Assignment on Data Warehouse (DW)
CSE512, SP 2023

Analyze and Design an e-commerce data warehouse for a nationwide chain of superstores of Bangladesh. There are many suppliers for the chain shop. The suppliers has sup-id, name, type of products to supply (clothes, machineries, food etc.), address (street, city, district). The chain shops serve the customers physically from the superstores and each superstore has its own system for all kinds of transactions (sale, procurement from the suppliers etc.). There are 10 million registered customers and each customer has customer id, name, NID, address (House no., street, thana, city, district, division and age group. A customer can purchase many items in a single transaction with transaction id, transaction type (cash or card), timestamp id, time of the day, day of the week, date, week, month, year, quantity, unit price and total price. Each item has an item id, name, type, country of manufacture. You have to design and implement a warehouse for this chain of superstores to support policy decision process and knowledge discovery. Perform the following tasks for the warehouse.

Analysis

1. Analyze the operational database in different sources and select the appropriate entities and attributes for uploading to the data warehouse.
2. Analyze the activities required in the source sites to design the wrappers for data integration, data extraction, pre-processing and upload to DW.

Submission: Report in pdf format to google classroom **Deadline** **Apr 20, 2023**

Design

Task 1: Design the architecture of the warehouse and explain the sources, preprocessing, noise reduction, transformation and uploading.

Task 2: Design the star schema for the warehouse **using the scenario and the data set** and explain how the data of the superstore database will be collected to the DW (source driven or destination driven).

Task 3: Prepare a mapping of your design with the DW given in VIS and propose any change (if required).

Submission: Report in pdf format to google classroom. **Deadline Apr 20, 2023**

Task 3: Perform Task 1-3 in VIS

Task 4: The following are the list of analytics:

1. store and time dimensional financial analytics,
2. customer and time dimensional financial analytics,
3. item and time dimensional financial analytics,
4. store and time dimensional inventory analytics,
5. item and time dimensional inventory analytics.

Every student must perform one of the above analytics. The number will selected as follows:

((Left most 7 digit of your ID) mod 5) + 1

Submission: May 5 2023 in VIS