Worksheets

**pandas**

Total questions: 118

Worksheet time: 1hrs 10mins

Instructor name: Mohd. Khan

Name

Class

Date

1. Which is not a feature of series

a) Mutable data

b) Multiple rows

c) Immutable size

d) Homogeneous data

2. Dataframe have

a) Mutable size

b) Mutable data

c) All of above

d) 2D Array

3. Which package should be needed for series

a) Random

b) Maths

c) Statistic

d) Pandas

4. Full form of NaN is

a) Not a Number

b) Not a Null

c) None of these

d) Not a Numeric

5. Dataframe can be created using

a) List
b) All of these
c) Array
d) Dictionary

6. Full form of CSV file is

a) Common Separated Value
b) Common System Value
c) Comma Separated Vault
d) Comma Separated Value

7. Not a function of Dataframe

a) loc()
b) Tail()
c) Head()
d) multi()

8. which functions used to transfer data from dataframe to CSV files

a) from_dataframe()
b) df_csv()
c) to_csv()
d) to_data()

9. Dataframe can contain multiple series

a) False
b) True

10. Which is true for series.

a) Size is mutable,Values is mutable
b) none
c) size is immutable,values is mutable
d) size is mutable,values is immutable.

11. Series.tail(3) will return how many values.

a) 5 values
b) none
c) 3 values from front
d) 3 values from last

12. Series.head() will return how many rows.

a) 5
b) 4
c) 2
d) 3

13. To extract subset from Series,the following function is used

a) row()
b) all
c) loc()
d) column()

14. we can analyze the data in pandas with :

a) Both
b) none
c) Series
d) Dataframe

15. Series in Pandas is

a) 3 Dimensional array
b) 1 Dimensional Array
c) 2 Dimensional array
d) none of above

16. Minimum number of argument we require to pass in pandas series ?

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

17. In data science, which of the python library are more popular ?

a) pandas
b) numpy
c) django
d) none

18. Which is not a feature of series

a) Multiple rows
b) Homogeneous data
c) Immutable size
d) Mutable data

19. Series can be created from

a) Dictionary
b) Array
c) All of them
d) Scatter value

20. Which package should be needed for series

a) Maths
b) Pandas
c) Statistic
d) Random

21. Full form of NaN is

a) Not a Number
b) None of these
c) Not a Numeric
d) Not a Null

22. Data structures in Pandas can be mutated in the terms of ____ but not of _____.

a) none of the above
b) size, value
c) value, size
d) semantic, size

23. pandas is a:

a) Dataframe
b) Data Structure
c) Library
d) Series

24. Write the output for the following:

```
import pandas as pd1
s = pd1.Series(5, index=[0, 1, 2, 3])
print(s)
```

a)
```
0 5
1 5
2 5
3 5
dtype: object
```

b)
```
1 5
2 5
2 5
4 5
dtype: int64
```

c)
```
0 5
1 5
2 5
3 5
dtype: int64
```

d)
```
0 5
```

25. write the output:

```
import pandas as pd1
s = pd1.Series([1,2,3])
t = pd1.Series([1,2,4])
u=s-t
print (u)
```

a)
```
0 0
1 0
2 -1
dtype: int64
```

b)
```
0 0
1 0
2 1
dtype: int64
```

c)
```
0 1
1 0
2 1
dtype: int64
```

d)
```
0 0
1 0
2 -1
dtype: float64
```

26. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.iloc[2:4])
```

a) none of the above

b)
```
c 3
d 4
dtype: int64
```

c)
```
b 2
c 3
d 4
dtype: int64
```

d)
```
b 3
c 4
dtype: int64
```

27. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.loc['b':'d'])
```

a)
```
b 2
c 3
d 4
dtype: object
```

b)
```
c 3
d 4
dtype: int64
```

c)
```
b 2
c 3
d 4
dtype: int64
```

d)
```
b 2
c 3
d 4
dtype: float64
```

28. Series is

a) A one dimensional structure

b) A two dimensional structure

c) A three dimensional structure

d) None of the Above

29. Data Frames is

a) Two Dimensional

b) None of the above

c) One Dimensional

d) Three Dimensional

30. Series is

a) Immutable

b) Mutable

31. Which is correct line to import pandas

a) import Pandas as pd

b) import pandas as pd

c) import panda as pd

32. s = pd.series([1,2,3,4,5],index=['a','b','c','d','e'])
print(s[:3] gives ?

a) a 1
b 2
c 3
d 4

b) a 1
b 2
c 3

c) a 3
b 4
c 5

33. series = pd.series(55)
print(series) gives ?

a) 1 51
2 52
3 53
4 54
5 55

b) 0 55
1 55
2 55
3 55
4 55

c) 0 51
1 52
2 53
3 54
4 55

34. s = pd.series(range(1,15,3),index = [x for x in 'abcde'])
print(s) gives ?

a) a 0
b 3
c 6
d 9
e 12

b) x 1
x 4
x 7
x 10
x 13

c) a 1
b 4
c 7
d 10
e 13

35. series = pd.series({'jan':31,"feb':29,'march':31})
print(series) gives ?

a) 0 31
1 29
2 31

b) jan 31
feb 29
march 31

c) 0 jan
1 feb
2 march

36. Data Frame contains ?

a) Data of same Types

b) Data of Different Types

37. s1 = pd.series([11,12,13,14])
s2 = pd.series([11,12,13,14],index=[1,2,3,4])
print(s1+s2) gives ?

a) Error

b) 0 22
1 24
2 26
3 28

c) 0 11
1 12
2 13
3 14
0 11
1 12
2 13
3 14

38. Best way to import the pandas module in your program ?

a) 3.from pandas import *

b) 1.import pandas

c) 4.All of the above

d) 2.import pandas as pd

39. Which is true for series.

a) Size is mutable,Values is mutable

b) size is immutable,values is mutable

c) none

d) size is mutable,values is immutable.

40. Series.tail(3) will return how many values.

a) 3 values from last
b) none
c) 5 values
d) 3 values from front

41. Series.head() will return how many rows.

a) 5
b) 4
c) 2
d) 3

42. To extract subset from Series,the following function is used

a) row()
b) all
c) loc()
d) column()

43. we can analyze the data in pandas with :

a) Dataframe
b) none
c) Series
d) Both

44. Series in Pandas is

a) 1 Dimensional Array
b) 2 Dimensional array
c) 3 Dimensional array
d) none of above

45. Minimum number of argument we require to pass in pandas series ?

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

46. Series can be created from

a) All of them
b) Array
c) Scatter value
d) Dictionary

47. Full form of NaN is

a) Not a Number
b) None of these
c) Not a Null
d) Not a Numeric

48. pandas is a:

a) Dataframe
b) Data Structure
c) Library
d) Series

49. Write the output for the following:

```
import pandas as pd1
s = pd1.Series(5, index=[0, 1, 2, 3])
print(s)
```

a) 0 5

b) 0 5
1 5
2 5
3 5
dtype: int64

c) 1 5
2 5
2 5
4 5
dtype: int64

d) 0 5
1 5
2 5
3 5
dtype: object

50. write the output:

```
import pandas as pd1
s = pd1.Series([1,2,3])
t = pd1.Series([1,2,4])
u=s-t
print (u)
```

a)
```
0 0
1 0
2 -1
dtype: int64
```

b)
```
0 0
1 0
2 -1
dtype: float64
```

c)
```
0 0
1 0
2 1
dtype: int64
```

d)
```
0 1
1 0
2 1
dtype: int64
```

51. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.iloc[2:4])
```

a)
```
b 2
c 3
d 4
dtype: int64
```

b)
```
c 3
d 4
dtype: int64
```

c) none of the above

d)
```
b 3
c 4
dtype: int64
```

52. DataFrame is .....

a) size mutable, data mutable

b) size immutable, data mutable

53. An empty DataFrame can be created by.....

a) passing arguments

b) without passing arguments

54. Which of the functions can be used to delete column/row from a DataFrame?

a) drop()

b) iloc()

c) at()

d) pop()

55. Full form of CSV file is

a) Comma Separated Value

b) Comma Separated Vault

c) Common Separated Value

d) Common System Value

56. Dataframe can contain multiple series

a) True

b) False

57. Which is the correct Pandas syntax to read in a csv file and assign it to a DataFrame df?

a) df = with open('file.csv') as pd.DataFrame

b) df = read_csv('file.csv')

c) df = read('file.csv', type = 'csv')

d) df = pd.read_csv('file.csv')

58. Which is true for series.

a) none
b) Size is mutable,Values is mutable
c) size is immutable,values is mutable
d) size is mutable,values is immutable.

59. Series.tail(3) will return how many values.

a) 5 values
b) 3 values from last
c) 3 values from front
d) none

60. Series.head() will return how many rows.

a) 2
b) 5
c) 4
d) 3

61. To extract subset from Series,the following function is used

a) column()
b) all
c) row()
d) loc()

62. we can analyze the data in pandas with :

a) Both
b) Dataframe
c) Series
d) none

63. Series in Pandas is

a) 2 Dimensional array
b) 1 Dimensional Array
c) 3 Dimensional array
d) none of above

64. Minimum number of argument we require to pass in pandas series ?

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

65. Which is not a feature of series

a) Homogeneous data
b) Mutable data
c) Immutable size
d) Multiple rows

66. Series can be created from

a) Dictionary
b) Array
c) All of them
d) Scatter value

67. Which package should be needed for series

a) Pandas
b) Random
c) Statistic
d) Maths

68. Full form of NaN is

a) None of these

b) Not a Number

c) Not a Numeric

d) Not a Null

69. Write the output for the following:

```
import pandas as pd1
s = pd1.Series(5, index=[0, 1, 2, 3])
print(s)
```

a)
```
1 5
2 5
2 5
4 5
dtype: int64
```

b)
```
0 5
1 5
2 5
3 5
dtype: int64
```

c) 0 5

d)
```
0 5
1 5
2 5
3 5
dtype: object
```

70. write the output:

```
import pandas as pd1
s = pd1.Series([1,2,3])
t = pd1.Series([1,2,4])
u=s-t
print (u)
```

a)
```
0 0
1 0
2 1
dtype: int64
```

b)
```
0 0
1 0
2 -1
dtype: int64
```

c)
```
0 1
1 0
2 1
dtype: int64
```

d)
```
0 0
1 0
2 -1
dtype: float64
```

71. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.iloc[2:4])
```

a) none of the above

b)
```
b 3
c 4
dtype: int64
```

c)
```
c 3
d 4
dtype: int64
```

d)
```
b 2
c 3
d 4
dtype: int64
```

72. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.loc['b':'d'])
```

a) b 2
c 3
d 4
dtype: object

b) b 2
c 3
d 4
dtype: float64

c) c 3
d 4
dtype: int64

d) b 2
c 3
d 4
dtype: int64

73. Pandas is_____.

a) a Python 2D plotting library which produces publication-quality figures in a variety of hardcopy formats and interactive environments across platforms.

b) a Python library that implements a range of machine learning, preprocessing, cross-validation and visualization algorithms using a unified interface.

c) a Python library that is built on NumPy and provides easy-to-use data structures and data analysis tools for the Python programming language.

74. Give the output of the following code:

```
>>>import pandas as pd
>>>dict1 = {'AR' : 100, 'VR' : 200, 'AI' : 300}
>>>ser = pd.Series(dict1)
>>>print(ser[1])
```

a) 200

b) 100

c) AR

d) VR

75. Consider the following code for creating a series:

```
import pandas as pd
dict1 = {'AR' : 100, 'VR' : 200, 'AI' : 300, 'DS':400, 'NLP':500}
ser = pd.Series(dict1)
```

What will be the print statement to get the following output:

```
AI      300
NLP     500
dtype: int64
```

a) print(ser[[2,4]])

b) print(ser[[2:4]])

c) print(ser[2,4])

d) print(ser[2,3,4])

76. Which of the following commands is used to install Pandas?

a) python install python

b) pip install pandas

c) python install pandas

d) pip install python-pandas

77. Consider the following code for creating a series:

```
import pandas as pd
dict1 = {'AR' : 100, 'VR' : 200, 'AI' : 300, 'DS':400, 'NLP':500}
ser = pd.Series(dict1)
```

What will be the print statement to get the following output:

```
VR    200
AI    300
dtype: int64
```

(1 Point)

a) print(ser[1:3])

b) print(ser[[1,3]])

c) print(ser[1,3])

d) print(ser[[1:3]])

78. Have you enjoyed the quiz on Pandas Series?

a) No

b) Yes

79. Missing data in panda series and dataframe can be filled with a _____ value.

Ans. _____

80. Basic feature of series are

a) immutable size, immutable data

b) Homogeneous data  Size Immutable  Values of Data Mutable

c) Hetrogenous data  Size Immutable  Values of Data Mutable

81. DataFrame is _____

a) like a two dimensional array with heterogeneous data

b) module

c) photo frame with data

82. Which function from the options given below can read the dataset from a large text file?

a) read_csv
b) read_pickle
c) read_json
d) read_hdf

83. Which among the following options can be used to create a DataFrame in Pandas?

a) All of the above
b) A python dict
c) An ndarray
d) A scalar value

84.
```
>>> df3.loc[:,df3.isnull().any()]
```

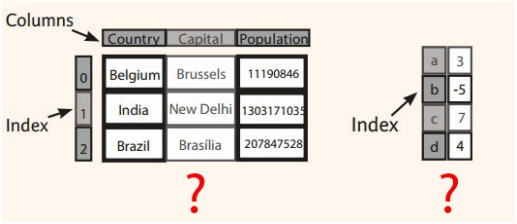
What does this line of code do?

a) Selects cols with NaN
b) Selects cols without NaN
c) Selects cols with any vals >1
d) Selects cols with vals > 1

85.
```
>>> df3.fillna(df3.mean())
```

What does this line of code do?

a) Replaces values with others
b) Drops NaN values
c) Fills NaN values with a predetermined value

86. 


_____ is a two-dimensional labelled data structure with columns of potentially different types, while_____ is a one-dimensional labelled array capable of holding any data type

a) Series, DataFrame

b) DataFrame, Series

87. ____iterates over the DataFrame columns, returning a tuple with the column name and the content as a Series.

____returns a tuple with row index and row data as a Series object.

a) df.iterrows(), df.items()

b) df.iteritems(), df.iterrows()

88.
```
>>> df2.duplicated('Type')
```

What does this line of code do?

a) Checks index duplicates

b) Checks duplicates

c) Drops duplicates

d) Returns unique values

89. Write the output for the following:

```
import pandas as pd1
s = pd1.Series(5, index=[0, 1, 2, 3])
print(s)
```

a)
```
0 5
1 5
2 5
3 5
dtype: int64
```

b)
```
1 5
2 5
2 5
4 5
dtype: int64
```

c) 0 5

d)
```
0 5
1 5
2 5
3 5
dtype: object
```

90. write the output:

```
import pandas as pd1
s = pd1.Series([1,2,3])
t = pd1.Series([1,2,4])
u=s-t
print (u)
```

a)
```
0 0
1 0
2 -1
dtype: int64
```

b)
```
0 1
1 0
2 1
dtype: int64
```

c)
```
0 0
1 0
2 -1
dtype: float64
```

d)
```
0 0
1 0
2 1
dtype: int64
```

91. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.iloc[2:4])
```

a)
```
b 3
c 4
dtype: int64
```

b)
```
b 2
c 3
d 4
dtype: int64
```

c)
```
c 3
d 4
dtype: int64
```

d) none of the above

92. write the output:

```
import pandas as pd
s=pd.Series([1,2,3,4],index=['a','b','c','d'])
print(s.loc['b':'d'])
```

a) b 2
c 3
d 4
dtype: float64

b) b 2
c 3
d 4
dtype: int64

c) c 3
d 4
dtype: int64

d) b 2
c 3
d 4
dtype: object

93. To extract subset from Series,the following function is used

a) row()

b) all

c) column()

d) loc()

94. df['List3']=df['List1']+df['List2']

a) sytax not correct

b) Can not use arithmatic operators

c) Syntax is correct

95. print(df.tail()) will give output

a) first five rows

b) only last 01 row

c) last five rows

d) Error in syntax

96. print(df.iloc[5]) will give output

a) Error in syntax

b) only 5th row

c) 0-4 rows

d) first five rows

97. Predict the output:
obj2=pd.Series([3.5,5.,6.5,8.])
print(obj2.size,obj2.hasnans)

a) 3 True

b) 4 False

c) 4 True

d) 4 F

98. What will be the output of following code?
**stu={'A':44,'B':44,'C':45,"D":47}**
**s8=pd.Series(stu)**
**print(s8[ :2]*100)**

a) **A 4300**
**B 3000**
**dtype: int64**

b) **A 4300**
**B 4000**
**dtype: int64**

c) **A 3300**
**B 4000**
**dtype: int64**

d) **A 4400**
**B 4000**
**dtype: int64**

99. A Dataframe contains Heterogeneous data

a) true

b) false

100. A Dataframe Data is Immutable

a) false

b) true

101. A Dataframe Size is Mutable

a) true

b) false

102. A Dataframe has axes column index (axis=0) row index (axes=1)

a) true

b) false

103. A data frame can be created using:

a) Series

b) Lists

c) Dictionary

d) A numpy 2D array

104. df['Tid'] & df.Tid are same

a) false

b) true

105. **Full form of NaN is**

a) None of these

b) Not a Number

c) Not a Null

d) Not a Numeric

106. **Which of the following commands is used to convert array named "grades" into data frame named "df_grades"?**

a) df_grades = pd.DataFrame("grades")

b) df_grades = pd.DataFrame(grades)

c) df_grades = grades

d) grades = pd.DataFrame(df_grades)

107. Type the syntax that returns the top 5 rows in DataFrame df with the native Pandas function (not slicing):

Ans. _____

108. Which of the functions can be used to delete column/row from a DataFrame?

a) iloc()

b) pop()

c) at()

109. Which of the following statement/s will give 3 rows from bottom of the dataframe ?

a) print(df.tail(3))

b) print(df.tail[3])

c) Print(df.tail())

110. i) Which of the following statement will delete rank2 row from the dataframe ?

a) df.del(“rank2”)

b) delete df(rank2)

c) df.drop(rank2)

d) df.drop(‘rank2’)

111. The instructor wants to add a new column, Marks to the dataframe. The values of the marks will be 12, 22, 21, 24. Help him to choose the correct command to do so.

a) a) Df.columns=[12, 22, 21, 24]
b) c) df.loc[marks] = [12, 22, 21, 24]
c) b) df['marks'] = [12, 22, 21, 24]
d) d) Both (b) and (c) are correct

112. Is boolean Indexing possible in Data Frame

a) true
b) false

113. The axis=1 identifies a DataFrame's _____

a) Rows
b) Data Types
c) Columns
d) Values

114. _____ attribute is used to specify column labels

a) columns()
b) columns
c) column ()
d) column

115. To get number of elements in a DataFrame _____ attribute may be used.

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

116. To extract a row / column from a DataFrame ___ function may be used

a) All of the above
b) row()
c) column()
d) loc()

117. The insert function requires ____ number of arguments in DataFrame.

a) 2
b) 4
c) 1
d) 3

118.

| | Name | Age |
|---|---|---|
| 111 | Ram | 34 |
| 222 | Syam | 38 |
| 333 | Rohit | 36 |

For the given DataFrame df, what will be the code to get the value 38?

a) None of the above
b) df . iloc [1,1]
c) df . loc [111]
d) df . Age [222]

**Answer Keys**

1. b) Multiple rows
2. c) All of above
3. d) Pandas
4. a) Not a Number
5. b) All of these
6. d) Comma Separated Value
7. d) multi()
8. c) to_csv()
9. b) True
10. c) size is immutable,values is mutable
11. d) 3 values from last
12. a) 5
13. c) loc()
14. a) Both
15. b) 1 Dimensional Array
16. a) 1
17. a) pandas
18. a) Multiple rows
19. c) All of them
20. b) Pandas
21. a) Not a Number
22. c) value, size
23. c) Library
24. c) 0 5 1 5 2 5 3 5 dtype: int64
25. a) 0 0 1 0 2 -1 dtype: int64
26. b) c 3 d 4 dtype: int64
27. c) b 2 c 3 d 4 dtype: int64
28. a) A one dimensional structure
29. a) Two Dimensional
30. b) Mutable
31. b) import pandas as pd
32. b) a 1 b 2 c 3
33. b) 0 55 1 55 2 55 3 55 4 55
34. c) a 1 b 4 c 7 d 10 e 13
35. b) jan 31 feb 29 march 31
36. a) Data of same Types
37. a) Error
38. c) 4.All of the above
39. b) size is immutable,values is mutable
40. a) 3 values from last
41. a) 5
42. c) loc()
43. d) Both
44. a) 1 Dimensional Array
45. d) 1
46. a) All of them
47. a) Not a Number
48. c) Library

49. b) 0 5 1 5 2 5 3 5 dtype: int64

50. a) 0 0 1 0 2 -1 dtype: int64

51. b) c 3 d 4 dtype: int64

52. a) size mutable, data mutable

53. b) without passing arguments

54. d) pop() , a) drop()

55. a) Comma Separated Value

56. a) True

57. d) df = pd.read_csv('file.csv')

58. c) size is immutable,values is mutable

59. b) 3 values from last

60. b) 5

61. d) loc()

62. a) Both

63. b) 1 Dimensional Array

64. c) 0

65. d) Multiple rows

66. c) All of them

67. a) Pandas

68. b) Not a Number

69. b) 0 5 1 5 2 5 3 5 dtype: int64

70. b) 0 0 1 0 2 -1 dtype: int64

71. c) c 3 d 4 dtype: int64

72. d) b 2 c 3 d 4 dtype: int64

73. c) a Python library that is built on NumPy and provides easy-to-use data structures and data analysis tools for the Python programming language.

74. a) 200

75. a) print(ser[[2,4]])

76. b) pip install pandas

77. a) print(ser[1:3])

78. n/a

79. NAN, not a number

80. b) Homogeneous data  Size Immutable  Values of Data Mutable

81. a) like a two dimensional array with heterogeneous data

82. a) read_csv

83. a) All of the above

84. a) Selects cols with NaN

85. c) Fills NaN values with a predetermined value

86. b) DataFrame, Series

87. b) df.iteritems(), df.iterrows()

88. b) Checks duplicates

89. a) 0 5 1 5 2 5 3 5 dtype: int64

90. a) 0 0 1 0 2 -1 dtype: int64

91. c) c 3 d 4 dtype: int64

92. b) b 2 c 3 d 4 dtype: int64

93. d) loc()

94. c) Syntax is correct

95. c) last five rows

96. b) only 5th row

97. b) 4 False

98. b) **A 4300 B 4000 dtype: int64**

99. a) true

100. a) false

101. a) true

102. b) false

103. b) Lists, c) Dictionary, a) Series, d) A numpy 2D array

104. b) true

105. b) Not a Number

106. b) df_grades = pd.DataFrame(grades)

107. df.head(), df.head(5)

108. b) pop()

109. a) print(df.tail(3))

110. d) df.drop('rank2')

111. c) b) df['marks'] = [12, 22, 21, 24]

112. a) true

113. c) Columns

114. b) columns

115. a) size

116. d) loc()

117. d) 3

118. d) df . Age [222] , b) df . iloc [1,1]