



Siddaganga Institute of Technology, Tumakuru – 572 103

(An Autonomous Institution affiliated to VTU, Belagavi, Approved by AICTE, New Delhi)

Second Semester Master of Business Administration Examinations, November 2022

Big Data Tools for Managers

Max. Marks: 50

Time: 1 ½ Hours

Note : Answer all questions.

- 1 In the early 60s, technology witnessed problem with velocity or real-time data assimilation. This inspired the evolution of _____ technology.
 - a) Database
 - b) Files
 - c) Big Data
 - d) None
- 2 Is Big Data usually unstructured and qualitative in nature?
 - a) True
 - b) False
- 3 Data in _____ bytes size called big data.
 - a) Peta (1 PB = 1000TB)
 - b) Giga (1 GB = 1000MB)
 - c) Tera (1TB = 1000GB)
 - d) Mega Byte (1MB = 1000KB)
- 4 What kind of data big data can manage _____ (Multiple Choice)?
 - a) Structured
 - b) Unstructured
 - c) Semi-structured
 - d) Mix-data
- 5 Database is an organized collection of data so that it can be easily manager and accessible.
 - a) True
 - False
- 6 Choose the primary characteristics of big data among the following:
 - a) Value
 - b) Variety
 - c) Volume
 - d) All of the above
- 7 Volume is one of the characteristics of big data. What does Volume refer to?
 - a) The hard disk or server capacity it can hold
 - b) The data that can be processed
 - c) The number of YouTube/Netflix videos that existed
 - d) The amount of data in variety of formats
- 8 To define Data, which statement is right from option?
 - a) Large Textual contents of files
 - b) Small piece of information
 - c) Various files in Folders
 - d) None
- 9 RDBMS stands for _____.
 - a) Relational Database Management System
 - b) Raw Database Management System
 - c) Relation Data Manipulation Software
 - d) None

- 10 Columns/Attributes/Field all are same in RDBMS.
a) True b) False
- 11 _____ tools and software used to write database related programming.
a) MySQL b) Notepad c) Jupyter d) None
- 12 To display data from students table, choose correct statement from below option.
a) SELECT # FROM students;
b) SELECT + FROM students WHERE Null;
c) SELECT * FROM students;
d) None
- 13 Which data manipulation command is used to combines the records from one or more tables?
a) Select b) PROJECT c) PRODUCT d) JOIN
- 14 In SQL – the function AVG, MIN, MAX, SUM, COUNT are called as _____.
a) Adjunct function
b) Set operation
c) Scaler operation
d) Aggregate function
- 15 Statement used to display table structures in MySQL.
a) DESCRIBE b) STR c) DISPLAY d) None
- 16 Which join used get the all the records from left table and matching records from right table?
a) LEFT JOIN b) INNER JOIN
c) CROSS JOIN d) None
- 17 UPDATE statement will update all the records if WHERE clause is not specified.
a) FALSE b) TRUE
- 18 Which statement used to remove complete table from MySQL?
a) DELETES b) TRUNCATE c) REMOVE d) DROP
- 19 DELETE statement without where clause delete all the records from Table.
a) TRUE b) FALSE
- 20 Statement used to get unique records from the Table.
a) UNIQUE b) DISTINCT
c) UNION d) None
- 21 Select software required to write and run R Code. (Multiple choice).
a) R b) Notebook c) R Studio d) Visual Studio
- 22 R studio has basically _____ window sessions.
a) Two b) Five c) Three d) Four
- 23 Select most common data structures are available in R (Multiple choice).
a) Vector b) Matrix c) DataFrame d) Integer
- 24 Function used to display values of the vectors or any variables in R.
a) print() b) abs() c) prod() d) sum()

Contd...

25 Operator used to access the variable/column of DataFrame?

- a) \$(dollar)
- b) #(hash)
- c) .(dot)
- d) None

26 Function used to transpose of the Matrix/DataFrame in R?

- a) t()
- b) T()
- c) c()
- d) None

27 Function to display top records from R DataFrame?

- a) head()
- b) print()
- c) tail()
- d) None

28 Haven package/library used to read SPSS/SAV file in R Code.

- a) True
- b) False

29 Which command is correct to create character ('A','B','C','D','E','F') vectors in R?

- a) c('A','B','C','D','E','F')
- b) c(A,B,C,D,E,F)
- c) c(AL,BL,CL,DL,EL,FL)
- d) None

30 Function used to read comma separated values in R.

- a) read.csv()
- b) read_csv()
- c) read.xls()
- d) None

31 Most common data types in Python programming languages (Multiple choice).

- a) int
- b) float
- c) bool
- d) Str

32 Python is general-purpose interpreted, interactive, object-oriented, and high-level programming language.

- a) True
- b) False

33 Software used to run Python code.

- a) Jupyter notebook
- b) R studio
- c) phpMyAdmin
- d) None

34 Select the Quotation which Python supports (Multiple selection).

- a) Single
- b) Double
- c) Triple
- d) Four

35 Which statement Python used to define sets of conditional statements?

- a) If..else
- b) having
- c) Where
- d) None

36 Which of the following is used to define a block of code in Python language?

- a) Indentation
- b) Key
- c) Brackets
- d) All of the mentioned

37 To define comments in python _____ operator/symbol used.

- a) //
- b) #
- c) /**/
- d) None

38 Read CSV file in Pandas, which functions used to read file and store as DataFrame.

- a) read_CSV
- b) read.CSV
- c) read_file
- d) None

39 Python reserved words are python keyword which can be used as Variable/Functions names.

- a) False
- b) True

40 A variable name can only contain alpha-numeric characters and underscores (A-z, 0-9, and _).

- a) True
- b) False

- 41 _____ is package manager in Python to download and install Python packages.
a) download b) pip c) install d) None
- 42 To define/create Function in Python which keyword _____ is used.
a) def b) describe c) colon d) None
- 43 _____ Pandas Function to get the frequency of categorical/labeled variable.
a) count b) freq c) value_counts d) None
- 44 Loop is useful for iterating elements of List/Set/Tuple.
a) False b) True
- 45 To view number of rows and columns for Pandas data, which is best option to get (assume data is Pandas DataFrame).
a) data.shape b) print(data) c) data d) None
- 46 In Python, Dictionaries are immutable.
a) False b) True
- 47 Sets always contains repeated value.
a) False b) True
- 48 To visualize graphs in Pandas which is most common library used.
a) Ploty b) Matplotlib
c) Seaborn d) None
- 49 len() functions to display the no of elements available in List/Tuple/Sets/Dictionary.
a) False b) True
- 50 The process of breaking out long-form text sentences and words called.
a) Stem b) Cluster c) Bag d) Tokens
-