Design of An Online Movie Streaming Database

CMPE353/CMSE354

Database Management Systems

Labs 2-7

Fall 2022-23

Task: You are asked to design a relational database according to the given schema diagram below using Oracle Live SQL software. The task has several subtasks and each of them are described below. Each task will be carried out during weekly lab hours (Labs 2-6). Therefore, each subtask must be completed weekly in accordance with the set lab dates (see dates below).

Problem: The aim is developing a database scheme which represents an example of a used dataset in an online streaming service such as Netflix. The schema diagram below is a minimized sample form of a real database that is used in an online streaming service. In these services users after logging into their account are provided with a list of movies. In this list usually the most basic information about each movie including the name, duration, casts, rate and the number of votes are shown to the users. These information are retrieved from database at the time of fetching of the webpage.

In our given design there is a main table with the name of "movie". In this table each movie is recognized with a five-digit length ID which starts with 1 (i.e. 1xxxx). Each movie also has a name, budget, popularity rate, revenue, and votes count. Budgets are in million dollar. Other informations of each movie including the language, the producer company, the country of origin and the cast are given in separate tables.

Information of the other tables are in the following format.

Language table:

Each language is determined with a five-digit unique ID starting with 2 (i.e. 2xxxx), a two-letter ISO code, and the complete name of the language.

Genre table:

In this table different genres are provided with a five-digit unique ID starting with 3 (i.e. 3xxxx).

Country table:

Each country is determined with a five-digit unique ID starting with 4 (i.e. 4xxxx), a two-letter ISO code, and the complete name of that country.

Gender table:

In this table females are shown with code 1 and males with code 2.

Person table:

In this table each person is provided with a five-digit unique ID starting with 5 (i.e. 5xxxx).

Production company:

In this table different companies are provided with a five digit unique ID starting with 6 (i.e. 6xxxx).

Following criteria are considered in the database design.

1. Each movie has a specific language; however, there may be many movies in the same language.
2. For the sake of simplicity, it is assumed that each movie is produced in a country; however, each country may produce many movies.

You can find other similar relations among the tables.
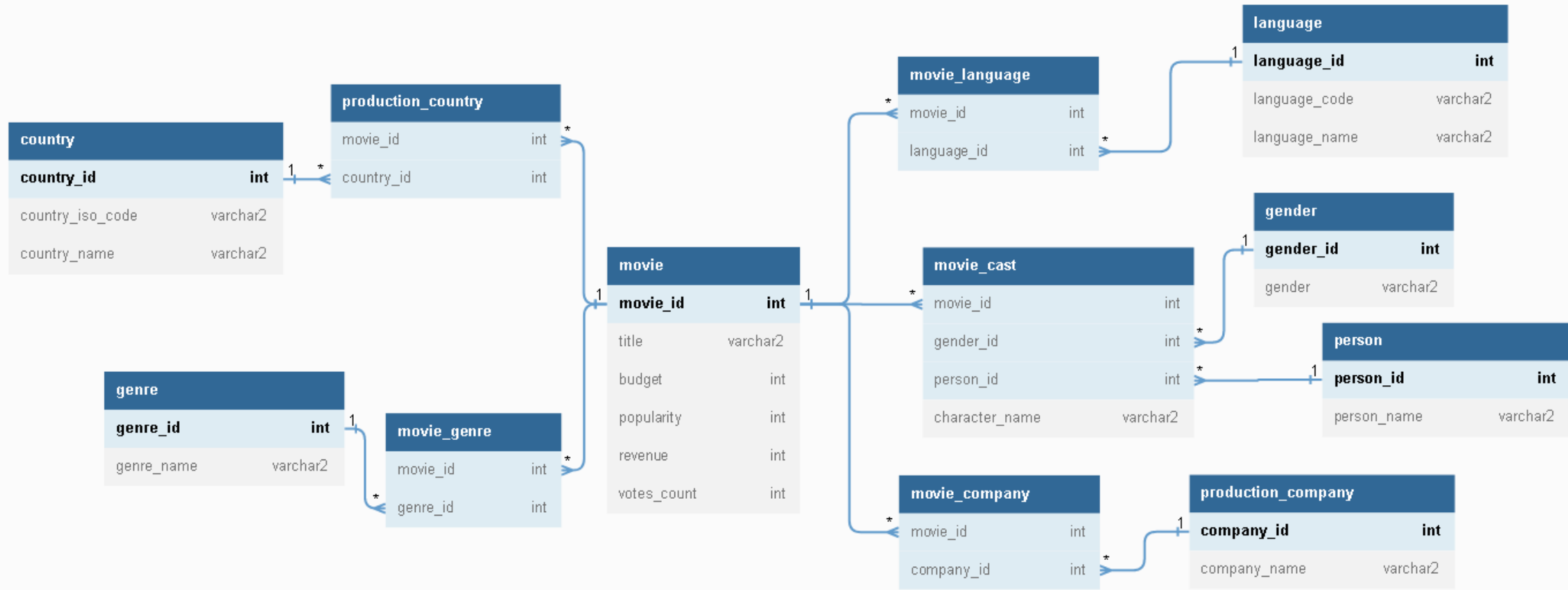
LAB 2 (7 Nov-11 Nov):
- SUBTASK 1: Fill in the following tables: (some are filled for your guidance)
  You have to fill these tables and submit your work during the lab.
- SUBTASK 2: Build the tables in Oracle Live SQL with your account according to the tables above.

movie

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id	√		varchar2(5)	NO	-	10001
title			varchar2(20)	NO	-	Don't Breath
budget			binary_float	YES		9.95
popularity			float(2)	YES		7.12
revenue			binary_float	YES		157.1
votes_count			int	YES		85

movie_language

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id		√	varchar2(5)	NO	-	10001
language_id		√	varchar2(5)	NO	-	20001

production_country

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id		√	varchar2(5)	NO	-	10001
country_id		√	varchar2(5)	NO	-	30001

movie_cast

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id						
gender_id						
person_id						
character_name						

movie_company

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id						
company_id						

Language

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
language_id						
language_code						
language_name						

movie_genre

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
movie_id						
genre_id						

country

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
country_id						
country_iso_code						
country_name						

production_company

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
company_id						
company_name						

genre

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
genre_id						
genre_name						

gender

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
gender_id						
gender						

person

Column Name	Primary Key	Foreign Key	Data Type	Nullable	Default	Sample Entry
person_id						
person_name						

LAB 3 (5 Dec-9 Dec):

- Insert sample data into the tables. (At least 10 records for each table)

1. How many movies are in the database?

2. Which movie does have the highest budge? Return the name of movie.

3. What is the total revenue of the available movies in the database?

4. Which genre does have the most number of movies? Specify the genre name.

LABs 4 and 5 (12 Dec-16 Dec / 19 Dec-23 Dec):

- Write an SQL query for the following questions:

1. In which language are most of the movies produced? Return the name of the language, **NOT** its ID.
2. Which country has produced the most number of movies? Return the **ID** and the **name of country** and the **number of movie** it has produced.
3. Which company has produced the most number of movies? Return the name of the company.
4. Which movie does have the lowest revenue? How many men and women are in the cast of this movie?
5. Find the number of movies each production company has produced. Return the name of company with the number of movies.
6. How many men and women are in the cast of the first movie in the database?
7. What is the genre of the most voted movie?
8. Which company did the most investment in the movie production? Hint: First find the summation of the budget of all of the movies produced by each company, then find the maximum value)

LAB 6 (26 Dec-30 Dec):

- Write a trigger to save all modification on table movie. (Update and Delete).

LAB 7 (2 Jan-6 Jan):

- Finalization and show of all work done throughout the semester.