

**DATA SCIENCE LAB**

**Experiment No.: 1**

**Aim**

**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

**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

**Procedure**

**1**. **How to create Series with nd array**

import pandas as pd

import numpy as np

arr=np.array([4,5,8,2,9])

s=pd.Series(arr)

print(s)

**2.How to create Series with Mutable index**

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)

**3.** **Creating a series from a Dictionary**

import pandas as pd

d={'Name':'Sree','Class':'RMCA','RollNo':41}

s=pd.Series(d)

print(s)

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

import pandas as pd

s=pd.Series([1,2,3,4,5,6])

print("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

print("Multiply by 2")

print(s\*2)

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

print("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

print("Square of the numbers")

print(s\*\*2)

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

print("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

print("greater number")

print(s[s>2])

**7.** **Addition of two series**

import pandas as pd

s1=pd.Series([1,2,3,4],index=['a','b','c','d'])

s2=pd.Series([6,7,8,10],index=['a','b','c','d'])

print("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

print("Add s1 and s2")

print(s1+s2)

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

import pandas as pd

import numpy as np

arr=np.array([1,2,3,4,5,6,7,8,9])

s=pd.Series(arr)

print(s)

print(s.head(1))

print(s.tail())

**9.** **Print the values from index 0 to 5**

import pandas as pd

import numpy as np

arr=np.array([1,2,3,4,5,6,7,8,9])

s=pd.Series(arr)

print(s)

print(s.loc[:5])

**10.** **Selection Using loc, iloc index label**

import pandas as pd

s=pd.Series(['a','b','c','d','e','f','g','h','i','j','k'])

print(s.loc[1:5])

print(s.iloc[1:5])

**11.Retrieve subsets of data using slicing**

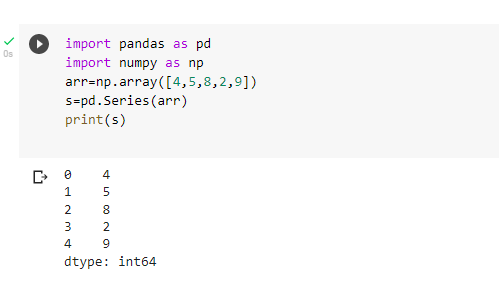
import pandas as pd

s=pd.Series(['a','b','c','d','e','f','g','h','i','j','k'])

print(s[1:10:2])

**Output Screenshot**

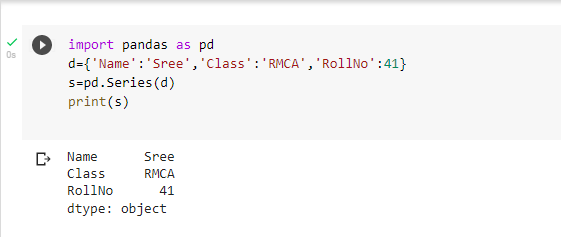
**1.**

****

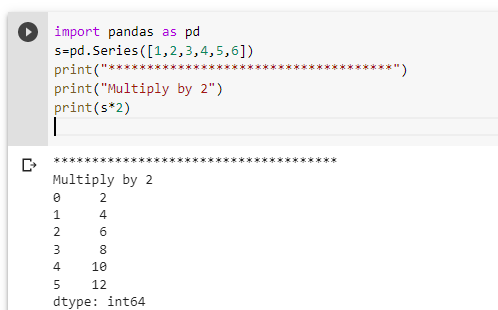
**2.**

****

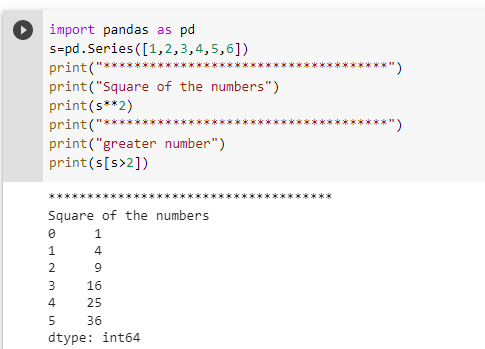
**3.**

****

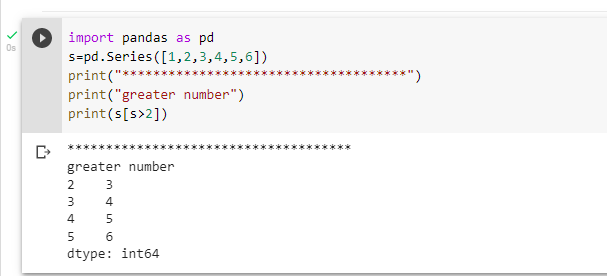
**4.**

****

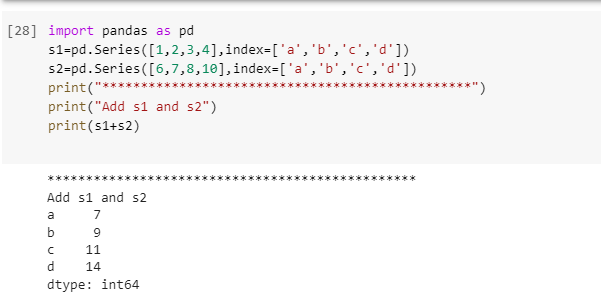
**5.**

****

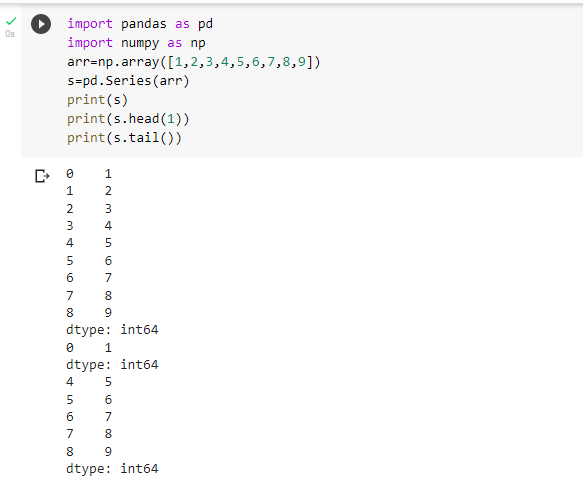
**6.**

****

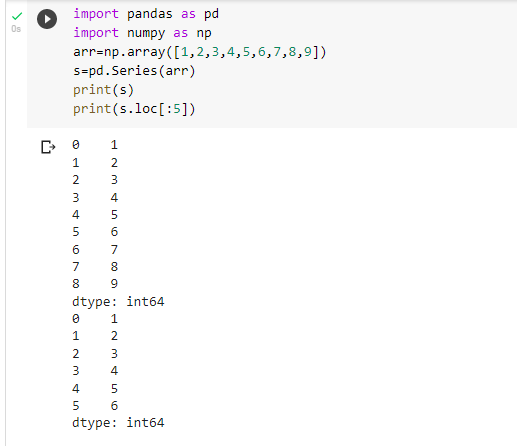
**7.**

****

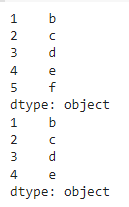
**8.**

****

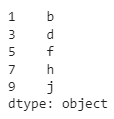
**9.**

****

**10.**

****

**11.**

****

**1.create Dataframe From Series**

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

**2. DataFrame from List of Dictionaries**

import pandas as pd

l=[{'Name':'sachin','city':'kerala'},

{'Name':'virat','city':'tamilnadu'}]

d=pd.DataFrame(l)

print(d)

**Output**

Name city

0 sachin kerala

1. virat tamilnadu

**3.Display the first 5 rows of data frame**

import pandas as pd

empdata = {'empid':[1,2,3,4,5,6],'ename':['Vimal','Sachin','Bav','Kumar','Ravy','Sunil']}

df=pd.DataFrame(empdata)

print(df)

print(df.head(5))

**Output**

empid ename

0 1 Vimal

1 2 Sachin

2 3 Bav

3 4 Kumar

4 5 Ravy

5 6 Sunil

empid ename

0 1 Vimal

1 2 Sachin

2 3 Bav

3 4 Kumar

4. 5 Ravy

**4.Select the last two columns of the data frame**

import pandas as pd

empdata = {'empid':[1,2,3,4,5,6], 'ename':['Vimal','Sachin','Bav','Kumar','Ravy','Sunil']}

df=pd.DataFrame(empdata)

print(df)

df.loc[0:5]

print(df.tail(2))

**Output**

empid ename

0 1 Vimal

1 2 Sachin

2 3 Bav

3 4 Kumar

4 5 Ravy

5 6 Sunil

empid ename

4 5 Ravy

5 6 Sunil

**6. Demonstrate deletion, and renaming of columns**

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)

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

**7 Demonstrate concat, Merge operations in data frame**

import pandas as pd

s= pd.Series([10,20,30,40])

df=pd.DataFrame(s)

df.columns=['List1']

df['List2']=40

df1=df.drop('List2',axis=1)

df2=df.drop(index=[2,3],axis=0)

print(df)

print(" After deletion::")

print(df1)

print (" After row deletion::")

print(df2)

**Output**

List1 List2

0 10 40

1 20 40

2 30 40

3 40 40

After deletion::

List1

0 10

1 20

2 30

3 40

After row deletion::

List1 List2

0 10 40

1. 20 40

**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

**Out put**

Original DataFrames:

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\_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

Join the said two dataframes along rows:

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

0 S4 Scarlette Fisher 201

1 S5 Carla Williamson 200

2 S6 Dante Morse 198

3 S7 Kaiser William 219

4 S8 Madeeha Preston 20