

Total No. of printed pages = 3

CSE 1818 PE 61

91423

Roll No. of candidate

--	--	--	--	--	--	--	--	--	--

Bina Chowdhury Central Library
Girijananda Chowdhury University
Hatkhowapara, Azara, Ghy-17

2023

B.Tech. 8th Semester End-Term Examination

CSE

BIG DATA ANALYTICS

(New Regulation w.e.f. 2017-18 & New Syllabus w.e.f. 2018-19)

Full Marks – 70

Time – Three hours

The figures in the margin indicate full marks for the questions.

Answer question No. 1 and any *four* from the rest.

1. Answer the following multiple choice questions : (10 × 1 = 10)
- (i) The word 'Big data' was coined by
- (a) Roger Mougaldas (b) John Philips
(c) Simon Woods (d) Martin Green
- (ii) Data in _____ bytes size is called Big Data.
- (a) Tera (b) Giga
(c) Peta (d) Meta
- (iii) Concerning the Forms of Big Data, which one of these is odd?
- (a) Structured (b) Unstructured
(c) Processed (d) Semi-Structured
- (iv) Listed below are the three steps that are followed to deploy a Big Data Solution except
- (a) Data Ingestion (b) Data Processing
(c) Data dissemination (d) Data Storage
- (v) Which industries below listed employ the use of so-called "Big Data" in their day to day operations?
- (a) Weather forecasting (b) Marketing
(c) Healthcare (d) All of the above

[Turn over

- (vi) 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
- (vii) _____ has the world's largest Hadoop cluster.
- (a) Apple (b) Datamatics
(c) Facebook (d) None of the above
- (viii) The feature of big data that refers to the quality of the stored data is _____.
- (a) Variety (b) Volume
(c) Variability (d) Veracity
- (ix) _____ is general-purpose computing model and runtime system for distributed data analytics.
- (a) MapReduce (b) Drill
(c) Oozie (d) None of the above
- (x) All of the following accurately describe Hadoop, except:
- (a) Open-source (b) Real-time
(c) Java-based (d) Distributed computing approach

Answer any four from the following

Bina Chowdhury Central Library
Girijananda Chowdhury University
Hatkhowapara, Azara, Ghy-17

(3 × 5 = 15)

2. Answer the following

- (a) What are the properties of Big-data system?
(b) Differentiate between vertical and horizontal scaling?
(c) How distributed file systems meet the storage requirement checklist?

(5 × 3 = 15)

3. Answer the following

- (a) Write about different 5Vs that define a Big data system.
(b) What is Big data analytics ?
(c) Write on low latency reads and updates.
(d) Write about different types of data.
(e) Explain graph schema.

(3 × 5 = 15)

4. Answer the following

- (a) Is there any scalability issue with relational database? Justify your answer.
(b) What is CAP theorem ?
(c) Write on data robustness and fault tolerance.

5. Answer the following (3 × 5 = 15)

- (a) How debugability is accomplished in the Lambda Architecture?
- (b) What is immutable schema?
- (c) What are the advantages of using immutable schema in Big data systems?

6. Answer the following (3 × 5 = 15)

- (a) "Unstructured data is rawer than normalized data" – justify your answer.
- (b) Write short note on Hadoop distributed file system.
- (c) Write short note on MongoDB.

Bina Chowdhury Central Library
Girijananda Chowdhury University
Hatkhowapara, Azara, Ghy-17

