1. Which of the following byte sizes is referred to as big data?
   1. Peta
   2. Giga
   3. Tera
   4. Meta
2. What are the various categories that big data can be divided into?
   1. 2 (Two)
   2. 5 (Five)
   3. 3 (Three)
   4. 4 (Four)
3. From the options listed below, identify the quintessential attributes that define big data.
   1. Value
   2. Variety
   3. Volume
   4. All the above
4. Identify the incorrect assertion from the following statements:
   1. Non-Relational databases necessitate the definition of schemas prior to data insertion
   2. NoSQL databases are designed to permit data insertion without a predefined schema.
   3. NewSQL databases are engineered to facilitate data insertion without a predefined schema.
   4. All the options.
5. Volume constitutes a fundamental attribute of big data. To what does the term “Volume” specifically pertain?
   1. The hard disk or server capacity it can hold.
   2. The data that can be processed
   3. The number of YouTube/Netflix videos that existed.
   4. The amount of data in variety of formats
6. Data can be defined as a compilation of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   1. Extensive textual content within files
   2. Diverse files organized in folders
   3. Minute pieces of information
   4. None
7. In the context of big data, how many fundamental 'V’s are recognized as its defining characteristics?
   1. 6 (Six)
   2. 2 (Two)
   3. 5 (Five)
   4. 4 (Four)
8. A database is a systematically arranged collection of data, designed to facilitate efficient management and accessibility.
   1. True
   2. False
9. What does the acronym DBMS represent?
   1. Database Management System
   2. Databases Management Software
   3. Data Manipulation Software
   4. None of the above
10. In the context of a Relational Database Management System (RDBMS), are the terms Columns, Attributes, and Fields synonymous?
    1. True
    2. False
11. Which data manipulation command is employed to joins the records and columns from one or more tables?
    1. SELECT
    2. PROJECT
    3. PRODUCT
    4. JOIN
12. Having clause can be use without GROUP BY in SELECT clause of aggregation.
    1. True
    2. False
13. Identify the statements used to display table structures in MySQL
    1. DESCRIBE
    2. STR
    3. DISPLAY
    4. None
14. In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as\_\_\_\_\_\_\_\_\_\_
    1. Adjunct function
    2. Set operation
    3. Scaler operation
    4. Aggregate function
15. Which join is utilized to retrieve the intersecting records from both the Employee table and the Employee-Employment-History table?
    1. LEFT JOIN
    2. INNER JOIN
    3. CROSS JOIN
    4. None
16. The HAVING clause is analogous to the WHERE clause, but it is exclusively utilized in conjunction with the GROUP BY clause.
    1. False
    2. True
17. \_\_\_\_\_\_\_\_\_ statement is employed to delete all the records from a table.
    1. DELETES
    2. TRUNCATE
    3. REMOVE
    4. DROP
18. Which of the following statement is a valid expression in SQL?
    1. SELECT NULL FROM school.
    2. SELECT \* FROM school;
    3. SELECT , FROM school WHEN City = NULL;
    4. SELECT # FROM school;
19. DELETE statement without where clause delate all the records from Table.
    1. TRUE
    2. FALSE
20. \_\_\_\_\_\_\_ used to get unique records from the Table.
    1. UNIQUE
    2. DISTINCT
    3. UNION
    4. None
21. R Studio essentially comprises\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ window sessions.
    1. 5(Five)
    2. 4(Four)
    3. 9(Nine)
    4. 6(Six)
22. Which function is used to generate a sequence of numbers with designated incremental or decremental values?
    1. seq(from=, to=, by=)
    2. seq(from=, to=, length.out=)
    3. seq(from=, to=, increment\_by=)
    4. None
23. Which function is used to obtain the absolute values of vectors?
    1. print()
    2. abs()
    3. prod()
    4. sum()
24. \_\_\_\_\_\_\_\_ operator is utilized to access a variable or column within a Data Frame?
    1. $(dollor)
    2. #(hash)
    3. .(dot)
    4. None
25. \_\_\_\_\_\_\_\_\_ function is utilized to transpose a Matrix or Data Frame in R?
    1. t()
    2. T()
    3. c()
    4. None
26. Identify a function to display the structure of DataFrame/List?
    1. str()
    2. summary()
    3. describe()
    4. None
27. \_\_\_\_\_\_\_\_\_\_\_\_\_ function to get quick summary of Data Frame?
    1. describe()
    2. summary()
    3. table()
    4. None
28. tail() functions to display records/observation from Top of the DataFrame?
    1. True
    2. False
29. which command is correct to create numeric (111, 222, 333, 444, 555, 666) vectors in R
30. c(‘111’,’222’,’333’,’444’,’555’,’666’)
31. c(111, 222, 333, 444, 555, 666)
32. c(111L, 222L, 333L, 444L, 555L, 666L)
33. None
34. Select a right function that used to read comma delimited values in R?
    1. read.csv()
    2. read\_csv()
    3. read.xls()
    4. None
35. Python is a versatile, interpreted, interactive, object-oriented, and high-level programming language
    1. True
    2. False
36. Identify the permissible quotation marks that Python sanctions for utilization within correct Python scripting. (Multiple selections allowed)
    1. Single
    2. Double
    3. Triple
    4. Four
37. Which of the following is utilized to delineate a block of code in the Python programming language?
    1. Indentation
    2. Key
    3. Brackets
    4. All of the mentioned
38. To define comments in Python which operator/symbol used
    1. //
    2. #
    3. /\*….\*/
    4. None
39. To import a comma-separated or delimited file in Pandas, which function is utilized to read the file and store it as a DataFrame?
    1. read\_csv
    2. read.csv
    3. read\_file
    4. None
40. Python reserved keywords are permissible for use as variable or function names.
    1. False
    2. True
41. A variable name is restricted to include only alphanumeric characters and underscores (A-Z, 0-9, and \_).
    1. True
    2. False
42. \_\_\_\_\_\_\_\_\_\_ is the package manager in Python utilized for downloading and installing Python packages.
    1. download
    2. pip
    3. install
    4. None
43. Identify the function utilized to display a summary of a Pandas Data Frame.
    1. summary
    2. describe
    3. head
    4. None
44. \_\_\_\_\_\_\_\_\_\_\_\_\_ function is utilized to obtain the frequency distribution of a categorical or labeled variable.
    1. count
    2. freq
    3. value\_counts
    4. None
45. To ascertain the number of rows and columns in a Pandas Data Frame, which method is the most optimal? (assume data is panda DataFrame)
    1. data.shape
    2. print(data)
    3. data
    4. None
46. In Python, Dictionaries are immutable that never change after it get created.
    1. False
    2. True
47. What is the output of the following code snippet?

temp = [10, 102, 30333]

temp = tuple(temp)

temp[0] = 999

print(temp)

* 1. [10, 102, 30333]
  2. (10,102,30333)
  3. (10,102,30333)
  4. Error

1. \_\_\_\_\_\_\_libraries to generate graphs on pandas DataFrame.
   1. Plotys
   2. Matplotlib
   3. Seapy
   4. None
2. What arithmetic operators cannot be used with Python strings?
   1. +
   2. –
   3. \*
   4. All of them mentioned
3. Strings in Python are immutable, signifying that a string cannot be altered once it has been created.
   1. True
   2. False
4. A \_\_\_\_\_\_\_\_\_\_\_ is a series of observations recorded over a specified duration.
   1. Theano
   2. Scikit
   3. Time Series
   4. Tensor
5. The process of breaking out long-form text into sentences and words called.
   1. Stem
   2. Cluster
   3. Bag
   4. Tokens
6. Typical text mining tasks include?
   1. Text Categorization
   2. Text Clustering
   3. Entity Relation Modelling
   4. All of the above
7. Which command is executed to ascertain the length of the string variable str?
   1. str.len()
   2. len(str)
   3. size(str)
   4. str.size()