

QPCODE:

Reg.No :

Name :

MAHATMA GANDHI UNIVERSITY, KOTTAYAM
MGU-UGP (Honours)
Third Semester Examination
(2025 Admission Onwards)
MG3DSEMATA08: Introductory Course on Data Science

Duration: 2 Hours

Maximum Marks: 70

Remember(K), Understand(U), Apply(A), Analyse(An), Evaluate(E), Create(C), Skill(S), Interest(I) and Appreciation(Ap)

Students should attempt at least one question from each course outcome to enhance their overall outcome attainability.

Part A

Short Answer Questions

Answer any 5 questions

Each question carries 2 marks

1. Give examples for structured data and unstructured data [U] [1]
2. What is data mining [U] [1]
3. Define data breach [R] [2]
4. Define biometric information. [U] [2]
5. Define rule-induction [U] [3]
6. Mention any four tools used for data visualization [A] [3]
7. What is the difference between traditional data and big data [An] [4]
8. Define a graph database. [U] [4]

[2x5 = 10]

Part B

Short Essay Questions

Answer any 5 questions

Each question carries 6 marks

9. Explain some challenges in data science [U] [1]
10. Define arithmetic mean, median and mode [U] [1]
11. Why is data referred to as the 'new oil' [A] [2]
12. A bank uses an AI system to approve loans. A customer is denied a loan but receives no explanation. Discuss the accountability issues involved and suggest governance measures. [A] [2]
13. Calculate the coefficient of correlation and comment on the relationship:
X: 10, 20, 30, 40, 50
Y: 45, 40, 35, 30, 25" [An] [3]
14. A healthcare organization wants to track the spread of an infectious disease and identify high-risk patients. Explain how a graph database can help achieve this objective. [E] [4]
15. Explain different chart types available in Qlik Sense. [An] [4]
16. Explain ACID properties [An] [4]

[6x5 = 30]

Part C

Essay Questions

Answer any 3 questions

Each question carries 10 marks

17. Explain the four levels of data [R] [1]
18. Discuss the major ethical issues in data science, including privacy, security, fairness, transparency, and accountability. [U] [2]
19. Explain the concept of correlation and discuss different types of correlation with examples. [An] [3]
20. A hospital plans to predict the likelihood of diseases using patient records. Explain how the six-step Data Science process can be applied to this problem. [An] [3]

21. Explain Hadoop framework

[U] [4]

22. A government agency wants to analyze large-scale census and public service data. Discuss how Apache Spark can help in data integration, analysis, and policy decision-making.

[S] [4]

[10x3 = 30]