

Unit-I

1. Data in _____ bytes size is called Big Data.

- A. Tera
- B. Giga
- C. Peta
- D. Meta

View Answer

Ans : C

Explanation: data in Peta bytes i.e. 10^{15} byte size is called Big Data.

2. How many V's of Big Data

- A. 2
- B. 3
- C. 4
- D. 5

View Answer

Ans : D

Explanation: Big Data was defined by the “3Vs” but now there are “5Vs” of Big Data which are Volume, Velocity, Variety, Veracity, Value

3. Transaction data of the bank is?

- A. structured data
- B. unstructured data
- C. Both A and B
- D. None of the above

View Answer

Ans : A

Explanation: Data which can be saved in tables are structured data like the transaction data of the bank.

4. In how many forms BigData could be found?

- A. 2
- B. 3
- C. 4
- D. 5

View Answer

Ans : B

Explanation: BigData could be found in three forms: Structured, Unstructured and Semi-structured.

5. Which of the following are Benefits of Big Data Processing?

- A. Businesses can utilize outside intelligence while taking decisions
- B. Improved customer service
- C. Better operational efficiency
- D. All of the above

View Answer

Ans : D

Explanation: All of the above are Benefits of Big Data Processing.

6. Which of the following are incorrect Big Data Technologies?

- A. Apache Hadoop
- B. Apache Spark
- C. Apache Kafka
- D. Apache Pytarch

View Answer

Ans : D

Explanation: Apache Pytarch is incorrect Big Data Technologies.

7. The overall percentage of the world's total data has been created just within the past two years is ?

- A. 80%
- B. 85%
- C. 90%
- D. 95%

View Answer

Ans : C

Explanation: The overall percentage of the world's total data has been created just within the past two years is 90%.

8) 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

Ans: Data Cleansing

9) 3V's are not sufficient to describe big data.

- a) True
- b) False

Ans: True

10. Communicative and collaborative is one among the key skill sets and behavioral characteristics of a data scientist [True / False]?

- a. True
- b. False

Answer : a

11. ----- are the sources of Bigdata [select all that apply]

- I. Book
- II. Facebook
- III. Genome sequence
- IV. Video Surveillance

Ans:

12. BI analyses the past data and make future predictions True/False ?

- a. True
- b. False

Answer : b

12. In which phase of data analytics ETLT is performed?

Ans: Phase 2 Data preparation is done in this phase. An analytical sandbox is used in this to perform analytics for the entire duration of the project. While you explore, preprocess and condition data, modeling follows suit. To get the data into the sandbox, you will perform ETLT (extract, transform, load and transform).

A. Discovery

B. Model Planning

C. Model Building

D. Data Preparation

13. In which data analytics lifecycle phase is an analytic sandbox prepared?

Phase 2 — Data preparation: Phase 2 requires the presence of an analytic sandbox, in which the team can work with data and perform analytics for the duration of the project. The team needs to execute extract, load, and transform (ELT) or extract, transform and load (ETL) to get data into the sandbox.

A. Data Preparation

B. Model Planning

C. Model Building

D. Discovery

14. In which phase would the team expect to invest most of the project time?

A. Data Preparation

B. Model Planning

C. Model Building

D. Discovery

15. In which phase would the team expect to invest least time of the project time?

A. Data Preparation

B. Model Planning

C. Model Building

D. Discovery

16. from following tools which tool is used for Model building?

a. Hadoop b. Octave c. OpenRefine d. All of Above

Ans B

17. from following tools which tool is used for Data preparation

a. Alpine Miner b. Excel c. Matlab d. Weka

Ans . A

18. To determine if the project was completed on time and within budget, is the key role of _____

A. Project Sponsor

B. Project Manager

C. Data Engineer

D. Data Scientist

19. How many Phases are there in Data Analytics Lifecycle?

A. 3

B. 6

C. 7

D. Any

20. In data Analytics life cycle we can move back and refine the work done. True or False

A. True

B. False

21. What are the key outputs from Analytics Projects?

A. PPT

B. report

C. code

D. All of above

22. _____ provides subject matter expertise for analytical techniques, data modeling and applying valid analytical techniques to give business problems.

A. Project Sponsor

B. Project Manager

C. Data Engineer

D. Data Scientist