

Roll No. :

JAMIA HAMDARD

Total Pages = 2 Pages

Department of Computer Science & Engineering
School of Engineering Sciences & Technology
New Delhi-110062

B.C.A. - 5th Semester (3rd Year)

Supplementary/Improvement Examination 2023 (For Odd Semester Exam 2022)

Paper Name : Introduction to Data Science and Big Data **Paper Code** : BCA PE523
Time : 3 Hours **Maximum Marks** : 75

Section - A (35 Marks)

All questions are compulsory. All questions carry 1 Marks

Q1.	Big data is explained with 5Vs. True/False
Q2.	Traditional database management tools can handle the size and complexity of big data. True/False
Q3.	Data Science is the same as Data Analytics. True/False
Q4.	Businesses can utilize information from big data to maintain a competitive advantage. True/False
Q5.	Data analytics helps accountants to analyze data and detect fraud. True/False
Q6.	Big data analytics technologies are necessary to extract valuable insights from the data. True/False
Q7.	Focusing on business goals and how to use big data analytics technologies to meet them is the recommended best practice for managing big data analytics programs. True/False
Q8.	The interrelatedness of data and the amount of development work that will be needed to link various data sources can be handled using big data analytics. True/False
Q9.	For organizations that aren't currently looking to do big data analytics, there is little or no benefit to examining the data they're retaining and evaluating how it's being used. True/False
Q10.	Hadoop 2.0 allows live stream processing of real-time data. True/False
Q11.	The programming model, MapReduce, used by Hadoop is simple to write and test. True/False
Q12.	Apache Hadoop is an open-source software framework. True/False
Q13.	Hive is a SQL-based data warehouse system for Hadoop. True/False
Q14.	Hive supports custom extensions written in Java. True/False
Q15.	Amazon Web Service Elastic MapReduce (EMR) is Amazon's packaged Hadoop offering. True/False
Q16.	Data Cleaning emphasizes on prediction, based on well-known properties learned from the training data. True/False
Q17.	Machine learning emphasizes on prediction, based on well-known properties learned from the training data. True/False
Q18.	Big Data is just about any data more than 100GB of data. True/False
Q19.	Data in a Relational Database is structured. True/False
Q20.	Most social media platforms uses "Big Data" to determine the behavior of its users. True/False
Q21.	Data generated from online transactions is one of the example for volume of big data. True/False

Q22.	Hadoop MapReduce allows you to perform distributed parallel processing on large volumes of data quickly and efficiently. True/False
Q23.	R Programming is a Big data technology. True/False
Q24.	Big data is different than "Business Intelligence". True/False
Q25.	Big data is different than "intelligent data mining". True/False
Q26.	Hadoop run on Cross-platform. True/False
Q27.	What is full form of ACID ?
Q28.	What is full form of HDFS ?
Q29.	What is full form of SQL ?
Q30.	What is full form of YARN ?
Q31.	What is the name of the programming framework originally developed by Google that supports the development of applications for processing large data sets in a distributed computing environment?
Q32.	Which organization has the world's largest Hadoop cluster ?
Q33.	The most popular high-level Java API in Hadoop Ecosystem is _____
Q34.	Hadoop is a framework that works with a variety of related tools. Apart from MapReduce and Hbase, common cohorts include _____
Q35.	Method by which companies analyze customer data in an effort to identify patterns and discover relationships between different data elements is often referred to as _____

Section - B (20 Marks)

Attempt any FOUR questions. All questions carry 5 Marks

Sl.No.	Question	Marks	CO	BL
Q1.	What is big data? Explain its importance in computer science..	5	CO1	L2
Q2.	Explain CAP Theorem and ACID property.	5	CO2	L1
Q3.	Explain Aggregate Data Model and Graph Database Model.	5	CO3	L3
Q4.	Explain with proper example Crowd sourcing analytics	5	CO4	L5
Q5.	What is Hadoop ? Explain HDFS and Map Reduce .	5	CO4	L4
Q6.	Explain how Big data can help in Credit Risk Management	5	CO6	L6

Section - C (20 Marks)

Attempt any TWO questions. All questions carry 10 Marks

Sl.No.	Question	Marks	CO	BL
Q1.	Explain 5 V of big data.	10	CO1	L2
Q2.	How is big data analyzed?	10	CO2	L1, L4
Q3.	Compare Inter and Trans firewall analytics	10	CO3	L3, L6
Q4.	Explain any 5 application of Big Data.	10	CO4	L5

----- The End -----