Lab 2b: Designing a Data Warehouse

I. Setting

In this lab, you are going to design a simple but optimized star schema for a retail business of your choice (e.g., video rental stores). To do so, you will have to follow the four main steps for designing a DW, we described during the lectures:

- Identify the business process
- Identify the grain
- Choose the dimensions
- Choose the measures

The business process should be defined according to these two rules:

- **High impact:** Keep in mind that the outcome is a report that will support the decision-making to increase the profit of the company.
- **Low risk:** Make sure that the data are relatively easy to collect and store.

The grain is the most important part of the design. Here are some tips to define a good granularity:

- Define the fact table (center of the star): What does one fact represent? "For instance, it could be a line item in a single transaction, such as: one row represents a movie rented by a customer from an employee in a store on a day."
- Choose the most atomic or granular level that you have with your data (i.e., data that cannot be break it down anymore). Such granular level allows data to be rolled up or down, aggregated, sliced, and diced. It is important, because we have to be prepared for any kind of query/question the end-user could request.

II. Implementation

Now use your favorite database diagram tool (e.g., SQL Server Management Studio) and start modeling your DW.

1) Build your fact table

- Identify the relevant dimensions that are going to be related to this table and specify their type.

- For instance, it could be: Product, Date, Customer, Store (or Internet), and Employee.

2) Build the dimension tables

- A dimension contains descriptive information regarding our data, such as: who, what, where, and when.

- Specify the attributes that define each dimension table. For instance, for DimProduct, it could be: ID, Name, Type, Price…

3) Define the measures

- Edit the fact table to add attributes that measure the performance of your business, including sell amount and costs.

4) Checking

To make sure that your DW is correctly implemented, verify that the design allows you to answer the questions as follows:

- How many sales did the company have?
- How many products did we sale in a store last month?
- How many products are bought by our top 10% of customers?
- Who are our top 10% best sellers/employee last year?
- Which one of our stores have the highest amount of sells?
- Which products are the most lucrative?
- What was the most lucrative day, month, or year?
- …

III. Deliverable

Write a short report, in which you:

1) Explain the business you chose to model through a DW.

2) Present and describe the ER diagram implemented as solution to assess the business performance of the imaginary company.

3) Answer the question as follows: Which transformation need to be applied to the DW, if the company is interested in knowing the effect of promotions on sales?

4) Draw conclusions of your work, in terms of benefits and limitations of your model.