PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION DECEMBER 2022
(First Semester)

Branch – **COMPUTER SCIENCE**

DATA MINING AND ANALYTICS/DATA MINING AND ITS APPLICATIONS

Time: Three Hours   Maximum: 50 Marks

SECTION-A (5 Marks)
Answer ALL questions
ALL questions carry EQUAL marks   (5 x 1 = 5)

1. The data is stored, retrieved and updated in _____
   (i) OLAP   (ii) OLTP   (iii) SMTP   (iv) FTP

2. Which of the following is not a component of a data warehouse?
   (i) Metadata   (ii) Current detail data
   iii) Lightly summarized data   (iv) Component Key

3. Which of the following process includes data cleaning, data integration, data selection, data transformation, data mining, pattern evolution and knowledge presentation?
   (i) KDD process   (ii) ETL process
   iii) KTL process   (iv) MDX process

4. _____ supports basic OLAP operations, including slice and dice, drill-down, roll-up and pivoting.
   (i) Information processing   (ii) Analytical processing
   (iii) Data mining   (iv) Transaction processing

5. The core of the multidimensional model is the _____ which consists of a large set of facts and a number of dimensions.
   (i) Multidimensional cube   (ii) Dimensions cube
   (iii) Data cube   (iv) Data model

SECTION - B (15 Marks)
Answer ALL Questions
ALL Questions Carry EQUAL Marks   (5 x 3 = 15)

6. a. Summarize the steps in the process of knowledge discovery.
   OR
   b. What are the major issues in data mining? Explain.

7. a. Explain the snowflake schema for multidimensional databases.
   OR
   b. Describe the purpose of metadata repository.

8. a. Distinguish between the frequent itemsets and closed itemsets.
   OR
   b. Write down the classification by decision tree induction.

9   a   Mention the requirements of clustering in data mining.

OR

     b   Enumerate the implementation of k-means method.

10   a   What are the advantages of data analytics? Explain.

OR

     b   Elaborate the function of descriptive analytics.

SECTION -C (30 Marks)
Answer ALL questions
ALL questions carry EQUAL Marks   (5 x 6 = 30)

11   a   Draw and explain the architecture of a typical data mining system.

OR

     b   Compare the dimensionality reduction and numerosity reduction.

12   a   Illustrate the OLAP operations in the multidimensional data model.

OR

     b   Outline the steps to build a data warehouse.

13   a   Discuss the generating association rules from frequent item sets.

OR

     b   How will you prepare the data for classification and prediction? Explain.

14   a   What are the various types of data in cluster analysis? Explain.

OR

     b   Examine the implementation of CLARANS algorithm.

15   a   Compare the predictive and prescriptive analytics.

OR

     b   Elaborate the data mining in health care and medicine.

Z-Z-Z   END