

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2022
(Sixth Semester)

Branch – COMPUTER TECHNOLOGY

DATA ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (10 x 1 = 10)

- 1 Data in _____ bytes size is called Big Data
(i) Tera (ii) Giga (iii)Peta (iv) Meta
- 2 Transaction data of the Banks are _____
(i) Structured Data (ii) Unstructured Data
(iii) Both A & B (iv) None of the Above
- 3 A _____ node acts as the Slave and is responsible for executing a Task assigned to it by the JobTracker.
(i) Map Reduce (ii) Mapper
(iii) Task Tracker (iv) Job Tracker
- 4 _____ function is responsible for consolidating the results produced by each of the Map() functions/tasks.
(i) Reduce (ii) Map
(iii) Reducer (iv) All of the Mentioned.
- 5 A Framework group reducers input key in _____ stage.
(i) Sort (ii) Shuffle
(iii) Reduce (iv) None of the mentioned.
6. _____ maps input Key/Value pairs to a set of intermediate key/Value Pairs
(i) Reducer (ii) Mapper
(iii) Streaming (iv) None of the mentioned.
- 7 Which of the following techniques would perform better for reducing dimensions of a data set?
(i) Removing columns which have too many missing values
(ii) Removing columns which have high variance in data
(iii) Removing columns with dissimilar data trends
(iv) None of these
- 8 Which of the following algorithms cannot be used for reducing the dimensionality of data?
(i) t-SNE (ii) PCA (iii) LDA False (iv) None of these
- 9 Which of the following is a NoSQL Database Type?
(i) SQL (ii) Document Databases
(iii) JSON (iv) All of the Mentioned
- 10 Which of the following is the Wide Column Store?
(i) Cassandra (ii) Riak
(iii) MongoDB (iv) Redis

Cont...

SECTION - B (25 Marks)

Answer ALL questions.

ALL questions carry EQUAL Marks (5 x 5 = 25)

- 11 a Summarize the different Sources of Big Data
OR
b Elaborately Elucidate the Data Mining Process with appropriate Example.
- 12 a Illustrate Map – Reducing Task with Suitable Example.
OR
b Explain the Strategies to be followed to cope up with Node Failures.
- 13 a Demonstrate the Relational Algebra Operations.
OR
b Discuss on Matrix Multiplication with One Map Reduce Step.
- 14 a Explain Root – Mean Square Error of Decomposition Algorithm.
OR
b Write short notes on algorithms recommended for optimizing an arbitrary element.
- 15 a Write short notes on On-Disk Storage devices.
OR
b Elaborate the NewSQL database for Big Data Storage Technology.

SECTION -C (40 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 x 8 = 40)

- 16 a Discuss various types of Big Data with neat Sketch.
OR
b Explain the importance of Big data Analysis
- 17 a Demonstrate the Map Reduce tasks with appropriate example.
OR
b Analyze the Grouping of Keys for the Map Reduce Execution with neat sketch.
- 18 a Demonstrate Computing by Selection and Projection by Map Reduce with suitable example
OR
b Explain in detail on Union, Selection and Difference operations by Map Reduce with suitable Example.
- 19 a Illustrate Incremental Computation of a UV-Decomposition method
OR
b Explain Root – Mean Square Error of Decomposition Algorithm.
- 20 a Compile the various Big Data Storage strategies and devices
OR
b Explain in detail about NoSQL databases.

Z-Z-Z

END