

**(FOR THE CANDIDATES ADMITTED
DURING THE ACADEMIC YEAR 2020 ONLY)**

(NO. OF PAGES:2)

20 UCS 6E5

REG. NO:

**N.G.M.COLLEGE(AUTONOMOUS): POLLACHI
END-OF-SEMESTER EXAMINATIONS: MAY-2023**

**B.Sc.- COMPUTER SCIENCE (S.F.)
VI SEMESTER**

**MAXIMUM MARKS : 70
TIME : 3 HOURS**

PART-III

BIG DATA ANALYTICS

SECTION-A

(10 X 1 = 10 MARKS)

ANSWER THE FOLLOWING QUESTIONS.

MULTIPLE CHOICE QUESTIONS.

(K1)

1. On which of the following platforms does Hadoop run?
 - a) Debian
 - b) Cross-platoform
 - c) Bare metal
 - d) Unix-like
- 2.. ____ is the most popular high-level java API in Hadoop Ecosystem.
 - a) Scalding
 - b) HCatalog
 - c) Cascalog
 - d) Cascading
3. Which is not a type of NoSQL?
 - a. HBase
 - b. QBase
 - c. CouchDB
 - d. MangoDB
4. While querying a hive table for a column with MAP data type, the elements in the column are referenced by
 - a) Index
 - b) Key
 - c) Value
 - d) nested Index
5. Data in _____ bytes size is called big data
 - a) Meta
 - b) Giga
 - c) Tera
 - d) Peta

SHORT ANSWERS

(K2)

6. Define Big Data.
7. What is the advantage of Hadoop ecosystems?.
8. Compare NoSQL vs SQL.
9. What is of Hive.
10. What is the solution of big data?

(CONTD.....2)

SECTION – B**(5 X4 = 20 MARKS)****ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS****(K3)**

11. a) Given below is the data regarding the electrical consumption of an organization. It contains the monthly electrical consumption and the annual average for various years.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Avg
1979	23	23	2	43	24	25	26	26	26	26	25	26	25
1980	26	27	28	28	28	30	31	31	31	30	30	30	29
1981	31	32	32	32	33	34	35	36	36	34	34	34	34
1984	39	38	39	39	39	41	42	43	40	39	38	38	40
1985	38	39	39	39	39	41	41	41	00	40	39	39	45

If the above data is given as input, we have to write applications to process it and produce results such as finding the year of maximum usage, year of minimum usage, and so on. Write a Program using MapReduce framework using the sample data.

(OR)

b) How is an application submitted in Hadoop YARN? Explain with example.

12. a) How to install Apache Pig on a Hadoop cluster on Ubuntu? Give Step by Step Instruction.

(OR)

b) Write the Syntax of various PIG Commands in Hadoop with example.

13. a) Construct with example using Apache Pig Group Operator.

(OR)

b) Interpret with example using Apache Pig ORDER BY Operator.

14. a) Describe in brief JOIN Clause use in HIVE Syntax with example.

(OR)

b) Write a R Program to find Minimum and Maximum of list of numbers.

15. a) Big Data Analytics in Manufacturing Industry . Construct a Case Study.

(OR)

b) Examine a Case Study of Location Tracking using Big Data Analytics

SECTION –C**(4 X 10 = 40 MARKS)****ANSWER ANY FOUR OUT OF SIX QUESTIONS****(16th QUESTION IS COMPULSORY AND ANSWER ANY THREE QUESTIONS)****(FROM Qn. No : 17 to 21)****(K4 (or) K5)**

16. Analyze in detail about Big Data Characteristics with example.

17. Summarize in brief about Hadoop Architecture with example.

18. Conclude in brief about various Components of Hadoop Ecosystem with example.

19. Point out HBASE Architecture with example.

20. Describe the HIVE Architecture with neat diagram.

21. Discuss in detail about Transforming Business Processes with Big Data with example.