Worksheet-1&2 (2024-2025)

Name of the Student : **Sub : Informatics Practices (065)**

Class: XII **Date of Issue : 30/03/2024**

Topic: Unit 1 – Pandas (Series & DataFrame) **Date of Submission:**

I. Multiple Choice Questions

1. Which of the following statement will import pandas library?

(a) Import pandas as pd (b) import Pandas as py

(c) import pandas as pd (d) import panda as pd

2. To create an empty Series object, you can use:

(a) pd.Series(empty) (b) pd.Series(np.NaN)

(c) pd.Series( ) (d) all of these.

3. To specify datatype int16 for a Series object, you can write:

(a) pd.Series (data=array, dtype=int16) (b) pd.Series (data=array, dtype=numpy.int16)

(c) pd.Series (data=array.dtype=pandas.int16) (d) all of the above.

4. To get the number of dimensions of a Series object, _____ attribute is displayed.

(a) index (b) size (c) itemsize (d) ndim

5. To get the size of the datatype of the items in Series object, you can display _____ attribute.

(a) index (b) size (c) itemsize (d) ndim

6. To get the number of elements in a Series object, _____ attribute may be used.

(a) index (b) size (c) itemsize (d) ndim

7. To get the number of bytes of the Series data, _____ attribute is displayed.

(a) hasnans (b) nbytes (c) ndim (d) dbyte

8. To check if the Series object contains NaN values, _____ attribute is displayed.

(a) hasnans (b) nbytes (c) ndim (d) dbyte

9. To display third element of a Series object S, you will write _____

(a) S[:3] (b) S[2] (c) S[3] (d) S[:2]

10. To display first three elements of a Series object S, you may write _____

(a) S[:3] (b) S[3] (c) S[3rd] (d) all of these

11. To display last five rows of a Series object S, you may write _____

(a) head( ) (b) head(5) (c) tail( ) (d) Tail(5)

12. Which of the following statement is wrong?

(a) We can't change the index of the series.

(b) We can easily convert the list, tuple, and dictionary into a Series.

(c) A series represents a single column in memory.

(d) We can create empty Series.

13. What type of error is returned by the following statement?

```
import pandas as pa
pa.Series([1,2,3,4], index=['a','b','c'])
```

(a) Value Error (b) Syntax Error (c) Name Error (d) Logical Error

14. To display last five rows of a series object 'S', you may write:

(a) S.Head( ) (b) S.Tail(5) (c) S.Head(5) (d) S.tail( )

15. Missing data in Pandas object is represented through:

(a) Null (b) None (c) Missing (d) NaN

16. Given a Pandas series called Sequences, the command which will display the first 4 rows is _____

(a) print(Sequences.head(4)) (b) print(Sequences.Head(4))

(c) print(Sequences.heads(4)) (d) print(Sequences.Heads(4))

II. (a) Create the following Series Object "SCapital".

```
Hyd          Telangana
Chennai      Tamilnadu
Mumbai     Maharastra
dtype: object
```

(b) Create a Series "MyWord" with the individual characters of the word "Welcome"

```
0    W
1    e
2    l
3    c
4    o
5    m
6    e
dtype: object
```

(c) Consider the following Series "Market". Write the outputs:

```
45    Carrot
20    Potato
25     Onion
60    Chilly
dtype: object
```

(i) Market[1:4] (ii) Market[0:4:2] (iii) Market[3:1:-1]

(d) Create a Series Object "FullMarks" with the following specifications.
Index should be multiples of 3 between 20, use range function.
Give all the values as 500.

(e)Create a series by using following ndarrays.

(i) Use arange( ) function to get values from 10 to 20, with each difference 2.5
(ii) Use linspace( ) function to get values from 10 to 20. Total values should be 5
(iii)Use tile( ) function to repeat [2,3,4] for 2 times.

III. Mention attribute names and its values for the following Series "S"

```
45    Carrot
20    Potato
25     Onion
60    Chilly
dtype: object
```

Worksheet 2 – DataFrames Concept

I. Multiple Choice Questions

(1) Which of the following statement is wrong in context of DataFrame?
(a) Two dimensional size is Mutable.
(b) Can perform arithmetic operations on rows and columns
(c) Homogeneous tabular data structure.
(d) Create DataFrame from numpy ndarray.

(2) When we create a DataFrame from a list of Dictionaries the columns labels are formed by the
(a) Union of the keys of the dictionaries.
(b) Intersection of the keys of the dictionaries.
(c) Union of the values of the dictionaries.
(d) Intersection of the values of the dictionaries.

(3) If a DataFrame is created using a 2D dictionary, then the indexes/row labels are formed from _____
(a) dictionary's values (b) inner dictionary's keys
(c) outer dictionary's keys (d) none of these.

(4) If a DataFrame is created using a 2D dictionary, then the column labels are formed from _____
(a) dictionary's values (b) inner dictionary's keys
(c) outer dictionary's keys (d) none of these.

(5) Which of the following can be used to specify the data while creating a DataFrame?
(a) Series (b) List of Dictionaries (c) Structured ndarray (d) All of these.

(6) The axis 0 identifies a dataframe's _____
(a) rows (b) columns (c) values (d) datatypes

(7) The axis 1 identifies a dataframe's ____
(a) rows (b) columns (c) values (d) datatypes

(8) To get the number of elements in a dataframe, _____ attribute may be used.
(a) size (b) shape (c) values (d) ndim

(9) To get NumPy representation of a dataframe, _____ attribute may be used.
(a) size (b) shape (c) values (d) ndim

(10) To get a number representing number of axes in a dataframe, _____ attribute may be used.

(a) size (b) shape (c) values (d) ndim

(11) Which attribute is not used with DataFrame?

(a) Size (b) Type (c) Empty (d) Columns

(12) To get the transpose of a dataframe D1, you can write _____

(a) D1.T (b) D1.Transpose (c) D1.Swap (d) All of these

(13) To extract row/column from a database, _____ function may be used.

(a) row( ) (b) column( ) (c) loc( ) (d) All of these

(14) To display the 3rd, 4th and 5th columns from the 6th to 9th rows of a dataframe DF, you can write _____.

(a) DF.loc[6:9,3:5] (b) DF.loc[6:10,3:6] (c) DF.iloc[6:10,3:6] (d) DF.iloc[6:9,3:5]

(15) To change the 5th column's value at 3rd row as 35 in dataframe DF, you can wirte _____

(a) DF[4,6]=35 (b) DF[3,5]=35 (c) DF.iat[4,6]=35 (d) DF.iat[3,5]=35

(16) Which among the following options can be used to create a DataFrame in Pandas?

(a) A scalar value (b) An ndarray (c) A python dict (d) All of these

(17) Identify the correct statement:

(a) The standard marker for missing data in Pandas is NaN.

(b) Series act in a way similar to that of an array.

(c) Both (a) and (b)

(d) None of the above.

(18) Identify the correct option to select first four rows and second to fourth columns from a DataFrame : "Data" :

(a) display(Data.iloc[1:4,2:4]) (b) display (Data.iloc[1:5,2:5])

(c) print(Data.iloc[0:4,1:4]) (d) print(Data.iloc[1:4,2:4])

(19) To delete a column from a DataFrame, you may use _____ statement.

(a) remove (b) del (c) drop (d) cancel

(20) To delete a row from a DataFrame, you may use _____statement.

(a) remove (b) del (c) drop (d) cancel

(21) Sudhanshu has written the following code to create a DataFrame with Boolean index:

```
import numpy as np
import pandas as pd
df=pd.DataFrame (data=[[5,6,7]], index=[true, false, true])
print (df)
```

While executing the code, she is getting an error, help her to rectify the code:

(a) df=pd.DataFrame ([True, False, True], data=[5,6,7])

(b) df=pd.DataFrame (data=[5,6,7] ,index=[True, False, True])

(c) df=pd.DataFrame ([true, false, true], data=[5,6,7])

(d) df=pd.DataFrame (index=[true, false, true], data=[5,6,7])

II. Create the following DataFrames:

	Teacher	Subject	Exp
One	Akash	Maths	3
Two	Suresh	English	15
Three	Mohan	IP	5

	Private	Aided	Govt	ZP
AP	100	75	125	89
TN	98	92	130	92
TS	110	85	110	91

	RNo	SName	Marks
First	51	Lahari	55
Second	52	Chanakya	62
Third	53	Harish	52
Fourth	54	Neha	75

(i) DFTeacher (ii) DFSchool (iii) DF Student

III. Create a DataFrame in Python from the given list:

[['Divya','HR',95000],['Mamta','Marketing',97000], ['Payal','IT',980000], ['Deepak','Sales',79000]]

Also give appropriate column headings as shown below:

	Name	Department	Salary
0	Divya	HR	95000
1	Mamta	Marketing	97000
2	Payal	IT	980000
3	Deepak	Sales	79000

IV. Mention attribute names and its values for the following DataFrame "DF"

	P1	P2
Mon	IP	Maths
Tue	Che	Phy
Wed	IP	Phy

V. Question on DataFrame Operations.

	P1	P2	P3
Mon	IP	Maths	Eng
Tue	Che	Phy	Acc
Wed	IP	Phy	Maths

(1) Create the above DataFrame "TT"

(2) Display only the details of First 2 Rows.

	P1	P2	P3
Mon	IP	Maths	Eng
Tue	Che	Phy	Acc

(3) Display the details of rows Wed and Mon (in the order Wed, Mon)

	P1	P2	P3
Wed	IP	Phy	Maths
Mon	IP	Maths	Eng

(4) Display the details of Columns "P1" and "P3"

	P1	P3
Mon	IP	Eng
Tue	Che	Acc
Wed	IP	Maths

(5) Display the following output.

	P3	P1	P2
Tue	Acc	Che	Phy
Mon	Eng	IP	Maths

(6) Display the following Output.

	P3	P2
Mon	Eng	Maths
Tue	Acc	Phy
Wed	Maths	Phy

(7) Change Tuesday's 3rd Period to "Eng"

	P1	P2	P3
Mon	IP	Maths	Eng
Tue	Che	Phy	Eng
Wed	IP	Phy	Maths

(8) Add a new column "P4" with values Acc,BS,Maths

	P1	P2	P3	P4
Mon	IP	Maths	Eng	Acc
Tue	Che	Phy	Eng	BS
Wed	IP	Phy	Maths	Maths

(9) Add a row with index Thu, values Bio,Che,Eng,Acc

	P1	P2	P3	P4
Mon	IP	Maths	Eng	Acc
Tue	Che	Phy	Eng	BS
Wed	IP	Phy	Maths	Maths
Thu	Bio	Che	Eng	Acc

(10) Rename Mon to Monday

	P1	P2	P3	P4
Monday	IP	Maths	Eng	Acc
Tue	Che	Phy	Eng	BS
Wed	IP	Phy	Maths	Maths
Thu	Bio	Che	Eng	Acc

(11) Rename P2 to Period2

	P1	Period2	P3	P4
Monday	IP	Maths	Eng	Acc
Tue	Che	Phy	Eng	BS
Wed	IP	Phy	Maths	Maths
Thu	Bio	Che	Eng	Acc

(12) Delete the Column P3

	P1	Period2	P4
Monday	IP	Maths	Acc
Tue	Che	Phy	BS
Wed	IP	Phy	Maths
Thu	Bio	Che	Acc

(13) Delete Row with Index "Tue"

	P1	Period2	P4
Monday	IP	Maths	Acc
Wed	IP	Phy	Maths
Thu	Bio	Che	Acc

(14) Modify the Entire P4 values to “Games”

	P1	Period2	P4
Monday	IP	Maths	Acc
Wed	IP	Phy	Acc
Thu	Bio	Che	Acc

(15) Delete complete DataFrame