

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)
MSc DEGREE EXAMINATION DECEMBER 2024
(First Semester)

Branch - **COMPUTER SCIENCE**
DATA MINING AND ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	Which of the following can be considered as the correct process of Data Mining? a. Infrastructure, Exploration, Analysis, Interpretation, Exploitation b. Exploration, Infrastructure, Analysis, Interpretation, Exploitation c. Exploration, Infrastructure, Interpretation, Analysis, Exploitation d. Exploration, Infrastructure, Analysis, Exploitation, Interpretation	K1	CO1
	2	The primary use of data cleaning is: a. Removing the noisy data b. Correction of the data inconsistencies c. Transformations for correcting the wrong data d. All of the above	K2	CO1
2	3	Total fact table the star schema is composed of ____? a. 4 b. 3 c. 2 d. 1	K1	CO2
	4	What does OLTP stand for? a. Offline transaction processing b. Online transaction processing c. Outline traffic processing d. None	K2	CO2
3	5	A collection of one or more items is called as ____ a. Itemset b. Support c. Confidence d. Support Count	K1	CO3
	6	An itemset whose support is greater than or equal to a minimum support threshold is ____ a. Itemset b. Frequent Itemset c. Infrequent items d. Threshold values	K2	CO3
4	7	Which clustering algorithm does not require specifying the number of clusters in advance? a. K-Means b. DBSCAN c. Agglomerative d. Mean-Shift	K1	CO4
	8	Which clustering algorithm is sensitive to the order of the data points? a. K-Means b. DBSCAN c. Agglomerative d. Mean-Shift	K2	CO4
5	9	Data Analytics uses ____ to get insights from data. a. Statistical figures b. Numerical aspects c. Statistical methods d. None of the mentioned above	K1	CO5

Cont...

	10	_____ predict and recommend how best to achieve those results by analyzing data from previous event. a.Diagnostic analytics b. Descriptive analytics c. Predictive analytics d.Prescriptive analytics	K2	CO5
--	----	---	----	-----

SECTION - B (35 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Compare and contrast characterization and discrimination of data mining task primitives.	K2	CO1
		(OR)		
	11.b.	Identify various tasks of data cube aggregation in data reduction.		
2	12.a.	Inspect how will you convert tables and spreadsheets to a data cube.	K4	CO2
		(OR)		
	12.b.	Differentiate Star, Snowflakes and Fact constellation schemas.		
3	13.a.	Discuss in detail the apriori algorithm with an example.	K2	CO3
		(OR)		
	13.b.	Compare and contrast classification and prediction method.		
4	14.a.	List and discuss the requirements of clustering in data mining.	K4	CO4
		(OR)		
	14.b.	List and discuss the parameters required for DBSCAN?		
5	15.a.	Differentiate descriptive and diagnostic data analytics with suitable examples.	K4	CO5
		(OR)		
	15.b.	Distinguish Predictive and Prescriptive data analytics with suitable examples.		

SECTION -C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Develop and design the steps in the process of Knowledge Discovery and explain the architecture of data mining.	K6	CO1
2	17	Explain the data warehouse architecture with suitable diagram.	K5	CO2
3	18	How to improve the efficiency of the Apriori Algorithm? Explain in detail.	K5	CO3
4	19	Explain in detail the CLARANS clustering algorithm.	K5	CO4
5	20	Elaborate the purpose and use of data analytics in the field of agriculture and telecommunication industry.	K6	CO5