Seat No.: Enrolment No			-
		GUJARAT TECHNOLOGICAL UNIVERSITY	
a 1 •	4	M.E -II st SEMESTER-EXAMINATION - JULY- 2012	10
Subject code: 1722304 Date: 10/07/2			12
•	Subject Name: Data warehouse and data mining Time: 10:30 am – 13:00 pm Total Marks		
		•	70
Instr			
		empt all questions. Re suitable assumptions wherever necessary.	
		res to the right indicate full marks.	
Q.1	(a)	Differentiate between star schema, snowflake schema and fact constellation	07
		with the help of examples	
	(b)	Show a three-tier data warehousing architecture, with the help of a suitable diagram.	07
Q.2	(a)	Distinguish between Data warehouse and OLTP systems with the help of suitable diagram	07
	(b)	How information can be extracted from data warehouse? Discuss any one technique with example	07
	(b)	What do you mean by schema in data warehouse? List all the schema modeling techniques for data warehouse schemas. Explain any one from those listed techniques	07
Q.3	(a)	Explain Data Cleaning, Data Integration & Transformation.	07
Q.D	(b)	What is clustering? Explain with example OR	07
Q.3	(a)	Explain sampling method for data reduction.	07
	(b)	Explain ID3 Algorithm for decision trees.	07
Q.4	(a)	What is the 'Apriori property'? How is it used by the APRIORI algorithm? What are the drawbacks of the Apriori algorithm? Discuss with suitable example.	07
	(b)	Describe OLAP operations in the multidimensional data models. OR	07
Q.4	(a)	Discuss the steps involved in data Processing.	07
	(b)	How is CLARANS different from CLARA? Illustrate this using a small example.?	07
Q.5	(a)	What is the main principle in partitioning methods of clustering? Explain any	07
0.5	(1)	one with suitable example	0 -
Q.5	(b)	Discuss naïve Bayesian classification. Why is it called "naïve"? OR	07
Q.5		Write short notes on any two from following:-	(07

(a) Mining Association Rules.(b) BIRCH

(c) Data mining application and trends.

X2)