

(Pages : 3)

M – 5990

Reg. No. :

Name :

Second Semester M.A. Degree Examination, March 2022

Behavioural Economics And Data Science

BEDS 523 – FOUNDATIONS OF DATA SCIENCE

(2020 Admission)

Time : 3 Hours

Max. Marks : 75

SECTION – A

Answer **all** questions from this section. Each question carries **1** mark.

1. What is Data Science useful for?
2. What is datafication in data science?
3. In which step of data reduction redundant attributes are detected.
4. _____ works to remove the noise from the data that includes techniques like binning, clustering, and regression.
5. Clustering is an unsupervised learning method (True/false).
6. Which clustering technique needs the merging approach?
7. The final output of Hierarchical clustering is _____
8. When does k-means clustering stop creating or optimizing clusters?
9. What can we predict from scatter plots?
10. What is the use of Histograms?

(10 × 1 = 10 Marks)

P.T.O.



SECTION – B

Answer any **seven** questions. Each question carries **5** marks.

11. What is meant by fitting the model? Explain.
12. Explain is the main purpose of principal component analysis?
13. How do you do exploratory data analysis? Illustrate with an example.
14. What are the steps in the data science process? Explain any two.
15. How to Use ROC Curves for Classification.
16. What is linear regression used for? Explain with an example
17. With an example explain K nearest neighbor classifier.
18. Explain the concept of clustering tendency.
19. Explain in detail agglomerative clustering.
20. Explain the different methods used for visualization.

(7 × 5 = 35 Marks)

SECTION – C

Answer any **three** questions. Each question carries **10** marks.

21. Give a detailed account on Association Rule Mining Algorithm with suitable example.
22. Explain in detail the need of confusion matrix with an example.



23. Consider the following 5 transactions in a grocery store. Find all the items being frequently bought together. Extract all rules with confidence above 80% and support above 60% from the following data:

Customer	Items
1	{Q,W,E,R,T,Y}
2	{A,W,E,R,T,Y}
3	{Q,S,R,T}
4	{Q,P,D,R,Y}
5	{D,W,W,R,B,T}

24. Suppose we have 4 types of data objects and each has two attributes. Our goal is to group these objects into K=2 group of objects. Assume the initial centroids are objects O1 and O2.

DATA OBJECTS	ATTRIBUTE 1	ATTRIBUTE 2
O1	1	1
O2	2	1
O3	4	3
O4	5	4

25. What is the need of data visualization process? Which are the different tools used for visualization?

(3 × 10 = 30 Marks)

