B.SC. (I. T.) SEM. - V (CBCS - 2015 COURSE) : SUMMER - 2018

SUBJECT: DATA WAREHOUSING AND DATA MINING

Day : **Thursday** Time: **02.30 p.m. to 05.30 p.m.**

Date : **31/05/2018** **S-2018-0959** Max. Marks: 60

N.B.:

1) Answer **ANY SIX** questions.

2) Figures to the right indicate **FULL** marks.

3) Draw neat and labelled diagram **WHEREVER** necessary.

Q. 1 **a)** Explain data warehouse in detail. **(06)**

b) Explain role of Business Intelligence in Data Warehouse. **(04)**

Q. 2 **a)** Explain different operations that can be performed in OLAP with suitable examples. **(06)**

b) Explain the terms Staging Area, Data Marts and Cubes. **(04)**

Q. 3 The college wants to record the marks for the courses completed by students using the dimensions: **(10)**

i) Course

ii) Student

iii) Time and a measure Aggregate marks

Create a cube and describe the following OLAP operations:

i) Slice

ii) Dice

iii) Roll Up

iv) Drill Down

v) Pivot

Q. 4 **a)** With respect to Association Rule mining define: **(04)**

i) Support

ii) Confidence

b) A database has five transactions let min_sup = 60 % and min_*conf* = 80% **(06)**

Trans ID	Items Bought
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

Find all frequent item sets using Apriori algorithm and generated association rules along with their values for support and confidence.

P. T. O.

Q. 5 What is data mining? Briefly explain the knowledge discovery process? **(10)**

Q. 6 What is dimensional modeling? What are the different types of data modeling? Explain each one of them with an example. **(10)**

Q. 7 Suppose that a data warehouse for Big University consists of the following four dimensions: Student, Course, Semester and Instructor and two measures count and average grade. When at the lowest conceptual level (e.g. for a given student course, semester and instructor combination) the average grade measure stores the actual course grade of the student. At higher conceptual levels, average grade stores the average grade for the given combination. **(10)**

a) Draw a snowflake schema diagram for the data warehouse.

b) Starting with the base cuboid (student; course; semester; instructor), what specific OLAP operations (e.g., roll-up from semester to year) should one perform in order to list the average grade of CS courses for each Big University Student.

Q.8 Write short notes on: **(10)**

a) Fact Constellation

b) Data Mart

* * * * *