Total no. of pages: 02 Roll No.....

EIGHT SEMESTER B. Tech. [CO]

MID TERM EXAMINATION March- 2019

CO404 DATA WAREHOUSING & DATA MINING

Time: 01:30 Hours **Max. Marks: 30**

NOTE: Attempt any 4 questions.

Que 1. (a) Define Data Warehouse and describe key features of data warehouse.
(b) Define Data Marts. Compare and contrast Top-Down and Bottom-Up approach of Data Marts design. (3.5+4)

Que 2. (a) Compare and Contrast Type 2 and Type 3 slowly Changing dimensions. Justify with example.
(b) What is STAR Schema? You are the data design specialist on the data warehouse project team for a manufacturing company. Design a STAR schema to track the production quantities. Production quantities are normally analyzed along the business dimensions of product, time, parts used, production facility, and production run. State your assumptions. (3.5+4)

Que 3. (a) Explain any four of Dr. Codd's initial guidelines for OLAP. Give reasons why the selected for are significant for OLAP.
(b) What data does an information package contains? Explain Business metrics and Facts with five examples. Design IPD for problem stated in question 2(b). (3.5+4)

Ques 4. Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19,20, 21,22, 22, 25, 25, 25, 25, 30, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. What is the mean, median and mode of the data?

(b) Briefly describe nominal attributes, asymmetric binary attributes and Jaccard Coefficient with example. (3.5+4)

P.T.O.

Que 5. Consider the following transaction table:

Transaction ID	Items Bought
T001	{M, O, N, K, E, Y}
T002	{D, O, N, K, E, Y}
T003	{M, A, K ,E}
T004	{M, U, C, K, Y}
T005	{C, O, O, K, I ,E}

Write Apriori Algorithm. Compute all frequent itemsets, and find all strong association rules. Assuming min_sup=60% and min_conf=80%. (7.5)

END