
```

```
0  Sachin  Bhardwaj     NaN
1   Vinod       NaN   verma


column name: Name
column values:
0    Sachin
1     Vinod
Name: Name, dtype: object


column name: SirName
column values:
0    Bhardwaj
1         NaN
Name: SirName, dtype: object


column name: Sirname
column values:
0      NaN
1    verma
Name: Sirname, dtype: object
```

5

```
import pandas as pd
s=pd.Series([10,15,18,22])
df=pd.DataFrame(s)
df.columns=['list1']
df['list2']=20
df['list3']=df['list1']+df['list2']
print(df)
```

Output

```
list1  list2  list3

0     10     20     30

1     15     20     35

2     18     20     38

3     22     20     42
```

6

```
import pandas as pd
s=pd.Series([10,15,18,22])
df=pd.DataFrame(s)
df.columns=['list1']
df['list2']=20
```

```
df['list3']=df['list1']+df['list2']
print(df)
del df['list3']
print(df)
```

Output

```
list1  list2  list3
0     10     20     30
1     15     20     35
2     18     20     38
3     22     20     42
   list1  list2
0     10     20
1     15     20
2     18     20
3     22     20
```

7

```
import pandas as pd
dic1= { 'id': ['1','2','3','4','5'],'Value1': ['A','C','E','G','I'],
       'Value2':['B','D','f','H','J']}
dic2={'id':['2','3','6','7','8'],'Value1':['K','M','O','Q','S'],
      'Value2':['L','N','P','R','T']}
df1=pd.DataFrame(dic1)
df2=pd.DataFrame(dic2)
df3=pd.concat([df1,df2])
print(df3)
```

Output

```
id Value1 Value2
0  1      A      B
1  2      C      D
2  3      E      f
3  4      G      H
4  5      I      J
0  2      K      L
1  3      M      N
2  6      O      P
3  7      Q      R
4  8      S      T
```

```
import pandas as pd
dic1= { 'id': ['1','2','3','4','5'],'Value1': ['A','C','E','G','I'],
       'Value2':['B','D','f','H','J']}
dic2={'id':['2','3','6','7','8'],'Value1':['K','M','O','Q','S'],
      'Value2':['L','N','P','R','T']}
dic3= {'id':['1','2','3','4','5','7','8','9','10','11'],
       'Value3':[12,13,14,15,16,17,15,12,13,23]}
df1=pd.DataFrame(dic1)
df2=pd.DataFrame(dic2)
df3=pd.concat([df1,df2])
df4=pd.DataFrame(dic3)
df5=pd.merge(df3,df4,on='id')
print(df5)
```

Output

```
id Value1 Value2  Value3
0 1     A     B     12
1 2     C     D     13
2 2     K     L     13
3 3     E     f     14
4 3     M     N     14
5 4     G     H     15
6 5     I     J     16
7 7     Q     R     17
8 8     S     T     15
```

CodeText

8.

a)

```
import pandas as pd
student_data1={
    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Stor
ey','Bryce Jenson','Ed Bernal','Kwame Morin'],
    'marks':['200','210','190','222','199']
}
student_data2={
    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Wil
liamson','Dante Morse','Kaiser William','Madeeha Preston'],
    'marks':['201','200','198','219','201']
}

df1=pd.DataFrame(student_data1)
df1=pd.DataFrame(student_data2)
df3=pd.merge(df1,df2,on='id',how='outer')
```

```
print(df3)
```

Output

```
id              name marks Value1 Value2
0  s4  Scarlette Fisher   201    NaN    NaN
1  s5  Carla Williamson   200    NaN    NaN
2  s6       Dante Morse   198    NaN    NaN
3  s7    Kaiser William   219    NaN    NaN
4  s8   Madeeha Preston   201    NaN    NaN
5   2               NaN   NaN      K      L
6   3               NaN   NaN      M      N
7   6               NaN   NaN      O      P
8   7               NaN   NaN      Q      R
9   8               NaN   NaN      S      T
```

b)

```
import pandas as pd
student_data1={
    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Stor
ey','Bryce Jenson','Ed Bernal','Kwame Morin'],
    'marks':['200','210','190','222','199']
}
student_data2={
    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Wil
liamson','Dante Morse','Kaiser William','Madeeha Preston'],
    'marks':['201','200','198','219','201']
}

df1=pd.DataFrame(student_data1)
df1=pd.DataFrame(student_data2)
df3=pd.merge(df1,df2,on='id',how='inner')
print(df3)
```

Output

```
Empty DataFrame
Columns: [id, name, marks, Value1, Value2]
Index: []
```

c)

```
import pandas as pd
student_data1={
    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Stor
ey','Bryce Jenson','Ed Bernal','Kwame Morin'],
    'marks':['200','210','190','222','199']
}
```

```
student_data2={
    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Wil
liamson','Dante Morse','Kaiser William','Madeeha Preston'],
    'marks':['201','200','198','219','201']
}

df1=pd.DataFrame(student_data1)
df1=pd.DataFrame(student_data2)
df3=pd.merge(df1,df2,on='id',how='right')
print(df3)
```

Output

id	name	marks	Value1	Value2
0	2	NaN	NaN	K	L
1	3	NaN	NaN	M	N
2	6	NaN	NaN	O	P
3	7	NaN	NaN	Q	R
4	8	NaN	NaN	S	T

CodeText

```
d)
import pandas as pd
student_data1={
    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Stor
ey','Bryce Jenson','Ed Bernal','Kwame Morin'],
    'marks':['200','210','190','222','199']
}
student_data2={
    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Wil
liamson','Dante Morse','Kaiser William','Madeeha Preston'],
    'marks':['201','200','198','219','201']
}

df1=pd.DataFrame(student_data1)
df1=pd.DataFrame(student_data2)
df3=pd.merge(df1,df2,on='id',how='left')
print(df3)
```

Output

id	name	marks	Value1	Value2
0	s4	Scarlette Fisher	201	NaN	NaN
1	s5	Carla Williamson	200	NaN	NaN
2	s6	Dante Morse	198	NaN	NaN
3	s7	Kaiser William	219	NaN	NaN
4	s8	Madeeha Preston	201	NaN	NaN

CodeText

e)

```
import pandas as pd
student_data1={
    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Stor
ey','Bryce Jenson','Ed Bernal','Kwame Morin'],
    'marks':['200','210','190','222','199']
}
student_data2={
    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Wil
liamson','Dante Morse','Kaiser William','Madeeha Preston'],
    'marks':['201','200','198','219','201']
}

df1=pd.DataFrame(student_data1)
df1=pd.DataFrame(student_data2)
df3=pd.merge(df1,df2,right_index=True,left_index=True)
print(df3)
```

Output

	id_x	name	marks	id_y	Value1	Value2
0	s4	Scarlette Fisher	201	2	K	L
1	s5	Carla Williamson	200	3	M	N
2	s6	Dante Morse	198	6	O	P
3	s7	Kaiser William	219	7	Q	R
4	s8	Madeeha Preston	201	8	S	T