Manage a chain of Movie Rental Stores

Introduction

In this project you will write more advanced queries on a database designed to resemble a real-world database system - MySQL's Sakila Sample Database.

Development of the Sakila sample database began in early 2005. Early designs were based on the database used in the Dell whitepaper (Three Approaches to MySQL Applications on Dell PowerEdge Servers).

The Sakila sample database is designed to represent a DVD rental store. The Sakila sample database still borrows film and actor names from the Dell sample database.

Problem Description

You're writing SQL to manage a chain of movie rental stores, for example,

- Track the inventory level and determine whether the rental can happen
- Manage customer information and identify loyalty customers
- Monitor customers' owing balance and find overdue DVDs

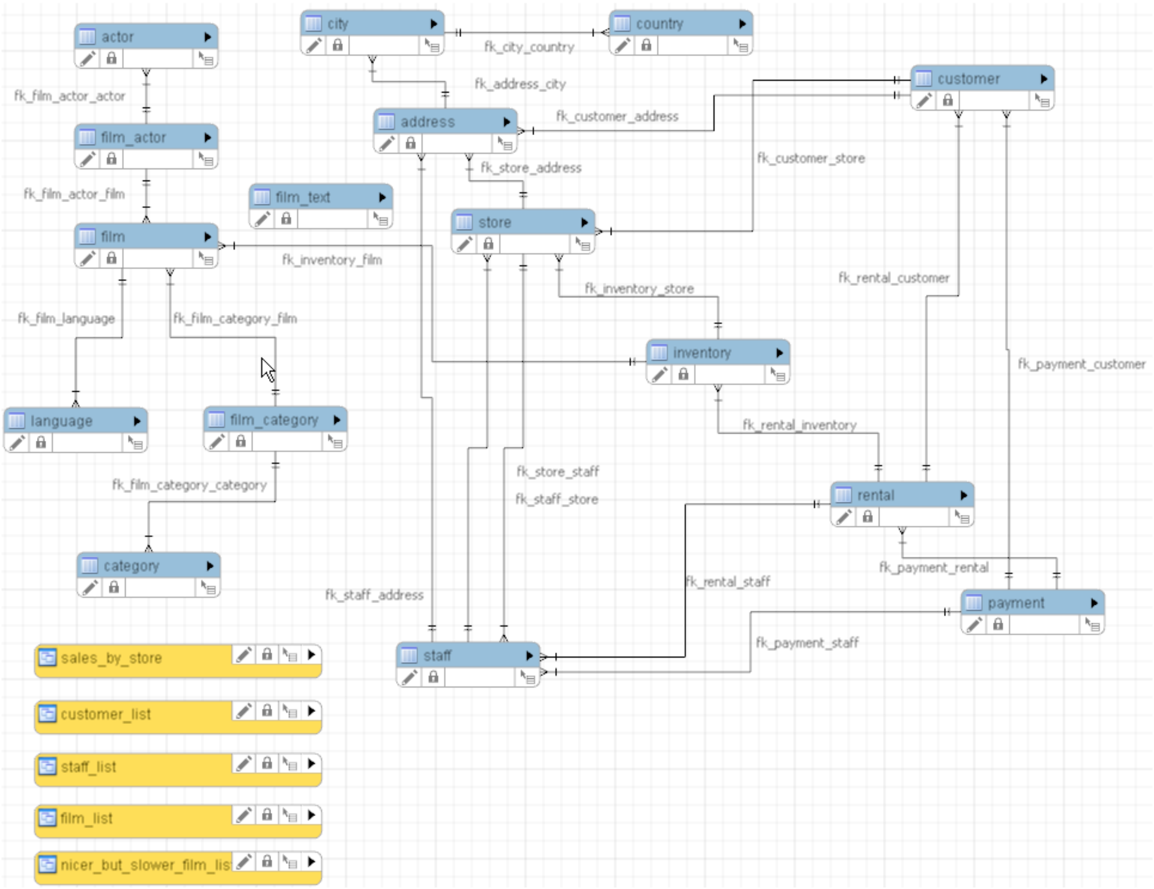
This project can be considered as a typical retail-related business case, because it has the main metrics you can find in any retailer's real database, such Walmart, Shoppers, Loblaws, Amazon…

Key Metrics:

- ➢ **Production information (in this project, it is the film)**
- ➢ **Sales information**
- ➢ **Inventory information**
- ➢ **Customer behavior information**

Data Structure

The following diagram provides an overview of the structure of the Sakila sample database.



Exercise 1 (for Day1-Day3):

1. Before doing any exercise, you should explore the data first. For Exercise 1, we will focus on the product, which is the film (DVD) in this project. Please explore the product-related tables (actor, film_actor, film, language, film_category, category) by using 'SELECT *' – do not forget to limit the number of records

Use table FILM to solve questions as below:

2. What is the largest rental_rate for each rating?
3. How many films in each rating category?
4. Create a new column film_length to segment different films by length:
   length < 60 then 'short'; length < 120 then 'starndard'; lengh >=120 then 'long'
   , then count the number of files in each segment.

Use table ACTOR to solve questions as below:

5. Which actors have the last name 'Johansson'
6. How many distinct actors' last names are there?

7. Which last names are not repeated? *Hint: use COUNT() and GROUP BY and HAVING*
8. Which last names appear more than once?

Use table FILM_ACTOR to solve questions as below:

9. Count the number of actors in each film, order the result by the number of actors with descending order
10. How many films each actor played in?

Exercise 2 (for Day4):

1. Before doing any exercise, you should explore the data first. For Exercise 1, we will focus on the product, which is the film (DVD) in this project. Please explore the product-related tables (actor, film_actor, film, language, film_category, category) by using ‘SELECT *’ – do not forget to limit the number of records;
2. Find language name for each film by using table Film and Language;
3. In table Film_actor, there are actor_id and film_id columns. I want to know the actor name for each actor_id, and film tile for each film_id. Hint: Use multiple table Inner Join
4. In table Film, there are no category information. I want to know which category each film belongs to. Hint: use table film_category to find the category id for each film and then use table category to get category name
5. Select films with rental_rate > 2 and then combine the results with films with rating G, PG-13 or PG

Exercise 3:

Let's look at sales first:

- The rental table contains one row for each rental of each inventory item with information about who rented what item, when it was rented, and when it was returned
- The rental table refers to the inventory, customer, and staff tables and is referred to by the payment table
- Rental_id: A surrogate primary key that uniquely identifies the rental

1. How many rentals (basically, the sales volume) happened from 2005-05 to 2005-08? Hint: use date between '2005-05-01' and '2005-08-31';
2. I want to see the rental volume by month. Hint: you need to use substring function to create a month column, e.g.
3. Rank the staff by total rental volumes for all time period. I need the staff's names, so you have to join with staff table

How about inventory?

4. Create the current inventory level report for each film in each store?
   - The inventory table has the inventory information for each film at each store
     - inventory_id - A surrogate primary key used to uniquely identify each item in inventory, so each inventory id means each available film.

5. When you show the inventory level to your manager, you manager definitely wants to know the film name. Please add film name for the inventory report.
   - Tile column in film table is the film name
   - Should you use left join or inner join? – this depends on how you want to present your result to your manager, so there is no right or wrong answer
   - Which table should be your base table if you want to use left join?

6. After you show the inventory level again to your manager, you manager still wants to know the category for each film. Please add the category for the inventory report.

- Name column in category table is the category name
- You need to join film, category, inventory, and film_category

7. Your manager is happy now, but you need to save the query result to a table, just in case your manager wants to check again, and you may need the table to do some analysis in the future
   - Use CREATE statement to create a table called as inventory_rep

8. Use your report to identify the film which is not available in any store, and the next step will be to notice the supply chain team to add the film into the store

Let's look at Revenue:

- The payment table records each payment made by a customer, with information such as the amount and the rental being paid for. Let us consider the payment amount as revenue and ignore the receivable revenue part
- rental_id: The rental that the payment is being applied to. This is optional because some payments are for outstanding fees and may not be directly related to a rental – which means it can be null;

9. How many revenues made from 2005-05 to 2005-08 by month?
10. How many revenues made from 2005-05 to 2005-08 by each store?
11. Say the movie rental store wants to offer unpopular movies for sale to free up shelf space for newer ones. Help the store to identify unpopular movies by counting the number of rental times for each film. Provide the film id, film name, category name so the store can also know which categories are not popular. Hint: count how many times each film was checked out and rank the result by ascending order.