Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
M.S. Systems Engineering at Wipro InfoTech (WIMS)
First Semester 2014 - 2015 (October 2014 to March 2015)
Mid semester Exam (Regular)
Answer Key

Course Number	:	SEWP ZG514
Course Title	:	Data Warehousing
Type of Exam	:	Closed Book
Weightage	:	30%
Duration	:	90 minutes
Date of Exam	:	21st December 2014

No. of Pages : 1
No. of Questions : 5

Session : AN

Note:
1. Please read and follow all the instructions given on the cover page of the answer script.
2. Start each answer from a fresh page. All parts of a question should be answered consecutively.
3. Appropriate assumptions can be made and explicitly mentioned .

---

1. Short Questions:- 5 * 2 = 10

   a. How do you characterize Data Warehouse?
   Inmon's definition or equivalent explanation is acceptable.

   b. Explain role-playing dimension with example.
   Role playing dimension is a single physical dimension table that plays different logical roles in a dimension model. (1 mark for definition + 1 mark for example)

   c. What is meant by grain of the business process?
   Grain indicates level of data detail in a dimensional model. The granularity of fact table constrains the level of analysis. A retail store recording daily sale of soaps can analyze product performance, but not which customer prefers which soap.

   d. Define main OLAP operations.
   Roll-up(drill-up)-Aggregation on a data cube, by climbing up a concept hierarchy for a dimension or by dimension reduction.
   Drill-down- Reverse of roll-up operation. navigate from less detailed to more detailed data.
   Slice- Performs a selection on one dimension of the given cube, resulting in a sub-cube.
   Dice- Define a sub-cube by performing a selection on two or more dimensions.
   Pivot(rotate)- visualization operation that rotates the data axes in a view ,in order to provide an alternative presentation of data.

   e. Consider manufacturing and insurance businesses. For each business, suggest a subject for data warehousing, a measure for fact table, and a dimension (other than time).
   Example can be inventory and claims, stock units and amount, warehouse location and claimant.

2. What are architectural components of data warehouse? Explain load manager and its operations. (2+3 marks)

Load Manager, Warehouse Manager, and Query Manager. (2 marks for definitions))

The load manager is the system component that performs all the operations necessary to support the extract and load process. It may be constructed using a combination of bespoke coding, Third-party tools and scripts. Generally load manager has an exclusive staging server where data operations are performed. The staging server is usually not accessible for non-
IT

Operations of load manager include

- Extracts the data from the source systems, e.g. FTP across the LAN
- Fast-load the extracted data into a temporary data store, likely to load the data into a relational database prior to applying transformations and checks.
- Perform simple transformations into a structure similar to the one in the data warehouse.

3. Define Star and Snowflake schemas. List relative advantages of each. (2+3 marks)

Star schema organizes the DW tables into facts and corresponding dimensions such that each dimension is a completely denormalized table and each dimension has foreign key relation with the fact table. In case of snowflake, the dimension entity is normalized. (2 marks)

Advantages of Star Schema are :

Easy for Users to Understand
Optimizes Navigation
Most Suitable for Query Processing
Can implement STARjoin and STARindex for performance enhancement

Advantages of snowflake schema are :

Small saving in storage space
Normalized structure are easier to update and maintain.

4. A major hospital wants to analyse revenues contributed from each doctor. They also need to report major health issues in the area to local authorities. Design a DW that will address their requirements. (5 marks).

Identification of appropriate measures in fact table – 1 mark
Identification of dimensions(Disease, Patient with demographics, Time, Doctor) – 2 marks
Revenues can be aggregated on Doctor – 1 mark
Diseases in the locality identified by patient demographic – 1 mark

5. A state government is keen on understanding tourism profile, such as tourist profile, tourist spending, seasonal volumes, at their tourist destinations. Propose a detailed DW design for their requirements.

Identification of appropriate measures in fact table – 1 mark
Identification of dimensions – 2 marks
Tourist dimension to include demographic, Time dimension to include season -2 mark