BUS 464 Business Data Management
Homework - Data Warehousing - I

Instructor: Nilesh Saraf

Fall 2015

This assignment will expose you to the SQL code that could be used to implement a Star-Schema based on the Sakila database. Note that the Sakila database is normalized and can be considered to be a relational DB underlying a Transactional Processing System (TPS). Complete the following steps:

1. A sandbox SQL code for a sandbox datawarehouse (DW) is provided to you (see file sakila_dwh.sql). Sketch the star schema implemented by it by examining the dimension and fact tables created. You can run it within your DBworkspace (db_your sfu id). Feel free to modify the code to refer to your database account and run it. Study how the surrogate keys are created when creating the fact and dimension tables (feel free to refer to online sources to understand the role of surrogate keys in a DW). Comment on the similarities between the contents of the DW with the actual Sakila DB by runnning a few SQL queries.

2. You will now create a smaller DW with the following dimensions (besides a time dimension that every DW needs to have.): store, film, staff. The *fact* would remain the same: rental count and rental duration (average). Write SQL so that it takes all the data needed from the Sakila DB version on the MySQL server. Thus, there should *not* be any INSERT statements to load fresh data into the DW.

3. Modify your SQL code to implement a slightly different star schema with two dimensions: store & actor, and fact: payment.

4. List 2 verification questions and 2 discovery questions that are most interesting for you, and can be approximately answered using the data in the Sakila DB – especially using any of the above star schemas.

Notes: Suffix all dimension tables with 'dim' and the fact table with 'fct' for ease of reading. Do not submit any data with the submission but only the SQL code devoid of INSERT statements. I should be able to create a replica of your DW simply by copying/pasting all your SQL code.