

QP CODE: 220200946



Reg No :

Name :

M.C.A. DEGREE EXAMINATION, MARCH 2022
Second Semester
Core - MCACT204 - DATA SCIENCE & BIG DATA ANALYSIS

2020 Admission Onwards

936029AE

Time: 3 Hours

Maximum: 75 Marks

Part A

*Answer any **ten** questions*

*Each question carries **3** marks*

1. Illustrate data Discrimination with Example.
2. What is an Attribute?
3. Discuss Data Reduction Methods.
4. Explain hash-based technique.
5. Explain the two steps in classification.
6. What are the properties of the data objects within a cluster?
7. Explain the different facets of data in Data Science?
8. Explain the three V's of Big Data?
9. What do you know about the term "Big Data"? Explain its applications?
10. Write about Hadoop Ecosystem ?
11. What are the issues with relational model and non-relational model?
12. What is MapReduce framework?

(10×3=30 marks)

Part B

*Answer **all** questions*

*Each question carries **9** marks*

13. a) What kinds of patterns can be mined in datamining ?Explain.

OR

- b) Explain basic preprocessing Methods.





14. a) Illustrate FP Growth principle with an example.

OR

b) Explain the importance of Classification in Data mining.

15. a) What are the different Data Science process?

OR

b) Explain the different facets of data in Data Science?

16. a) Explain cloud computing and big data? Elaborate on cloud computing features and its types.

OR

b) What is MapReduce Framework? Diagrammatically explain the working of Map Reduce Architecture.

17. a) Compare the features of RDBMS and big data solutions.

OR

b) Write notes on a) data types in Hive b) Variables, properties and queries in Hive.

(5×9=45 marks)

