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) Immutable size
c) Multiple rows
d) Homogeneous data

2. Dataframe have

a) 2D Array
b) Mutable size
c) Mutable data
d) All of above

3. Which package should be needed for series

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

4. Full form of NaN is

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

5. Dataframe can be created using

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

6. Full form of CSV file is

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

7. Not a function of Dataframe

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

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

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

9. Dataframe can contain multiple series

a) False
b) True

10. Which is true for series.

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

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

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

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

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

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

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

14. we can analyze the data in pandas with :

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

15. Series in Pandas is

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

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

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

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

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

18. Which is not a feature of series

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

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) Random
b) Maths
c) Statistic
d) Pandas

21. Full form of NaN is

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

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

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

23. pandas is a:

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

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

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

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

c)
```
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: float64
```

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

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

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

28. Series is

a) A one dimensional structure

b) A two dimensional structure

c) None of the Above

d) A three dimensional structure

29. Data Frames is

a) None of the above

b) Three Dimensional

c) Two Dimensional

d) One Dimensional

30. Series is

a) Immutable

b) Mutable

31. Which is correct line to import pandas

a) import pandas as pd

b) import panda as pd

c) import Pandas as pd

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

a) a 3
b 4
c 5

b) a 1
b 2
c 3

c) a 1
b 2
c 3
d 4

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

a) 0 55
1 55
2 55
3 55
4 55

b) 1 51
2 52
3 53
4 54
5 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) x 1
x 4
x 7
x 10
x 13

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

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

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

a) 0 jan
1 feb
2 march

b) jan 31
feb 29
march 31

c) 0 31
1 29
2 31

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) 2.import pandas as pd

c) 4.All of the above

d) 1.import pandas

39. Which is true for series.

a) size is mutable,values is immutable.

b) size is immutable,values is mutable

c) none

d) Size is mutable,Values is mutable

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

a) none

b) 3 values from front

c) 5 values

d) 3 values from last

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

a) 2

b) 5

c) 3

d) 4

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

a) loc()

b) column()

c) all

d) row()

43. we can analyze the data in pandas with :

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

44. Series in Pandas is

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

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

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

46. Series can be created from

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

47. Full form of NaN is

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

48. pandas is a:

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

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

c) 0 5
1 5
2 5
3 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: int64

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

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 3
c 4
dtype: int64

b) none of the above

c) b 2
c 3
d 4
dtype: int64

d) c 3
d 4
dtype: int64

52. DataFrame is .....

a) size immutable, data mutable

b) size mutable, data mutable

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

a) without passing arguments

b) passing arguments

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

a) drop()

b) pop()

c) at()

d) iloc()

55. Full form of CSV file is

a) Comma Separated Vault

b) Common Separated Value

c) Common System Value

d) Comma Separated 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 = pd.read_csv('file.csv')

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

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

58. Which is true for series.

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

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

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

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

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

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

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

62. we can analyze the data in pandas with :

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

63. Series in Pandas is

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

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

66. Series can be created from

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

67. Which package should be needed for series

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

68. Full form of NaN is

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

69. 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) 1 5
2 5
2 5
4 5
dtype: int64

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

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 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
```

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

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

c) none of the above

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: 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
```

73. Pandas is_____.

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

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

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) VR

c) 100

d) AR

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[2,3,4])

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

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

a) pip install python-pandas

b) python install python

c) python install pandas

d) pip install 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) Yes

b) No

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) Hetrogenous data Size Immutable Values of Data Mutable

c) Homogeneous data Size Immutable Values of Data Mutable

81. DataFrame is _____

a) module

b) like a two dimensional array with heterogeneous data

c) photo frame with data

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

a) read_json

b) read_csv

c) read_pickle

d) read_hdf

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

a) An ndarray

b) All of the above

c) A python dict

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 with any vals >1

c) Selects cols without NaN

d) Selects cols with vals > 1

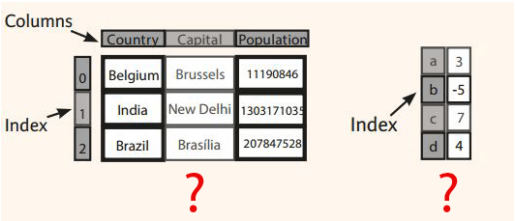
85. `>>> df3.fillna(df3.mean())`

What does this line of code do?

a) Replaces values with others

b) Fills NaN values with a predetermined value

c) Drops NaN values

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) DataFrame, Series

b) Series, DataFrame

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) Drops duplicates

b) Checks duplicates

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

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

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

d) 0 5

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: int64

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

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) none of the above

d) c 3
d 4
dtype: int64

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: int64

b) b 2
c 3
d 4
dtype: float64

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) column()

b) row()

c) all

d) loc()

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

a) sytax not correct

b) Syntax is correct

c) Can not use arithmatic operators

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

a) only last 01 row

b) last five rows

c) first five rows

d) Error in syntax

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

a) 0-4 rows
b) Error in syntax
c) only 5th row
d) first five rows

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

a) 4 F
b) 3 True
c) 4 True
d) 4 False

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 4000
dtype: int64

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

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

d) **A 4300**
B 3000
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) A numpy 2D array
b) Dictionary
c) Lists
d) Series

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

a) true
b) false

105. **Full form of NaN is**

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

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

a) df_grades = grades
b) df_grades = pd.DataFrame("grades")
c) grades = pd.DataFrame(df_grades)
d) df_grades = pd.DataFrame(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())
c) print(df.tail(3))

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

a) df.del(“rank2”)
b) df.drop(rank2)
c) delete df(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) b) df['marks'] = [12, 22, 21, 24]
b) d) Both (b) and (c) are correct
c) a) Df.columns=[12, 22, 21, 24]
d) c) df.loc[marks] = [12, 22, 21, 24]

112. Is boolean Indexing possible in Data Frame

a) false
b) true

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

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

114. _____ attribute is used to specify column labels

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

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

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

a) 2
b) 1
c) 4
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) df . Age [222]

b) df . iloc [1,1]

c) None of the above

d) df . loc [111]