Project on Data Warehouse

CSE 411, Fall 2020

Full marks: 20

Submission Deadline: 26th January, 2021

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 serves 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 agegroup. 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. The following reports will be generated from the warehouse.

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 and explain how the data of the superstore database will be collected to the DW (source driven or destination driven).

Task 3: Write three different cross tabulation for three different dimensions using total price/ quantity. Write SQL to find the cross-tabs.

Task 4: Show the structure of a 4 dimensionai data cube and explain.

Submission instruction:

1. Prepare a video on the assignment and duration must be 3-4 minutes presenting the tasks 1 – 4 and upload to the following drive location:

<https://drive.google.com/drive/folders/1mQ6KBLe0i4Ij5W13rvZ9r5cWGk9JZFnK>

**Deadline for upload is, 2:00PM, 26th January, 2021.**

**Presentation:** During class period on 26th January, 2021, you have to run and share the video. You may be asked question about your work immediately after the video presentation.