Data Warehousing project – Graduate Admission

Instruction Guide

Date: 04/26/2017

Group - 5

Instruction Guide to run the project

Instructions to run the ETL job:

Step 1. Run the otl_admission_dump.sql file. This creates the required oltp database along with the tables and its respective data populated.

Step 2. The following databases have been used for this project:

- oltp_admission
- source_db
- staging_db
- datamart_db

Step 2. Run the Source_DB_DDL.sql file. This is used to replicate the data structure into the ETL system creating the required tables.

Step 3. Run the Staging_DB_DDL.sql file. This file creates the necessary tables in the staging database in the ETL Process.

Step 4. Run the Datamart_DB_DDL.sql file. This file creates the necessary dimension and fact tables in the target/datamart database in the ETL Process.

Step 5. Open the main job: main-datamart-application.kjb in pentaho.

Step 6. IMPORTANT: Open fact-admission.ktr, GOTO dimdate1, and clear the cache for dim_date. Double click on a dimdate1. Click on browse for lookup table, and select dim_date. Next go to action -> ddl -> use current connection-> clear cache. This step is required because, we have changed the datatype for the_date column since the date fields in our model were of character type.

Step 7. Run the main job.