Design and Implementation of a Data Warehouse for a Retail Store

Report 2: Logical Design and Physical Design of the Data Warehouse Schema
Absolutely no discussions between project teams about the entire project.

Overall Objective of this Project

The objective of this project is to design and develop a data warehouse for a retail store chain called Dominick's Fine Food (DFF) that has many branches. We use their store level data to set up the data warehouse and help them analyze the data.

Project Scenario

The data for this project includes store-level data covering shelf management and pricing from all the branches of DFF.   The data can be accessed from the following web site: https://www.chicagobooth.edu/research/kilts/datasets/dominicks

Overall Objective of this Report

Report 2 will focus on logical design and physical design of the warehouse structure <u>based on</u> the business questions that have been selected for your group. The tasks involved are:

- Remember that you are given the architecture for this project. You need to follow Dimensional modeling scheme, Independent data marts style, HOLAP style for the warehouse and SQL Server 2016 as the infrastructure for implementation.
- Use Dimensional Modeling technique to create the star schemas (could create snowflake schemas if needed). Look at Kimball's methodology of creating independent conformable data marts that has been discussed in class. (Hint: Stay away from snowflake, if possible.) Make sure to use Kimball's process of data mart design in total.
- Develop the data warehouse schema using Visio, Lucid Chart or any other drawing tool (target DB is SQL Server 2016).
- Provide a Mapping table#1 from SOURCE FILES TO STAGING TABLES that includes {source file name, source_file_attribute, mapping, staging_table_type (relation), staging_table_name, staging_table_attribute}. Remember that mapping function could be transformation functions and/or copy, replicate etc. functions
- Provide another Mapping table#2 from STAGING TABLES TO PRESENTATION SERVER TALBES that includes {staging_table, staging_table_attribute, mapping, Data Mart_table_type (fact table/dimension table), Data Mart_table_name, Data Mart_table_attribute}. Remember that mapping function here could be copy plus other functions.
- Make sure that your BQs (chosen for you from Report 1) correspond to the structure of independent data marts. Justify each schema with logic and also with data to explain how it supports the BQs.
- Develop a brief physical design plan by describing data aggregate plan, indexing plan, data standardization plan, and storage plan (remember that your data marts will grow in future for DFF). Keep this up to two paragraphs.
- Submit a report that shows all these tasks (i) – (vi) that you have done.

Submit

A report showing all these tasks that you have done. If needed, put the screen shots in the report. Visio outputs are also required.

Section 1.   **Introduction**
- Talk about Logical design for DFF

Section 2.   **Overview of Kimball's Methodology**
- Describe each step of the Methodology
- Why is it important to follow this methodology to create independent data marts.

Section 3.   **DW Logical Design (Star Schema Design)**
- Use Dimensional Modeling technique to create the star schemas (could create snowflake schemas if needed). Look at Kimball's methodology of creating independent conformable data marts that has been discussed in class. (Hint: Stay away from snowflake, if possible.) Make sure to use Kimball's process of data mart design in total.
- Develop the data warehouse schema using Visio, Lucid chart, or any other drawing tool (target DB is SQL Server 2016).
- Make sure that your BQs (chosen for you from Report 1) correspond to the structure of independent data marts. Justify each schema with logic and also with data to explain how it supports the BQs.

Section 4.   **Develop Two Mapping Tables (Sources-to-Data Marts)**
- Provide a Mapping table#1 from SOURCE FILES TO STAGING TABLES that includes {source file name, source_file_attribute, mapping, staging_table_type (relation), staging_table_name, staging_table_attribute}. Remember that mapping function could be transformation functions and/or copy, replicate etc. functions.

- Provide another Mapping table#2 from STAGING TABLES TO PRESENTATION SERVER TALBES that includes {staging_table, staging_table_attribute, mapping, Data Mart_table_type (fact table/dimension table), Data Mart_table_name, Data Mart_table_attribute}. Remember that mapping function here could be copy plus other functions. See a typical example below:

Section 5.   **Physical Design**
- Develop a brief physical design plan by describing data aggregate plan, indexing plan, data standardization plan, and storage plan (remember that your data marts will grow in future for DFF). Keep this up to two paragraphs.

Grading
- Quality of the presentation   20%
- Quality of the content logical design   70%
  - **MUST follow the steps given by Kimball Methodology to create the Dimension models for all data marts**
- Quality of the content physical design   10%