Pandas - Series & Dataframes

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import glob
import re
import math
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"


import warnings
warnings.filterwarnings("ignore")
```

Series

Create Series

```
# Create series from Nump Array
v = np.array([1,2,3,4,5,6,7])
s1 = pd.Series(v)
s1
```

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
dtype: int32
```

```
#Datatype of Series
s1.dtype
```

```
dtype('int32')
```

```
# number of bytes allocated to each item
s1.itemsize
```

```
C:\Users\DELL\Anaconda3\lib\site-packages\ipykernel_launcher.py:2: FutureWarning: Ser

4
```

```
# Number of bytes consumed by Series
s1.nbytes
```

```
28
```

```
# Shape of the Series
s1.shape
```

```
(7,)
```

```
# number of dimensions
s1.ndim
```

```
1
```

```
# Length of Series
len(s1)
```

```
7
```

```
s1.count()
```

```
7
```

```
s1.size
```

```
7
```

```
# Create series from List
s0 = pd.Series([1,2,3],index = ['a','b','c'])
s0
```

```
a    1
b    2
c    3
dtype: int64
```

```
# Modifying index in Series
s1.index = ['a' , 'b' , 'c' , 'd' , 'e' , 'f' , 'g']
s1
```

```
a    1
b    2
c    3
d    4
e    5
f    6
g    7
dtype: int32
```

```
# Create Series using Random and Range function
v2 = np.random.random(10)
ind2 = np.arange(0,10)
s = pd.Series(v2,ind2)
v2 , ind2 , s
```

```
(array([0.87790351, 0.21256923, 0.2833476 , 0.84976498, 0.17274437,
        0.36953613, 0.92661933, 0.13005525, 0.25394528, 0.43563311]),
 array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9]),
 0    0.877904
 1    0.212569
 2    0.283348
 3    0.849765
 4    0.172744
 5    0.369536
 6    0.926619
 7    0.130055
 8    0.253945
 9    0.435633
 dtype: float64)
```

```
# Creating Series from Dictionary
dict1 = {'a1' :10 , 'a2' :20 , 'a3':30 , 'a4':40}
s3 = pd.Series(dict1)
s3
```

```
a1    10
a2    20
a3    30
a4    40
dtype: int64
```

```
pd.Series(99, index=[0, 1, 2, 3, 4, 5])
```

```
0    99
1    99
2    99
3    99
4    99
5    99
dtype: int64
```

Slicing Series

[ ] ↳ *10 cells hidden*

Append Series

[ ] ↳ *7 cells hidden*

Operation on Series

[ ] ↳ *20 cells hidden*

DataFrame

[ ] ↳ *140 cells hidden*

Merge Dataframes

[ ] ↳ *34 cells hidden*

Working with Dates in Pandas

[ ] ↳ *38 cells hidden*

Importing multiple CSV files in DataFrame

[ ] ↳ *33 cells hidden*

LIKE OPERATION IN PANDAS

[ ] ↳ *8 cells hidden*

Regex in Pandas dataframe

[ ] ↳ *6 cells hidden*

Replace values in dataframe

[ ] ↳ *11 cells hidden*

Group By

[ ] ↳ *11 cells hidden*

Loading Data in Chunks

[ ] ↳ *3 cells hidden*

Stack & unstack in Pandas

[ ] ↳ *5 cells hidden*

PIVOT Tables

[ ] ↳ *3 cells hidden*

Hierarchical indexing

[ ] ↳ *13 cells hidden*

Crosstab in Pandas

[ ] ↳ *4 cells hidden*

Row & Column Bind

[ ] ↳ *9 cells hidden*