

AKSHAT SHRIVASTAVA

ROLL- 2305591

CSE-35

AD-LAB-05

Assignment No: 0.6 (Pandas)

1. Write a Pandas program to add, subtract, multiple and divide two Pandas Series.

```
import pandas as pd
s1 = pd.Series([1,2,3,4,5])
s2 = pd.Series([9, 8, 6, 5])

print('Addition:\n',s1+s2)
print('Subtraction:\n', s1-s2)
print('Multiplication:\n', s1 * s2)
print('Division:\n', s1/s2)
```

```
Addition:
0      10.0
1      10.0
2       9.0
3       9.0
4       NaN
dtype: float64
Subtraction:
0      -8.0
1      -6.0
2      -3.0
3      -1.0
4       NaN
dtype: float64
```

```
Multiplication:
0      9.0
1     16.0
2     18.0
3     20.0
4      NaN
dtype: float64
Division:
0      0.111111
1      0.250000
2      0.500000
3      0.800000
4      NaN
dtype: float64
```

2. Write a Pandas program to convert a dictionary to a Pandas series.

```
import pandas as pd
s1 = pd.Series([1,2,3,4,5])
d= s1.to_dict()
d = {"a": 10, "b": 20, "c": 30}
s2=pd.Series(d)
print(s2)
print(type(s2))

a      10
b      20
c      30
dtype: int64
<class 'pandas.core.series.Series'>
```

3. Write a Pandas program to convert a NumPy array to a Pandas series.

```
import numpy as np
import pandas as pd
arr= np.array([1,2,3,4,5,6,7,8,9])
s=pd.Series(arr)
print(s)
print(type(s))

0      1
1      2
```

```

2    3
3    4
4    5
5    6
6    7
7    8
8    9
dtype: int64
<class 'pandas.core.series.Series'>

```

4. Write a Pandas program to convert a given Series to an array.

```

import pandas as pd
s1 = pd.Series([1,2,3,4,5])
s2 = pd.Series([9, 8, 6, 5])
s3=s2.to_numpy()
print(s3)
print(type(s3))

[9 8 6 5]
<class 'numpy.ndarray'>

```

5. Write a Pandas program to convert Series of lists to one Series.

```

import pandas as pd
s_lists = pd.Series([[ 'Red', 'Green'], [ 'Blue'], [ 'Yellow', 'White']])
single_series = s_lists.explode().reset_index(drop=True)
print(single_series)

0    Red
1    Green
2    Blue
3    Yellow
4    White
dtype: object

```

6. Write a Pandas program to create a dataframe from a dictionary and display it.

```

import pandas as pd
d = { 'Col1': [1, 2], 'Col2': [3, 4]}
df = pd.DataFrame(data=d)
print(df)

```

| | Col1 | Col2 |
|---|------|------|
| 0 | 1 | 3 |
| 1 | 2 | 4 |

7. Write a Pandas program to get the first 3 rows of a given DataFrame.

```
import pandas as pd

df = pd.DataFrame({
    "A": [1, 2, 3, 4, 5],
    "B": [10, 20, 30, 40, 50]
})

result = df.head(3)

print(result)
```

| | A | B |
|---|---|----|
| 0 | 1 | 10 |
| 1 | 2 | 20 |
| 2 | 3 | 30 |

8. Write a Pandas program to select the 'name' and 'score' columns from the following DataFrame. Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima',  
'Katherine', 'James', 'Emily', 'Michael',  
'Matthew', 'Laura',  
  
'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan,  
8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes',  
'no', 'no', 'yes']}
```

```
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

Expected Output:

Select specific columns:

name score

a Anastasia 12.5

b Dima 9.0

c Katherine 16.5

... h Laura NaN

i Kevin 8.0

j Jonas 19.0

```
import pandas as pd
import numpy as np

exam_data = {
    'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily',
            'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
    'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
    'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
    'qualify': ['yes', 'no', 'yes', 'no', 'no',
               'yes', 'yes', 'no', 'no', 'yes']
}

labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

df = pd.DataFrame(exam_data, index=labels)

result = df[['name', 'score']]

print("Select specific columns:")
print(result)
```

Select specific columns:

| | name | score |
|---|-----------|-------|
| a | Anastasia | 12.5 |
| b | Dima | 9.0 |
| c | Katherine | 16.5 |
| d | James | NaN |
| e | Emily | 9.0 |
| f | Michael | 20.0 |
| g | Matthew | 14.5 |
| h | Laura | NaN |
| i | Kevin | 8.0 |
| j | Jonas | 19.0 |

9. Write a Pandas program to select the rows where the score is missing, i.e. is NaN. Sample Python dictionary data and list labels:

```

exam_data = {'name': ['Anastasia', 'Dima',
'Katherine', 'James', 'Emily', 'Michael',
'Matthew', 'Laura',

'Kevin', 'Jonas'],

'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan,
8, 19],

'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],

'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes',
'no', 'no', 'yes']}

labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', labels = ['a', 'b',
'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

```

Expected Output:

Rows where score is missing:

attempts name qualify score

d 3 James no NaN

h 1 Laura no NaN

```
import pandas as pd
import numpy as np

exam_data = {
    'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily',
             'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
    'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
    'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
    'qualify': ['yes', 'no', 'yes', 'no', 'no',
                'yes', 'yes', 'no', 'no', 'yes']
}

labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

df = pd.DataFrame(exam_data, index=labels)

result = df[df['score'].isna()]

print("Rows where score is missing:")
print(result)
```

Rows where score is missing:

| | name | score | attempts | qualify |
|---|-------|-------|----------|---------|
| d | James | NaN | 3 | no |
| h | Laura | NaN | 1 | no |

10. Write a Pandas program to append a new row 'k' to data frame with given values for each column. Now

delete the new row and return the original DataFrame. Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima',  
'Katherine', 'James', 'Emily', 'Michael',  
'Matthew', 'Laura',  
  
'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan,  
8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes',  
'no', 'no', 'yes']}  
  
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

Values for each column will be:

name : "Suresh", score: 15.5, attempts: 1,
qualify: "yes", label: "k" Expected Output:

Append a new row:

Print all records after insert a new record:

attempts name qualify score

a 1 Anastasia yes 12.5

b 3 Dima no 9.0

..... j 1 Jonas yes 19.0

k 1 Suresh yes 15.5

```
import pandas as pd
import numpy as np

exam_data = {
    'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily',
             'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
    'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
    'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
    'qualify': ['yes', 'no', 'yes', 'no', 'no',
                'yes', 'yes', 'no', 'no', 'yes']
}

labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

df = pd.DataFrame(exam_data, index=labels)

df.loc['k'] = ['Suresh', 15.5, 1, 'yes']
```

```
print("Print all records after insert a new record:")
print(df)
```

```
df = df.drop('k')
```

```
print("\nDataFrame after deleting row 'k':")
print(df)
```

Print all records after insert a new record:

| | name | score | attempts | qualify |
|---|-----------|-------|----------|---------|
| a | Anastasia | 12.5 | 1 | yes |
| b | Dima | 9.0 | 3 | no |
| c | Katherine | 16.5 | 2 | yes |
| d | James | NaN | 3 | no |
| e | Emily | 9.0 | 2 | no |
| f | Michael | 20.0 | 3 | yes |
| g | Matthew | 14.5 | 1 | yes |
| h | Laura | NaN | 1 | no |
| i | Kevin | 8.0 | 2 | no |
| j | Jonas | 19.0 | 1 | yes |
| k | Suresh | 15.5 | 1 | yes |

DataFrame after deleting row 'k':

| | name | score | attempts | qualify |
|---|-----------|-------|----------|---------|
| a | Anastasia | 12.5 | 1 | yes |
| b | Dima | 9.0 | 3 | no |
| c | Katherine | 16.5 | 2 | yes |
| d | James | NaN | 3 | no |
| e | Emily | 9.0 | 2 | no |
| f | Michael | 20.0 | 3 | yes |
| g | Matthew | 14.5 | 1 | yes |
| h | Laura | NaN | 1 | no |
| i | Kevin | 8.0 | 2 | no |
| j | Jonas | 19.0 | 1 | yes |