WORKSHEET 5 SQL

# https://www.databasestar.com/wp-content/uploads/2019/11/movies_erd.pngRefer the following ERD and answer all the questions in this worksheet. You have to write the queries using MySQL for the required Operation.

**Table Explanations:**

* The **movie** table contains information about each movie. There are text descriptions such as title and overview. Some fields are more obvious than others: revenue (the amount of money the movie made), budget (the amount spent on creating the movie). Other fields are calculated based on data used to create the data source: popularity, votes\_avg, and votes\_count. The status indicates if the movie is Released, Rumoured, or in Post-Production.
* The **country** list contains a list of different countries, and the **movie\_country** table contains a record of which countries a movie was filmed in (because some movies are filmed in multiple countries). This is a standard many-to-many table, and you’ll find these in a lot of databases.
* The same concept applies to the **production\_company** table. There is a list of production companies and a many-to-many relationship with movies which is captured in the **movie\_company** table.
* The **languages** table has a list of languages, and the **movie\_languages** captures a list of languages in a movie. The difference with this structure is the addition of a **language\_role** table.
* This **language\_role** table contains two records: Original and Spoken. A movie can have an original language (e.g. English), but many Spoken languages. This is captured in the **movie\_languages** table along with a role.
* **Genres** define which category a movie fits into, such as Comedy or Horror. A movie can have multiple genres, which is why the **movie\_genres** table exists.
* The same concept applies to **keywords**, but there are a lot more keywords than genres. I’m not sure what qualifies as a keyword, but you can explore the data and take a look. Some examples as “paris”, “gunslinger”, or “saving the world”.
* The cast and crew section of the database is a little more complicated. Actors, actresses, and crew members are all people, playing different roles in a movie. Rather than have separate lists of names for crew and cast, this database contains a table called **person**, which has each person’s name.
* The **movie\_cast** table contains records of each person in a movie as a cast member. It has their character name, along with the **cast\_order**, which I believe indicates that lower numbers appear higher on the cast list.
* The **movie\_cast** table also links to the gender table, to indicate the gender of each character. The gender is linked to the **movie\_cast** table rather than the **person** table to cater for characters which may be a different gender than the person, or characters of unknown gender. This means that there is no gender table linked to the **person** table, but that’s because of the sample data.
* The **movie\_crew** table follows a similar concept and stores all crew members for all movies. Each crew member has a job, which is part of a **department** (e.g. Camera).

# QUESTIONS:

* 1. Write SQL query to show all the data in the Movie table.

select \* from movie;

* 1. Write SQL query to show the title of the longest runtime movie.

select title from movie order by runtime desc limit 1;

* 1. Write SQL query to show the highest revenue generating movie title.

select title from movie order by revenue desc limit 1;

* 1. Write SQL query to show the movie title with maximum value of revenue/budget.

select title,budget from movie order by budget desc limit 1;

* 1. Write a SQL query to show the movie title and its cast details like name of the person, gender, character name, cast order.

select movie.title,movie\_cast.person\_id, gender.gender, movie\_cast.character\_name, movie\_cast.cast\_order from movie join movie\_cast on movie.movie\_id=movie\_cast.movie\_id join gender on movie\_cast.gender\_id=gender.gender\_id

* 1. Write a SQL query to show the country name where maximum number of movies has been produced, along with the number of movies produced.

select country\_name,count(country\_name) as no\_of\_movies\_produced from movie as a join production\_country as c on a.movie\_id=c.movie\_id join country on c.country\_id=country.country\_id group by country\_name order by count(country\_name) desc;

* 1. Write a SQL query to show all the genre\_id in one column and genre\_name in second column.

select \* from genre;

* 1. Write a SQL query to show name of all the languages in one column and number of movies in that particular column in another column.

select language\_name,movie\_id,count(language\_name) from movie\_languages as m join language as l on m.language\_id=l.language\_id group by language\_name order by count(language\_name) desc;

* 1. Write a SQL query to show movie name in first column, no. of crew members in second column and number of cast members in third column.

select title,count(cr.person\_id),count(ca.person\_id) from movie as m join movie\_crew as cr on m.movie\_id=cr.movie\_id  full outer join movie\_cast as ca on m.movie\_id=ca.movie\_id group by title

* 1. Write a SQL query to list top 10 movies title according to popularity column in decreasing order.

select title,revenue from movie order by popularity desc;

* 1. Write a SQL query to show the name of the 3rd most revenue generating movie and its revenue.

select title from movie order by revenue desc;

* 1. Write a SQL query to show the names of all the movies which have “rumoured” movie status.

select title from movie where movie\_status='rumored';

* 1. Write a SQL query to show the name of the “United States of America” produced movie which generated maximum revenue.

select title,revenue from movie as m join production\_country as pc on m.movie\_id=pc.movie\_id join country as c on pc.country\_id=c.country\_id where country\_name='United States of America' order by revenue desc;

* 1. Write a SQL query to print the movie\_id in one column and name of the production company in the second column for all the movies.

select movie\_id,company\_name from movie\_company as mc join production\_company as pc on mc.company\_id=pc.company\_id;

* 1. Write a SQL query to show the title of top 20 movies arranged in decreasing order of their budget.

select title from movie order by budget limit 20;