

Reg. No.									
----------	--	--	--	--	--	--	--	--	--



Manipal Institute of Technology, Manipal

(A Constituent Institute of Manipal University)



VII SEMESTER B.TECH (COMPUTER SCIENCE AND ENGINEERING)

END SEMESTER EXAMINATION, NOV/DEC 2015

SUBJECT: BUSINESS INTELLIGENCE AND ITS APPLICATIONS

[CSE 437]

REVISED CREDIT SYSTEM

Time: 3 Hours

26-11-2015

MAX. MARKS: 50

Instructions to Candidates:

- ❖ Answer **ANY FIVE FULL** questions.
- ❖ Missing data, if any, may be suitably assumed.

1A. Define business intelligence. Explain the process of turning data into actionable plan.

1B. Define semi structured data. Explain possible solutions for extracting information from stored unstructured data and semi structured data.

1C. Describe any six OLAP operations on multidimensional data. Give example for each.

(2+5+3)

2A. Give the William H. Inmon definition of data warehouse. Explain key terms in the definition.

2B. Why data quality matters? Explain with four examples.

2C. Give the difference between ER modelling and Dimensional modelling.

(3+4+3)

3A. Draw the ER diagram for the following: You set up a database company, Art-Base that builds a product for art galleries. The core of this product is a database with a schema that captures all the information that galleries need to maintain.

Galleries keep information about artists, their names (which are unique), birthplaces, age, and style of art. For each piece of artwork, the artist, the year it was made, its unique title, its type of art (e.g., painting, lithograph, sculpture, photograph), and its price must be stored. Pieces of artwork are also classified into groups of various kinds, for example, portraits, still life's, works by Picasso, or works of the 19th century; a given piece may belong to more than one group. Each group is identified by a unique name (like those just given) that describes the group. Finally, galleries keep information about customers. For each customer, galleries keep that person's unique name, address, total amount of dollars spent in the gallery (very important!), and the artists and groups of art that the customer tends to like.

3B. Is the OLTP database design optimal for a data warehouse? Justify.

3C. List and explain the characteristics of a good business metrics in KPI. Medicines in Gulf countries are 100% costlier compared to Asia pacific region which is 50%. Identify the metric components.

(5+2+3)

4A. The Company you work for wants to digitize their time cards. You have been asked to design the database for submitting and approving time cards.

- *timecard* should have hours worked and date submitted
 - Each *timecard* is associated with exactly one *employee*
 - Each *timecard* should have a unique id
 - Each *timecard* has a status: it is either approved, not approved, or pending
 - Each *employee* has a unique id
 - Each *manager* has a unique id and a name
 - Each *manager* is in charge of multiple employees
 - Each *manager* approves time cards for multiple employees
- i) Draw the conceptual data model for the above problem.
 - ii) Draw the logical data model for the above problem.
 - iii) Draw the physical data model for the above problem.
 - iv) Convert the logical data model to star schema.
 - v) Identify the RCD in above problem and tackle it in star schema.

4B. Suppose that a data warehouse for Big University consists of the following four dimensions: **student**(student_id, student_name, address, major, status, university), **course**(course_id, course_name, department), **semester**(semester_id, semester, year) and **instructor**(instructor_id, dept, rank), and two measures **count** and **avg_grade**. When at the lowest conceptual level (e.g., for a given student, course, semester, and instructor combination), the **avg_grade** measure stores the actual course grade of the student. At higher conceptual levels, **avg_grade** stores the average grade for the given combination.

- i) Draw a snowflake schema diagram for the data warehouse.
- ii) Starting with the base cuboid [**student, course, semester, instructor**], what specific OLAP operations should one perform in order to list the average grade of CS courses for each Big University student.
- iii) If each dimension has five levels (including **all**), such as “**student<major<status<university<all**”, how many cuboids will this cube contain (including the base and apex cuboids)?
- iv) Identify SCD and explain how to tackle them.
- v) Identify the different types of fact in snowflake schema. (5+5)

5A. Explain i) Degenerate dimension ii) Junk dimension

5B. Explain all the common perspectives of reporting that apply at all levels of the enterprise.

5C. Faculty want to improve the student performance, write the balanced scorecard for the above requirement. (2+4+4)

6A. What is a dashboard? List the attributes of dashboard.

6B. What are the visible components of designer window when working on a SSIS project? Explain each of them.

6C. Write a note on the following :-

- i) List report
- ii) Matrix report
- iii) Gauge report (2+5+3)