Data Warehousing Assignment—Part I

Toon Calders   Hatem Haddad

Deadline: 07 November 2016

1   Practical information

Deadline: 07 November 2016
Group size: Up to three students
How to submit: Upload solution at `uv.ulb.ac.be`

2   Objectives

The goals of this assignment are:

1. Create a dimensional fact model;

2. Translate it to the relational model; take into account potential changes that may occur in the dimensions.

3   Database Description

A DVD rental company maintains a database with the current information about its rentals. For every customer, personal information is stored (first name, last name). The address table contains address information for customers, staff and stores. The film table is a list of all films potentially in stock. The inventory table is about the actual in-stock copies of each film. A film can have one or more categories. The payment table records each payment made by a customer with information about the payment date and the amount. The rental table is about rentals with information about who rented what item, when it was rented, and when it was returned. The attribute last_update in each table represents the date and time that the row was created or most recently updated. Stores have multiple inventories. A rental has zero or one payment. Figure 1 shows the database tables used by the DVD rental company. A copy of this database is available on the course website under the name **dvdrental**.

4   Problem Description

The company would like to keep track of its rentals and how they evolve over time. Therefore it is decided to extract the rental information at regular intervals from the DVD rentals database and to store it into a small data warehouse to allow for historical analysis of the rental evolution. The data warehouse should allow for storing the historical data over a long time and should anticipate potential changes in the database. Based on the data stored in the data warehouse it should be possible to answer analysis queries such as the following:

1. Give the average rental amount per category per year;

2. Give the average rental amount per city per country;

3. Give the monthly average rental per category per shop per country;

4. Give the total and the average rental amount, per category for films released since 1990 per city and per country.

The above queries are examples to illustrate potential use of the datawarehouse only. All information in the OLTP database should go into the OLAP data warehouse. Throughout the lifetime of the data warehouse there may be changes to the following tables: payment, customer, staff and store. Accordingly, your model should be such that these changes can be handled correctly in the data warehouse. For the other tables there may be inserts, but the existing tuples are never updated or deleted.
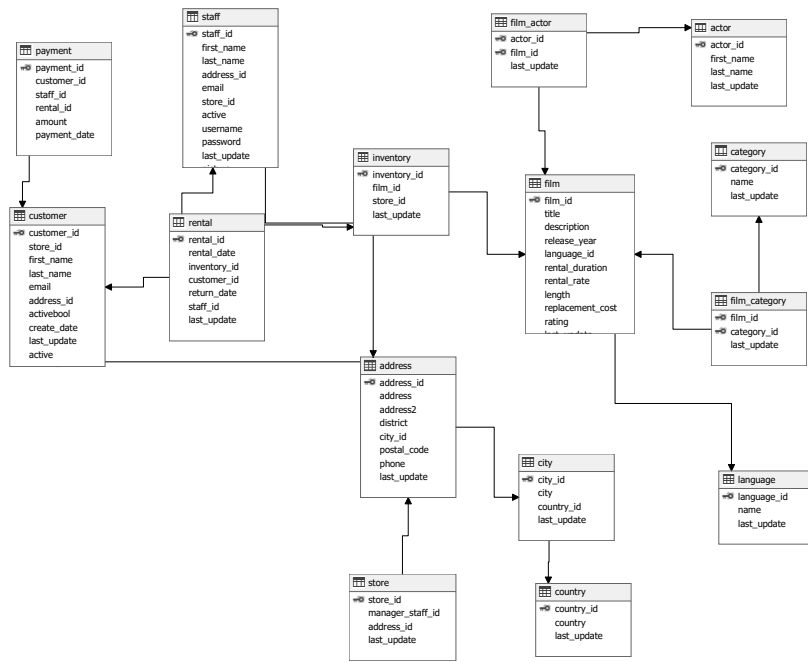
Figure 1: Database tables of the Assignment1 database.

5   Deliverables

You should deliver a report (as a .pdf), containing (length indication is purely indicative):

1. A **cover page** with the list of group members, including student ID;

2. Your **Dimensional Fact Model** with a short explanation (total: 1p to 1.5p);

3. Your **Relational Model** with a short explanation where necessary. The model has to reflect potential changes in the dimension tables (total: 1.5p);