

## BIG DATA

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. Apache Kafka is an open-source platform that was created by?

- A. LinkedIn
- B. Facebook

C. Google  
D. IBM  
View Answer  
Ans : A

Explanation: Apache Kafka is an open-source platform that was created by LinkedIn in the year 2011.

9. What was Hadoop named after?

A. Creator Doug Cutting's favorite circus act  
B. Cuttings high school rock band  
C. The toy elephant of Cutting's son  
D. A sound Cutting's laptop made during Hadoop development  
View Answer  
Ans : C

Explanation: Doug Cutting, Hadoop creator, named the framework after his child's stuffed toy elephant.

10. What are the main components of Big Data?

A. MapReduce  
B. HDFS  
C. YARN  
D. All of the above  
View Answer  
Ans : D

Explanation: All of the above are the main components of Big Data.

11. All of the following accurately describe Hadoop, EXCEPT \_\_\_\_\_

A. Open-source  
B. Real-time  
C. Java-based  
D. Distributed computing approach  
View Answer  
Ans : B

Explanation: Apache Hadoop is an open-source software framework for distributed storage and distributed processing of Big Data on clusters of commodity hardware.

12. \_\_\_\_\_ has the world's largest Hadoop cluster.

- A. Apple
- B. Datamatics
- C. Facebook
- D. None of the above

[View Answer](#)

Ans : C

Explanation: Facebook has many Hadoop clusters, the largest among them is the one that is used for Data warehousing.

13. Facebook Tackles Big Data With \_\_\_\_\_ based on Hadoop.

- A. Project Prism
- B. Prism
- C. Project Big
- D. Project Data

[View Answer](#)

Ans : A

Explanation: Prism automatically replicates and moves data wherever it's needed across a vast network of computing facilities.

14. \_\_\_\_\_ is general-purpose computing model and runtime system for distributed data analytics.

- A. Mapreduce
- B. Drill
- C. Oozie
- D. None of the above

[View Answer](#)

Ans : A

Explanation: Mapreduce is general-purpose computing model and runtime system for distributed data analytics.

15. The examination of large amounts of data to see what patterns or other useful information can be found is known as

- A. Data examination
- B. Information analysis
- C. Big data analytics
- D. Data analysis

[View Answer](#)

Ans : C

Explanation: The examination of large amounts of data to see what patterns or other useful information can be found is known as Big data analytics.

16. Big data analysis does the following except?

- A. Collects data
- B. Spreads data
- C. Organizes data
- D. Analyzes data

[View Answer](#)

Ans : D

Explanation: Big data analysis does the following except Spreads data.

17. What makes Big Data analysis difficult to optimize?

- A. Big Data is not difficult to optimize
- B. Both data and cost effective ways to mine data to make business sense out of it
- C. The technology to mine data
- D. None of the above

[View Answer](#)

Ans : B

Explanation: Both data and cost effective ways to mine data to make business sense out of it makes Big Data analysis difficult to optimize.

18. The new source of big data that will trigger a Big Data revolution in the years to come is?

- A. Business transactions
- B. Social media
- C. Transactional data and sensor data
- D. RDBMS

[View Answer](#)

Ans : C

Explanation: The new source of big data that will trigger a Big Data revolution in the years to come is Transactional data and sensor data.

19. The unit of data that flows through a Flume agent is

- A. Log
- B. Row
- C. Record
- D. Event

[View Answer](#)

Ans : D

Explanation: The unit of data that flows through a Flume agent is Event.

20. Listed below are the three steps that are followed to deploy a Big Data Solution except

- A. Data Processing
- B. Data dissemination
- C. Data Storage
- D. Data Ingestion

[View Answer](#)

Ans : B

Explanation: Listed below are the three steps that are followed to deploy a Big Data Solution except Data dissemination.