

b. Demonstrate different arithmetic, logical and filtering operations on Dataframes. 8 3 4 2

25. a. Predict different univariate visualization for categorical data with assumption on your own relevant dataset. 8 3 5 2

(OR)

b. Apply the need for dashboards in detail. 8 4 5 2

PART – C (1 × 15 = 15 Marks)

Answer ANY ONE Questions

26. Use Pandas to perform the following operations on the given data label.

- | | | | | |
|---|---|---|---|---|
| i. Calculate the sum of the examination attempts by the students. | 5 | 4 | 4 | 2 |
| ii. Calculate the mean score for each different student in dataframe. | 5 | 4 | 4 | 2 |
| iii. Select the rows where number of attempts in the examination is less than two and score greater than 15 | 5 | 4 | 4 | 2 |

Sample 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],

'quality': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']

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

27. Demonstrate Numpy array with the given set of values and perform the operations. 15 4 3 2

$$X = \begin{bmatrix} 3, 6, 9, 12, \\ 15, 18, 21, 24, \\ 27, 30, 33, 36, \\ 39, 42, 45, 48, \\ 51, 54, 57, 60 \end{bmatrix}$$

- Return the array of elements in the third column (from all rows)
- Return a sub-array comprising of Odd rows and even columns of the above array.
- Build a new array from the above array swapping first and last, second and third columns
- Swap rows and columns in reverse order
- Sum of elements

* * * * *

Reg. No.

B.Tech. DEGREE EXAMINATION, MAY 2023

Third Semester

21CSS202T – FUNDAMENTALS OF DATA SCIENCE

(For the candidates admitted from the academic year 2022-2023 onwards)

Note:

- Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.
- Part - B** and **Part - C** should be answered in answer booklet.

Time: 3 Hours

Max. Marks: 75

PART – A (20 × 1 = 20 Marks)

Answer ALL Questions

- | | | | | |
|--|---|---|---|---|
| 1. What does eye() function in the numpy package return? | 1 | 2 | 1 | 2 |
| (A) A null matrix (B) An identity matrix (C) A diagonal matrix (D) A symmetric matrix with only 1s and 0s | | | | |
| 2. Which of the following statement is false? | 1 | 2 | 1 | 2 |
| (A) ndarray is also known as Axis array (B) ndarray.dataitem size is the buffer containing actual elements of the array (C) Numpy main object is the homogeneous multidimensional array (D) In numpy, dimensions are called axes | | | | |
| 3. How to convert numpy array to list? | 1 | 1 | 1 | 2 |
| (A) array.list() (B) array.list (C) list.array() (D) list(array) | | | | |
| 4. Which of the following is not an application for data science? | 1 | 2 | 1 | 2 |
| (A) Recommendation systems (B) Image and speech recognition (C) Online price comparison (D) Privacy checker | | | | |
| 5. What is the purpose of using 'P' in a PDB command? | 1 | 1 | 2 | 2 |
| (A) Display variable (B) Print variable (C) Quit from debug mode (D) Clear the variable | | | | |
| 6. Which module in python supports regular expressions? | 1 | 2 | 2 | 2 |
| (A) re (B) regex (C) pyregex (D) regexpy | | | | |
| 7. What does the following re-search do? | 1 | 1 | 2 | 2 |
| (A) Matches a pattern at the start of the string (B) Matches a pattern at any position in the string (C) Matches a pattern only at the end of string (D) Such a function does not exist | | | | |

8. Json array elements can be accessed by
(A) Name (B) Key
(C) Value (D) Index
9. Which of the following is a valid for loop attributes of Django template system?
(A) forloop.reverse (B) forloop.firstitem
(C) forloop.counter() (D) forloop.lastitem
10. Which of the following password validators are not provided by default Django?
(A) NumericPasswordValidator (B) CommonPasswordValidator
(C) MinimumLengthValidator (D) MaximumLengthValidator
11. What does {{forloop.counter}} prints?
(A) It will not print if for loop variable is not defined
(B) It will count the number of times loop run.
(C) It prints the integer value of number of times the loop executed
(D) It prints the number of times as a string
12. This template {##} is used for
(A) It is comment in template language
(B) It is used for business logic
(C) It will raise an exception
(D) It is used for adding HTML
13. Which of the following statement for creating dataframe is valid?
(A) df=pd.dataframe(dict1) (B) df=pd.DataFrame(dict1)
(C) df=pd.dataFrame(dict1) (D) df=pd.DataFrame(dict1)
14. Which of the following is not an attribute of dataframe?
(A) Axes (B) Empty
(C) Transpose (D) Size
15. What is the output for the following statement?
df.iloc[2:7,:3]
(A) Display 2nd and 7th row and first 3 columns
(B) Display 3rd to 7th row and first 3 columns
(C) Display 2nd and 7th row and first 4 columns
(D) Display 3rd to 7th row and first 4 columns
16. The function to fill in NaN values with 100, 1F DF is the dataframe object
(A) fillna(100) (B) df.fillna(100)
(C) df.fillna(0) (D) df.fill(100)
17. By default hist() uses bin value of
(A) 5 (B) 10
(C) 15 (D) 20
18. Which of the following is used to set the limit of X-axis of the plot?
(A) Limit() (B) Limits()
(C) X Limit() (D) X Limits()

19. Consider the following code to display a bar chart.
X = range(0, 40, 8)
Y = range(10, 100, 10)
plt.bar(x,y)
while executing it, is producing error why?
(A) Both the sequences X and Y do not start with
(B) Both the sequence X and Y are no of same shape
(C) Values produced by range is not considered for chart
(D) Both the sequences X and Y have values with different intervals

20. Which statement is not true about scatter chart?
(A) Also known as correlation chart
(B) It is primarily used to represent relation between two sets of data
(C) The closer the dots to line, the weaker the relation will be
(D) It uses dots to represent data points

PART – B (5 × 8 = 40 Marks)

Answer ALL Questions

Marks BL CO PO

21. a. Construct a python program to perform arithmetic operations on matrices. 8 3 1 2

(OR)

- b. Illustrate in detail about data science process. 8 4 1 2

22. a. Construct a python program with class named bank account which represents a bank account, having attributes.
.accountNumber (numeric type)
.name(name of the account owner as string type)
.balance
create a constructor with parameters : account Number, name, balance.
Also Add a Deposit() method which manages the deposit actions and
Withdrawal() method which manages withdrawal actions. Create a
bankFees () method to apply the bank fees with a percentage of 5% of the
balance account and display it using display() method.

(OR)

- b. Apply CRUD operations in Sqlite with suitable samples. 8 4 2 2

23. a. Employ suitable examples to demonstrate HTTP request and HTTP response. 8 3 3 2

(OR)

- b. Write in detail about views with a suitable example in Django. 8 4 3 2

24. a. Illustrate in detail about the general techniques for handling large volume of data. 8 4 4 2

(OR)