

**PUNE VIDYARTHI GRIHA'S  
COLLEGE OF SCIENCE & TECHNOLOGY**

**Final Examination Sept-2020  
T.Y.BSc.Computer Science (Semester-VI)**

**Subject – Data Science**

1. Data types are differed on basis of-----
  - a. the way of look
  - b. the type of operations
  - c. the type of operators used
  - d. the way of storage
2. Which of the following step is performed by data scientist after acquiring the data ?
  - a. Data Cleansing
  - b. Data Integration
  - c. Data Replication
  - d. All of the Mentioned
3. Which of the following command allows you to update the repository ?
  - a. Push
  - b. Pop
  - c. Update
  - d. None of the mentioned
4. Which of the following systems record changes to a file over time ?
  - a. Record Control
  - b. Version Control
  - c. Forecast Control
  - d. None of the mentioned
5. Which of the following is good way of performing experiments in data science ?
  - a. Measure variability
  - b. Un Generalize to the problem
  - c. Don't Have Replication
  - d. Not analysing
6. Which of the following package is used for reading excel data ?
  - a. Xlsx
  - b. Xlsc
  - c. read.sheet
  - d. All of the Mentioned
7. Which of the following Approach should be used to ask Data Analysis Question?
  - a. Data Cleansing
  - b. Data Integration
  - c. Data Replication
  - d. All of the Mentioned

8. Binary attribute are -----
- This takes only two values. In general, these values will be 0 and 1
  - The natural environment of a certain species
  - Systems that can be used without knowledge of internal operations
  - None of these
9. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Which of the following statement is true in following case?
- Feature F1 is an example of nominal variable.
  - Feature F1 is an example of ordinal variable.
  - It doesn't belong to any of the above category.
  - Both of these
10. Which of the following is not a NoSQL database?
- SQL Server
  - MongoDB
  - Cassandra
  - None of these
11. NoSQL databases is used mainly for handling large volumes of \_\_\_\_\_ data.
- Unstructured
  - Structured
  - Semi-Structured
  - All of these
12. Most NoSQL databases support automatic \_\_\_\_\_ meaning that you get high availability and disaster recovery.
- Processing
  - Scalability
  - Replication
  - All of the mentioned
13. Which of the following is not a data visualization Technique
- Histogram
  - Pie Chart
  - Box Plot
  - Docx File
14. Which of these is not a High-Level Language
- Fortan
  - PHP
  - Java
  - C
15. Which of the following step is performed by data scientist after Cleaning the data ?
- Data Processing
  - Data Collection
  - Data Analysis
  - Data Modelling

16. \_\_\_\_\_ data is difficult to manipulate and typically needs to be processed in some way before it can be used in standard data analysis software.
- Structured data
  - Unstructured data
  - Summerized data
  - Frequency data
17. Which of the following function is used for determining missing values?
- any
  - Nan
  - not
  - all
18. Point out the correct statement.
- Primary data is original source of data
  - Secondary data is original source of data
  - Questions are obtained after data processing steps
  - None of the Mentioned
19. \_\_\_\_\_ measures asymmetry about the mean of the probability distribution of a random variable.
- skewness
  - covariance
  - variance
  - Kurtosis
20. \_\_\_\_\_ shows all individual data points.
- Box-plot
  - scatter plot
  - line plot
  - pie chart
21. What is true about Data Visualization?
- Data Visualization is used to communicate information clearly and efficiently to users by the usage of information graphics such as tables and charts.
  - Data Visualization helps users in analyzing a large amount of data in a simpler way.
  - Data Visualization makes complex data more accessible, understandable, and usable.
  - Data visualization is an IDE
22. Which of the following is known as raw data?
- tidy data
  - processed data
  - clean data
  - formatted data
23. Data cannot be visualized using?
- graphs
  - charts
  - maps

- d. image
24. Data visualization is also an element of the broader \_\_\_\_\_.
- a. deliver presentation architecture
  - b. data presentation architecture
  - c. dataset presentation architecture
  - d. data process architecture
25. Which method shows hierarchical data in a nested format?
- a. Treemaps
  - b. Scatter plots
  - c. Population pyramids
  - d. Area charts