1. Big data is also changing at a faster rate,
   1. **False**
   2. True
2. Text Mining/Analysis can be used in:
   1. Detecting spam model
   2. Predicting stock Movements
   3. News stories categorization
   4. **All the above**
3. Semi-structured data, also known as having a schema-less or self-describing structure
   1. False
   2. **True**
4. Select wrong Elements of Big Data
   1. Volume
   2. Velocity
   3. Variety
   4. **VolPro**
5. \_\_\_\_\_\_\_\_\_\_\_ is the process of transforming unstructured text into a structured format to identify meaningful patterns and new insights.
   1. Data mining
   2. **Text Mining**
   3. File Mining
   4. Deep Mining
6. In which database, data is a blend between structured and unstructured data formats?
   1. Full-structured data
   2. Partial-structured data
   3. **Semi-structured data**
   4. Uni-structured data
7. The process of breaking out long-form text into sentences and words called?
   1. Stem
   2. Cluster
   3. Bag
   4. **Tokens**
8. Text mining is being used by large media companies, to clarify information and to provide readers with greater search experiences.
   1. **TRUE**
   2. FALSE
   3. Can be true or false
   4. Can not say
9. Typical text mining tasks include?
   1. text categorization
   2. text clustering
   3. entity relation modeling
   4. **All of the above**
10. Stemming: This refers to the process of separating the prefixes and suffixes from words to derive the root word form and meaning.
    1. **TRUE**
    2. FALSE
    3. Can be true or false
    4. Can not say
11. Most of the data in the world is in fact unstructured
    1. False
    2. **True**
12. The structured and unstructured data that is collected and analyzed in social media analytics programs includes all of the following, EXCEPT:
    1. Comments/reviews
    2. Demographic information
    3. **Transaction data**
    4. Number of likes and followers
13. Which of the following techniques can be used for the purpose of keyword normalization, the process of converting a keyword into its base form(Multiple Choice)?
    1. **Lemmatization**
    2. Levenshtein
    3. **Stemming**
    4. Soundex
14. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Python libraries used to perform text analysis
    1. **Spacy**
    2. Pandas
    3. Numpy
    4. None
15. Tokenization refers to segmenting text into words, punctuations marks, numbers..etc
    1. **True**
    2. False
16. Select the right statement to install spacy library
    1. pip download spacy
    2. **pip install spacy**
    3. download spacy
    4. None
17. Big data refers to collection small and strictly structure type data.
    1. True
    2. **False**
18. Select the sources of data in Big Data eco systems (Multiple Choice).
    1. In/Out
    2. **Internal**
    3. **External**
    4. None
19. Structured data can be defined as the data that has a defined repeating pattern.
    1. **True**
    2. False
20. Is a new data challenge that requires leveraging existing systems differently, then we classify data as Big Data.
    1. **True**
    2. False
21. Big Data usually unstructured and ~70-80% data available in unstructured format.
    1. False
    2. **True**
22. Velocity is the speed in which data is process and becomes accessible.
    1. False
    2. **True**
23. Variability refers to the data which keeps on changing constantly.
    1. **True**
    2. False
24. Variety describes one of the biggest challenges of \_\_\_\_\_\_.
    1. **Big data**
    2. Data science
    3. Data integration
    4. None of the mentioned above
25. The data that can be processed, stored, and retrieved in a fixed format called \_\_\_\_\_,
    1. **Structured Data**
    2. Unstructured Data
    3. Semi-structured Data
    4. None of the mentioned above
26. Semi-structured data, also known as having a schema-less or self-describing structure
    1. False
    2. **True**
27. Select correct Elements of Big Data
    1. Volume
    2. Velocity
    3. Variety
    4. **All**
28. Text mining is process of transforming free of textual data into a structure format.
    1. False
    2. **True**
29. In which database, data is a blend between structured and unstructured data formats?
    1. Full-structured data
    2. Partial-structured data
    3. **Semi-structured data**
    4. Uni-structured data
30. The process of breaking out long-form text into sentences and words called?
    1. Stem
    2. Cluster
    3. Bag
    4. **Tokens**
31. Text mining is being used by large media companies, to clarify information and to provide readers with greater search experiences.
    1. **TRUE**
    2. FALSE
    3. Can be true or false
    4. Can not say
32. Typical text mining tasks include?
    1. text categorization
    2. text clustering
    3. entity relation modeling
    4. **All of the above**
33. Stemming: This refers to the process of separating the prefixes and suffixes from words to derive the root word form and meaning.
    1. **TRUE**
    2. FALSE
    3. Can be true or false
    4. Can not say
34. Most of the data in the world is in fact unstructured
    1. False
    2. **True**
35. The structured and unstructured data that is collected and analyzed in social media analytics programs includes all of the following, EXCEPT:
    1. Comments/reviews
    2. Demographic information
    3. **Transaction data**
    4. Number of likes and followers
36. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Python libraries used to perform text analysis
    1. **Spacy**
    2. Pandas
    3. Numpy
    4. None
37. Tokenization refers to segmenting text into words, punctuations marks, numbers..etc
    1. **True**
    2. False
38. Select the right statement to install spacy library
    1. pip download spacy
    2. **pip install spacy**
    3. download spacy
    4. None
39. Big data refers to collection large data and strictly structure type data.
    1. True
    2. **False**
40. Select the sources of data in Big Data eco systems (Multiple Choice).
    1. In/Out
    2. **Internal**
    3. **External**
    4. None
41. Structured data can be defined as the data that has a defined repeating pattern.
    1. **True**
    2. False
42. Is a new data challenge that requires leveraging existing systems differently, then we classify data as Big Data.
    1. **True**
    2. False
43. Big Data usually in structured format and ~70-80% data available in structured format.
    1. **False**
    2. True
44. SQL Stands for
    1. Structured Language
    2. Simple Query Language
    3. **Structured Query Language**
    4. None
45. Which data manipulation command is used to combines the records from one or more tables?
    1. SELECT
    2. PROJECT
    3. PRODUCT
    4. **JOIN**
46. The command used to delete a particular record in a Table
    1. UPDATE TABLE
    2. **DELETE FROM**
    3. TRUNCATE COLUMN
    4. DROP
47. Which of the following is a legal expression in SQL?
    1. SELECT NULL FROM SALES;
    2. **SELECT \* FROM SALES;**
    3. SELECT \* FROM SALES WHEN PRICE = NULL;
    4. SELECT # FROM SALES;
48. The virtual table that its created by data from the result of an SQL 'Select' statement is called
    1. **View**
    2. Synonym
    3. Sequence
    4. Transaction
49. Which statement used to update/change the value of any columns
    1. MODIFY statement
    2. **UPDATE statement**
    3. ALTER statement
    4. None
50. Which of the following commands can be used to delete a tables/relation if it is not needed anymore?
    1. UPDATE
    2. ALTER
    3. **DROP**
    4. DELETE
51. Having clause can be use without GROUP BY
    1. TRUE
    2. **FALSE**
52. Which of the following represents basic SQL DML statements
    1. Drop, Update, Delete
    2. Insert, Alter, Delete
    3. Insert, View, Create
    4. **Insert, Update, Delete**
53. WHERE & HAVING clause, both used to filter out the data
    1. **TRUE**
    2. FALSE
54. Numbers, Strings and DateTime are most common data types in SQL
    1. **TRUE**
    2. FALSE
55. In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as\_\_\_\_\_\_\_\_\_\_
    1. adjunct function
    2. set operation
    3. scaler operation
    4. **aggregate function**
56. Statement used to get unique records from the Table
    1. UNIQUE
    2. **DISTINCT**
    3. UNION
    4. None
57. Tuples & Rows are same in Table
    1. **TRUE**
    2. FALSE
58. Statement used to display table structures in MySQL
    1. **DESCRIBE**
    2. STR
    3. DISPLAY
    4. None
59. DELETE statement without where clause delate all the records from Table
    1. **TRUE**
    2. FALSE
60. Which join used get the common records from both the tables (table 1 & table 2).
    1. LEFT JOIN
    2. **INNER JOIN**
    3. CROSS JOIN
    4. None
61. \_\_\_\_\_\_\_\_\_\_\_ operator is used to combine the result-set of two or more SELECT statements.
    1. **UNION**
    2. JOIN
    3. MERGE
    4. None
62. HAVING clause is equivalent to WHERE clause but HAVING used with only GROUP BY clause.
    1. FALSE
    2. **TRUE**
63. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ operators used to specify multiple possible values for a column while defining condition with WHERE clause.
    1. LIKE
    2. **IN**
    3. AND
    4. OR
64. Which statement used to delete all the records from table
    1. DELETES
    2. **TRUNCATE**
    3. REMOVE
    4. DROP
65. UPDATE statement without WHERE clause update all records for selected columns.
    1. FALSE
    2. **TRUE**
66. SQL Stands for
    1. Structured Language
    2. Simple Query Language
    3. **Structured Query Language**
    4. None
67. Asterisk (\*) in SELECT statement indicates that Query will bring only few columns from the Table
    1. **FALSE**
    2. TRUE
68. The command used to delete a particular record in a Table
    1. UPDATE TABLE
    2. **DELETE FROM**
    3. TRUNCATE COLUMN
    4. DROP
69. Which of the following is a legal expression in SQL?
    1. SELECT NULL FROM students;
    2. **SELECT \* FROM students;**
    3. SELECT \* FROM students WHEN marks = NULL;
    4. SELECT # FROM students;
70. The virtual table that its created by data from the result of an SQL 'Select' statement is called
    1. **View**
    2. Synonym
    3. Sequence
    4. Transaction
71. Which statement used to INSERT the value for Table
    1. **INSERT statement**
    2. TRIGGER statement
    3. ALTER statement
    4. None
72. Which of the following commands can be used to delete a tables/relation if it is not needed anymore?
    1. UPDATE
    2. ALTER
    3. **DROP**
    4. DELETE
73. Numeric data can be inserted without Single/Double quotes with INSERT statement.
    1. **TRUE**
    2. FALSE
74. Which of the following represents basic SQL statements?
    1. Drop, Update, Delete
    2. Insert, Alter, Delete
    3. Insert, View, Create
    4. **Select, Insert, Update, Delete**
75. WHERE & HAVING clause, both used to filter out the data
    1. **TRUE**
    2. FALSE
76. Numbers, Strings and DateTime are most common data types in SQL
    1. **TRUE**
    2. FALSE
77. In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as\_\_\_\_\_\_\_\_\_\_
    1. adjunct function
    2. set operation
    3. scaler operation
    4. **aggregate function**
78. Statement used to get unique records from the Table
    1. UNIQUE
    2. **DISTINCT**
    3. UNION
    4. None
79. Columns & Attributes are same in Table
    1. **TRUE**
    2. FALSE
80. Statement used to display table structures in MySQL
    1. **DESCRIBE**
    2. STR
    3. DISPLAY
    4. None
81. DELETE statement without where clause delate all the records from Table
    1. **TRUE**
    2. FALSE
82. Which join used get all the records from right table and matching records from the left table (table 1 & table 2).
    1. LEFT JOIN
    2. **RIGHT JOIN**
    3. CROSS JOIN
    4. None
83. \_\_\_\_\_\_\_\_\_\_\_ operator is used to combine the result-set of two or more SELECT statements and duplicates allowed.
    1. **UNION ALL**
    2. JOIN
    3. MERGE
    4. None
84. HAVING clause is equivalent to WHERE clause but HAVING used with only GROUP BY clause.
    1. FALSE
    2. **TRUE**
85. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ operators used to specify multiple possible values for a column while defining condition with WHERE clause.
    1. LIKE
    2. **IN**
    3. AND
    4. OR
86. Which statement used to delete all the records from table
    1. DELETES
    2. **TRUNCATE**
    3. REMOVE
    4. DROP
87. UPDATE statement without WHERE clause update all records for selected columns.
    1. FALSE
    2. **TRUE**
88. Function used to load the package in memory/project of R is
    1. import()
    2. load.packages()
    3. **library()**
    4. None
89. install() is used to install the package in the memory/R environment?
    1. True
    2. **False**
90. R Studio has basically\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ window sessions.
    1. 2/Two
    2. **4/Four**
    3. 3/Three
    4. 1/One
91. Function used to get the project working directory/folder?
92. cd()
93. pwd()
94. show()
95. **getwd()**
96. function used to create vectors in R?
97. t()
98. **c()**
99. class()
100. None
101. Which command is correct to create numeric (1,2,3,4,5,6) vectors in R
102. c(‘1’,’2’,’3’,’4’,’5’,’6’)
103. **c(1,2,3,4,5,6)**
104. c(1L,2L,3L,4L,5L,6L)
105. None
106. Function used show the class of basic data types?
     1. print()
     2. **class()**
     3. show()
     4. c()
107. Operator used to create range of values i.e. 1:10?
     1. **Colon (:) operator**
     2. Special operator
     3. Semicolon(;) operator
     4. None
108. Function used to create sequence of number with specified increment/decrement values?
     1. **seq(from=, to=, by=)**
     2. seq(from=, to=, length.out=)
     3. seq(from=, to=, increment\_by=)
     4. None
109. Special variable/function **letters** used to get the small letter alphabets vectors directly?
     1. **True**
     2. False
110. Let’s assume num <- c(10,20,30,40,50) is numeric vector, need to access the 2nd element from vector which statement is correct?
     1. **num[2]**
     2. num[1:2]
     3. num[]
     4. None
111. Function used to sort the vector elements?
     1. c()
     2. **sort()**
     3. print()
     4. None

1. Function used to get absolute values of the vectors?
   1. print()
   2. **abs()**
   3. prod()
   4. sum()
2. Operators to check the elementwise conditions in R?
   1. Condition
   2. **Relational operators**
   3. Filter
   4. None
3. Matrix data structures has only one indexes?
   * 1. True
     2. **False**
4. Function used to get the dimension of Matrix/DataFrame?
   * 1. nrow()
     2. ncol()
     3. **dim()**
     4. None
5. Operator used to access the variable/column of DataFrame?
   * 1. **$(dollor)**
     2. #(hash)
     3. .(dot)
     4. None
6. Matrix element are arranged on column basis by specifying byrow=TRUE?
   * 1. True
     2. **False**
7. Combining two same matrixing by row using rbind() function in R?
   * 1. **True**
     2. False
8. List is collections of all types of data & data structures in R?
   * 1. **True**
     2. False
9. True Matrix multiplication perform with special operator
   * 1. **%\*%**
     2. %/%
     3. %in%
     4. None
10. Function used to transpose of the Matrix/DataFrame in R?
    * 1. **t()**
      2. T()
      3. c()
      4. None
11. Function used to read comma separated values in R?
    * 1. **read.csv()**
      2. read\_csv()
      3. read.xls()
      4. None
12. Function to display the structure of DataFrame/List?
    * 1. **str()**
      2. summary()
      3. describe()
      4. None
13. Function to get quick summary of DataFrame?
    * 1. describe()
      2. **summary()**
      3. table()
      4. None
14. Function to split the Dataframe based on specified column name?
    * 1. **split()**
      2. subset()
      3. get()
      4. None
15. tail() functions to display records/observation from Top of the dataframe?
    * 1. True
      2. **False**
16. The Following function is used to print first 6 records of the dataset
    * 1. **head**
      2. tail
      3. start
      4. initial
17. Function used to display the value of any variables (Multiple Selection)?
    * 1. **print()**
      2. **View()**
      3. display()
      4. get()
18. Python is a generate purpose interpreted, and high-level programming language.
    1. **True**
    2. False
19. Python is statement looks like English, this indicates python code is more readable and looks like English statement.
    1. False
    2. **True**
20. Python is only available for Windows PC/Laptop operating system.
    1. **False**
    2. True
21. Identify the software name which used to write python code and projects.
    1. RStudio
    2. **Jupyter Notebook**
    3. MS Word
    4. None
22. Python reserved words that cannot be used for variable name/ function name and reserved words are in lowercase always.
    1. False
    2. **True**
23. Identify invalid variable names
    1. **2myvar = “John”**
    2. Myvar = “John”
    3. MYVAR = “John”
    4. None
24. \_\_\_\_\_\_\_\_\_\_\_\_ function used to print the value of variables in Python
    1. **Print**
    2. Display
    3. View
    4. None
25. Python used \_\_\_\_\_\_\_\_\_\_\_\_\_\_ approach to indicates the block of codes.
    1. **Indentation**
    2. Brackets
    3. Block
    4. None
26. Select the quotations support by Python
    1. Single
    2. Double
    3. Triple
    4. **All of the above**
27. \_\_\_\_\_\_\_\_\_\_\_\_\_\_ symbol represents comment line in Python
    1. +
    2. /\* \*/
    3. **#**
    4. None
28. To download pandas library, select the right statements
    1. **pip install pandas**
    2. pip download pandas
    3. pip update pandas
    4. None
29. \_\_\_\_\_\_\_\_\_\_\_ functions to check type of data variable is holding
    1. **type()**
    2. check()
    3. print()
    4. None
30. \_\_\_\_\_\_\_\_\_\_\_\_\_ (if..else) statement used for checking condition with Python.
    1. **Conditional statement**
    2. Control Statement
    3. Loop statement
    4. None
31. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ keyword used to define functions in Python.
    1. If
    2. **def**
    3. fun
    4. None
32. Select common data structures available in Python (Multiple selection)
    1. Matrix
    2. **List**
    3. **Tuple**
    4. **Sets**
33. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structure is immutable.
    1. **Tuple**
    2. List
    3. Dictionary
    4. None
34. Set data structures always contains unique elements.
    1. False
    2. **True**

Plus(+) operators used to combine elements of \_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structures.

* 1. Set
  2. Dictionary
  3. **List**
  4. None

1. Len function is common function to find the total length of elements in data structures.
   1. False
   2. **True**
2. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structures contain element in form of keys: values.
   1. List
   2. **Dictionary**
   3. DataFrame
   4. None
3. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ approach to access elements from List & Tuple data structures.
   1. **Indexing**
   2. Columns
   3. Name
   4. None
4. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ function used to combine two set.
   1. Append
   2. Combine
   3. **Union**
   4. None
5. Identify the software name which used to write python code and projects.
   1. RStudio
   2. **Jupyter Notebook**
   3. MS Word
   4. None
6. Python reserved words that cannot be used for variable name/ function name and reserved words are in lowercase always.
   1. False
   2. **True**
7. Identify invalid variable names
   1. **My-var = “John”**
   2. Myvar = “John”
   3. MYVAR = “John”
   4. None
8. Python used \_\_\_\_\_\_\_\_\_\_\_\_\_\_ approach to indicates the block of codes.
   1. **Indentation**
   2. Brackets
   3. Block
   4. None
9. Select the quotations support by Python
   1. Single
   2. Double
   3. Triple
   4. **All of the above**
10. \_\_\_\_\_\_\_\_\_\_\_\_\_\_ symbol represents comment line in Python
    1. +
    2. /\* \*/
    3. **#**
    4. None
11. To download pandas library, select the right statements
    1. **pip install sklearn**
    2. pip download sklearn
    3. pip update sklearn
    4. None
12. \_\_\_\_\_\_\_\_\_\_\_ functions to check type of data variable is holding
    1. **type()**
    2. check()
    3. print()
    4. None
13. \_\_\_\_\_\_\_\_\_\_\_\_\_\_ statement used for writing conditional statement with Python.
    1. **If..else**
    2. For loop
    3. def
    4. None
14. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ keyword used to define functions in Python.
    1. If
    2. **def**
    3. fun
    4. None
15. Select common data structures available in Python (Multiple selection)
    1. Matrix
    2. **List**
    3. **Tuple**
    4. **Sets**
16. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structure is mutable(Multiple selection).
    1. Tuple
    2. **List**
    3. **Dictionary**
    4. **Set**
17. Set data structures always contains unique elements.
    1. False
    2. **True**
18. Plus(+) operators used to combine elements of \_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structures.
    1. Set
    2. Dictionary
    3. **List**
    4. None
19. Len function is common function to find the total length of elements in data structures.
    1. False
    2. **True**
20. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ data structures contain element in form of keys: values.
    1. List
    2. **Dictionary**
    3. DataFrame
    4. None
21. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ approach to access elements from List & Tuple data structures.
    1. **Indexing**
    2. Columns
    3. Name
    4. None
22. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ function used to add or update values in Dictionary.
    1. **update**
    2. Combine
    3. Union
    4. None
23. \_\_\_\_\_\_\_\_\_\_\_\_\_ brackets used to create set data structures.
    1. **{ }**
    2. [ ]
    3. ()
    4. None
24. Data Structures/Collections are useful containers to store and manipulate list of homogeneous or heterogeneous elements
    1. False
    2. **True**
25. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ statement to add library/package in python code.
    1. Load
    2. **import**
    3. Add
    4. None
26. Variable name can only contain alpha-numeric character and underscores(A-z, 0-9 and \_ )
    1. False
    2. **True**

=========================== Set-2 ==============================

1. \_\_\_\_\_\_\_\_\_\_\_ is a collection of data that is used in volume, yet growing exponentially with time
   1. Big databases
   2. Big DBMS
   3. Big Data Files
   4. **Big Data**
2. Which of the following are the Benefits of Big Data Processing?
   1. Businesses can utilize outside intelligence while taking decision
   2. Better operational efficiency
   3. Improve customer server
   4. **All of the above**
3. Big data analysis does the following except?
   1. Spread data
   2. **Analyze data**
   3. Organizes data
   4. Collect data
4. Which of the following is true about big data?
   1. Big data can be processed using traditional techniques
   2. **Big data refers to data sets that are at least a petabyte in size**
   3. Big data analysis does not involve reporting and data mining techniques
   4. Big data has low velocity meaning that it is generated slowly
5. How many V’s of Big Data
   1. 2
   2. 3
   3. 5
   4. **4**
6. Choose the primary characteristics of big data among the following
   1. Value
   2. Variety
   3. Volume
   4. **All of the above**
7. Database Management System is software used to store and retrieve the databases.
   1. **True**
   2. False
8. Which is DBMS software used to manipulated data & databases.
   1. **MySQL**
   2. SQL
   3. Data Studio
   4. None
9. Identify correct Full form of RDBMS
   1. Row Database Management System
   2. **Relational Database Management System**
   3. Right Database Management System
   4. None
10. Records are also known as Tuples in RDBMS
    1. **True**
    2. False
11. Which join used get all the records from right table and matching records from the left table (table 1 & table 2).
    1. LEFT JOIN
    2. **RIGHT JOIN**
    3. CROSS JOIN
    4. None
12. \_\_\_\_\_\_\_\_\_\_\_ operator is used to combine the result-set of two or more SELECT statements and duplicates allowed.
    1. **UNION ALL**
    2. JOIN
    3. MERGE
    4. None
13. Which statement used to insert the value for a Table
    1. **INSERT statement**
    2. TRIGGER statement
    3. ALTER statement
    4. None
14. Which join used get the common records from both the tables (table 1 & table 2).
    1. LEFT JOIN
    2. **INNER JOIN**
    3. CROSS JOIN
    4. None
15. Numeric data can be inserted without Single/Double quotes with INSERT statement.
    1. **TRUE**
    2. FALSE
16. Which statement used to delete all the records from table
    1. DELETES
    2. **TRUNCATE**
    3. REMOVE
    4. DROP
17. Which of the following represents four basic SQL statements?
    1. Drop, Update, Delete
    2. Insert, Alter, Delete
    3. Insert, View, Create
    4. **Select, Insert, Update, Delete**
18. DELETE statement without where clause delate all the records from Table
    1. **TRUE**
    2. FALSE
19. \_\_\_\_\_\_\_\_\_\_\_ operator is used to combine the result-set of two or more SELECT statements and duplicates allowed.
    1. **UNION ALL**
    2. JOIN
    3. MERGE
    4. None
20. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ operators used to specify multiple possible values for a column while defining condition with WHERE clause.
    1. LIKE
    2. **IN**
    3. AND
    4. OR
21. R Studio has basically\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ window sessions.
    1. 2/Two
    2. **4/Four**
    3. 3/Three
    4. 1/One
22. Function used show the value of variable.
    1. **print()**
    2. class()
    3. show()
    4. c()
23. Operator used to create range of values i.e. 11:20?
    1. **Colon (:) operator**
    2. Special operator
    3. Semicolon(;) operator
    4. None
24. True Matrix multiplication perform with special operator
    1. **%\*%**
    2. %/%
    3. %in%
    4. None
25. Function used to display the data in Excel like format
    1. print()
    2. display()
    3. **View()**
    4. None
26. Operator used to access the variable/column of DataFrame?
    1. **$(dollor)**
    2. #(hash)
    3. .(dot)
    4. None
27. Function used to transpose of the Matrix/DataFrame in R?
    1. **t()**
    2. T()
    3. c()
    4. None
28. Functions used to count the frequency of Factor/Character data?
    1. **table()**
    2. freq()
    3. count()
    4. Nonw
29. Function to get quick summary of DataFrame?
    1. describe()
    2. **summary()**
    3. table()
    4. None
30. Function to subset the data from Dataframe based on conditions?
    1. split()
    2. **subset()**
    3. get()
    4. None
31. which command is correct to create Integer (10,20,30,40,50,60) vectors in R
32. c(‘10’,’20’,’30’,’40’,’50’,’60’)
33. c(10,20,30,40,50,60)
34. **c(10L,20L,30L,40L,50L,60L)**
35. None
36. Function used to read comma separated values in R?
    1. **read.csv()**
    2. read\_csv()
    3. read.xls()
    4. None
37. To define comments in Python which operator/symbol used
    1. //
    2. **#**
    3. /\*….\*/
    4. None
38. Select the Quotation which Python supports (Multiple selection)
    1. '
    2. “
    3. **“ ” ” or ‘ ‘ ‘**
    4. “ ” ” ”
39. Python used Indentation to define a code block, like other programming used {} (brackets)
    1. **True**
    2. False
40. Read Exccel file in Pandas, which functions used to read file and store as DataFrame
    1. **read\_excel**
    2. read.excel
    3. read\_file
    4. None
41. Python reserved words are basic building blocks of the Python programming language
    1. False
    2. **True**
42. A variable name can only contain alpha-numeric characters and underscores (A-z, 0-9, and \_ )
    1. **True**
    2. False
43. Identify the valid statement for installing pandas packages.
    1. **pip install pandas**
    2. python install pandas
    3. pip setup pandas
    4. None
44. Function which used to display summary of Pandas DataFrame
    1. summary
    2. **describe**
    3. head
    4. None
45. value\_counts() function gives the frequency count of any selected columns from Pandas DataFrame.
    1. **True**
    2. False
46. In Python, Dictionaries are immutable
    1. **False**
    2. True
47. Set data structure allow to store unique values in Python.
    1. False
    2. **True**
48. Most common used graphics library to display graphs with Pandas DataFrame
    1. Seaborn
    2. Ploty
    3. **Matplotlib**
    4. None
49. What arithmetic operators cannot be used with Python strings?
    1. +
    2. **–**
    3. \*
    4. All of them mentioned
50. Strings are immutable in Python, which means a string cannot be modified.
    1. **True**
    2. False
51. The process of breaking out long-form text into sentences and words called?
    1. Stem
    2. Cluster
    3. Bag
    4. Bag
    5. **Tokens**
52. Identify the most common libraries used for performing text mining/analysis in Python? (Multiple choice)
    1. NLTK
    2. Spacy
    3. Sklearn
    4. **All of the above**
53. To return the length of string s what command do we execute? (assume str is string variable)
    1. **str.len()**
    2. len(str)
    3. size(str)
    4. str.size()
54. Write the output of the following code:

**>> L = [1,2,3,4,5,[6,7,8]]**

**>> print(L[5])**

* 1. **[6,7,8]**
  2. Error
  3. 6,7,8
  4. 5

1. How many V’s of Big Data
   1. 2
   2. 3
   3. 5
   4. **4**
2. Data in \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ bytes size called big data
   1. **Peta**
   2. Giga
   3. Tera
   4. Meta
3. Total forms of big data is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   1. 1
   2. 2
   3. **3**
   4. 4
4. Choose the primary characteristics of big data among the following
   1. Value
   2. Variety
   3. Volume
   4. **All of the above**
5. Point out the wrong statement:
   1. **Non-Relational databases require that schemas be defined before you can add data**
   2. NoSQL databases are built to allow the insertion of data without a predefined schema.
   3. NewSQL database are built to allow the insertion of data without a predefined schema.
   4. All of the options.
6. Volume is one of the characteristics of big data. What does Volume refer to?
   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**
7. Data is a collection of a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   1. Large Textual contents of files
   2. Various files in Folders
   3. **Small piece of information**
   4. None
8. Database is an organized collection of data so that it can be easily manage and accessible.
   1. **True**
   2. False
9. DBMS stands for
   1. **Database Management System**
   2. Databases Management Software
   3. Data Manipulation Software
   4. None
10. Columns/Attributes/Field all are same in RDBMS
    1. **True**
    2. False
11. Which data manipulation command is used to combines the records from one or more tables?
    1. SELECT
    2. PROJECT
    3. PRODUCT
    4. **JOIN**
12. Having clause can be use without GROUP BY
    1. True
    2. **False**
13. In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as\_\_\_\_\_\_\_\_\_\_
    1. Adjunct function
    2. Set operation
    3. Scaler operation
    4. **Aggregate function**
14. Statement used to display table structures in MySQL
    1. **DESCRIBE**
    2. STR
    3. DISPLAY
    4. None
15. Which join used get the common records from both the tables (table 1 & table 2).
    1. LEFT JOIN
    2. **INNER JOIN**
    3. CROSS JOIN
    4. None
16. HAVING clause is equivalent to WHERE clause but HAVING used with only GROUP BY clause.
    1. FALSE
    2. **TRUE**
17. Which statement used to delete all the records from table
    1. DELETES
    2. **TRUNCATE**
    3. REMOVE
    4. DROP
18. Which of the following is a legal expression in SQL?
    1. SELECT NULL FROM Employee;
    2. **SELECT \* FROM Employee;**
    3. SELECT , FROM Employee WHEN City = NULL;
    4. SELECT # FROM Employee;
19. DELETE statement without where clause delate all the records from Table
    1. **TRUE**
    2. FALSE
20. Statement used to get unique records from the Table
    1. UNIQUE
    2. **DISTINCT**
    3. UNION
    4. None
21. R Studio has basically\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ window sessions.
    1. 2/Two
    2. **4/Four**
    3. 3/Three
    4. 1/One
22. Function used to create sequence of number with specified increment/decrement values?
    1. **seq(from=, to=, by=)**
    2. seq(from=, to=, length.out=)
    3. seq(from=, to=, increment\_by=)
    4. None
23. Function used to get absolute values of the vectors?
    1. print()
    2. **abs()**
    3. prod()
    4. sum()
24. Operator used to access the variable/column of DataFrame?
    1. **$(dollor)**
    2. #(hash)
    3. .(dot)
    4. None
25. Function used to transpose of the Matrix/DataFrame in R?
    1. **t()**
    2. T()
    3. c()
    4. None
26. Function to display the structure of DataFrame/List?
    1. **str()**
    2. summary()
    3. describe()
    4. None
27. Function to get quick summary of DataFrame?
    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 (1,2,3,4,5,6) vectors in R
30. c(‘1’,’2’,’3’,’4’,’5’,’6’)
31. **c(1,2,3,4,5,6)**
32. c(1L,2L,3L,4L,5L,6L)
33. None
34. Function used to read comma separated values in R?
    1. **read.csv()**
    2. read\_csv()
    3. read.xls()
    4. None
35. Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language
    1. **True**
    2. False
36. Select the Quotation which Python supports (Multiple selection)
    1. Single
    2. Double
    3. **Triple**
    4. Four
37. Which of the following is used to define a block of code in Python 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. Read CSV file in Pandas, which functions used to read file and store as DataFrame
    1. **read\_csv**
    2. read.csv
    3. read\_file
    4. None
40. Python reserved words can not be used as Variable/Functions names
    1. False
    2. **True**
41. A variable name can only contain alpha-numeric characters and underscores (A-z, 0-9, and \_ )
    1. **True**
    2. False
42. \_\_\_\_\_\_\_\_\_\_\_\_\_ is package manager in Python to download & install Python packages.
    1. download
    2. **pip**
    3. install
    4. None
43. Function which used to display summary of Pandas DataFrame
    1. summary
    2. **describe**
    3. head
    4. None
44. \_\_\_\_\_\_\_\_\_\_\_\_\_ Function to get the frequency of categorical/labeled variable.
    1. count
    2. freq
    3. **value\_counts**
    4. None
45. To view no of rows & columns for Pandas data, which is best option to get (assume data is panda DataFrame)
    1. **data.shape**
    2. print(data)
    3. data
    4. None
46. In Python, Dictionaries are immutable
    1. **False**
    2. True
47. What will be the output of the following code snippet?

a = [1, 2, 3]

a = tuple(a)

a[0] = 2

print(a)

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

1. To visualize graphs in Pandas which is most common library used
   1. Ploty
   2. **Matplotlib**
   3. Seaborn
   4. None
2. What arithmetic operators cannot be used with Python strings?
   1. +
   2. **–**
   3. \*
   4. All of them mentioned
3. Strings are immutable in Python, which means a string cannot be modified.
   1. **True**
   2. False
4. A \_\_\_\_\_\_\_\_\_\_\_ is a sequence of observations over a certain period.
   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. Bag
   5. **Tokens**
6. Typical text mining tasks include?
   1. Text Categorization
   2. Text Clustering
   3. Entity Relation Modelling
   4. **All of the above**
7. To return the length of string s what command do we execute? (assume s is string variable)
   1. **s.len()**
   2. len(s)
   3. size(s)
   4. s.size()
8. In the early 60s, technology witnessed problem with velocity or real-time data assimilation. This inspired the evolution of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ technology.
   1. Database
   2. **Files**
   3. Big Data
   4. None
9. Is Big Data usually unstructured and qualitative in nature?
   1. **True**
   2. False
10. Data in \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ bytes size called big data
    1. Peta (1 PB = 1000TB)
    2. Giga (1 GB = 1000MB)
    3. Tera (1 TB = 1000GB)
    4. Mega Byte (1 MB = 1000KB)
11. What kind of data big data can manage \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (Multiple Choice)?
    1. **Structured**
    2. **Unstructured**
    3. **Semi-structured**
    4. Mix-data
12. Database is an organized collection of data so that it can be easily manage and accessible.
    1. **True**
    2. False
13. Choose the primary characteristics of big data among the following
    1. Value
    2. Variety
    3. Volume
    4. **All of the above**
14. Volume is one of the characteristics of big data. What does Volume refer to?
    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**
15. To define Data, which statement is right from option?
    1. Large Textual contents of files
    2. **Small piece of information**
    3. Various files in Folders
    4. None
16. RDBMS stands for
    1. **Relational Database Management System**
    2. Raw Databases Management System
    3. Relation Data Manipulation Software
    4. None
17. Columns/Attributes/Field all are same in RDBMS
    1. **True**
    2. False
18. \_\_\_\_\_\_\_\_\_\_ tools & software used to write database related programming.
    1. **MySQL**
    2. Notepad
    3. Jupyter
    4. None
19. To display data from students table, choose correct statement from below option.
    1. SELECT # FROM students;
    2. SELECT + FROM students WHERE Null;
    3. **SELECT \* FROM students;**
    4. None
20. Which data manipulation command is used to combines the records from one or more tables?
    1. SELECT
    2. PROJECT
    3. PRODUCT
    4. **JOIN**
21. In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as\_\_\_\_\_\_\_\_\_\_
    1. Adjunct function
    2. Set operation
    3. Scaler operation
    4. **Aggregate function**
22. Statement used to display table structures in MySQL
    1. **DESCRIBE**
    2. STR
    3. DISPLAY
    4. None
23. Which join used get the all the records from left table and matching records from right table.
    1. **LEFT JOIN**
    2. INNER JOIN
    3. CROSS JOIN
    4. None
24. UPDATE statement will update all the records if WHERE clause it not specified.
    1. FALSE
    2. **TRUE**
25. Which statement used to remove complete table from MySQL
    1. DELETES
    2. **TRUNCATE**
    3. REMOVE
    4. DROP
26. DELETE statement without where clause delate all the records from Table
    1. **TRUE**
    2. FALSE
27. Statement used to get unique records from the Table
    1. UNIQUE
    2. **DISTINCT**
    3. UNION
    4. None
28. Select software required to write & run R Code. (Multiple choice)
    1. **R**
    2. Notebook
    3. **R Studio**
    4. Visual Studio

1. R Studio has basically\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ window sessions.
   1. Two
   2. Five
   3. Three
   4. **Four**
2. Select most common data structures are available in R (Multiple choice)
   1. **Vector**
   2. **Matrix**
   3. **DataFrame**
   4. Integer
3. Function used to display values of the vectors or any variables in R?
   1. **print()**
   2. abs()
   3. prod()
   4. sum()
4. Operator used to access the variable/column of DataFrame?
   1. **$(dollor)**
   2. #(hash)
   3. .(dot)
   4. None
5. Function used to transpose of the Matrix/DataFrame in R?
   1. **t()**
   2. T()
   3. c()
   4. None
6. Function to display top records from R DataFrame?
   1. **head()**
   2. print()
   3. tail()
   4. None
7. haven package/library used to read SPSS/SAV file in R Code.
   1. **True**
   2. False
8. which command is correct to create character (‘A’,’B’,’C’,’D’,’E’,’F’) vectors in R
   1. **c(‘A’,’B’,’C’,’D’,’E’,’F’)**
   2. c(A,B,C,D,E,F)
   3. c(AL,BL,CL,DL,EL,FL)
   4. None
9. Function used to read comma separated values in R?
   1. read.csv()
   2. **read\_csv()**
   3. read.xls()
   4. None
10. Most common data types in Python programming language (Multiple choice)
    1. **int**
    2. **float**
    3. **bool**
    4. **str**
11. Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language
    1. **True**
    2. False
12. Software used to run Python code.
    1. **Jupyter notebook**
    2. R Studio
    3. phpMyAdmin
    4. None
13. Select the Quotation which Python supports (Multiple selection)
    1. **Single**
    2. **Double**
    3. **Triple**
    4. Four
14. Which statement python used to define sets of conditional statements
    1. **If..else**
    2. having
    3. where
    4. None
15. Which of the following is used to define a block of code in Python language?
    1. **Indentation**
    2. Key
    3. Brackets
    4. All of the mentioned
16. To define comments in Python \_\_\_\_\_\_\_\_\_\_\_\_\_ operator/symbol used
    1. //
    2. **#**
    3. /\*….\*/
    4. None
17. Read CSV file in Pandas, which functions used to read file and store as DataFrame
    1. read\_csv
    2. **read.csv**
    3. read\_file
    4. None
18. Python reserved words are python keyword which can be used as Variable/Functions names
    1. **False**
    2. True
19. A variable name can only contain alpha-numeric characters and underscores (A-z, 0-9, and \_ )
    1. **True**
    2. False
20. \_\_\_\_\_\_\_\_\_\_\_\_\_ is package manager in Python to download & install Python packages.
    1. download
    2. **pip**
    3. install
    4. None
21. To define/create Function in Python which keyword \_\_\_\_\_\_\_\_\_\_\_is used.
    1. **def**
    2. describe
    3. colon
    4. None
22. \_\_\_\_\_\_\_\_\_\_\_\_\_ Pandas Function to get the frequency of categorical/labeled variable.
    1. count
    2. freq
    3. **value\_counts**
    4. None
23. Loop is useful for iterating elements of List/Set/Tuple
    1. False
    2. **True**
24. To view no of rows & columns for Pandas data, which is best option to get (assume data is panda DataFrame)
    1. **data.shape**
    2. print(data)
    3. data
    4. None
25. In Python, Dictionaries are immutable
    1. **False**
    2. True
26. Sets always contains repeated value.
    1. False
    2. **True**
27. To visualize graphs in Pandas which is most common library used
    1. Ploty
    2. **Matplotlib**
    3. Seaborn
    4. None
28. len() functions to display the no of elements available in List/Tuple/Sets/Dictionary.
    1. False
    2. **True**