

Erik Safford

Spring 2019

Write Up

My schema contained 11 tables, 1 table dealt with the primary key movie_id and all the single entry columns, with the other 10 tables corresponding to the 5 multiple entry columns. With these created tables, the data is in second normal form (2NF).

movies – movie_id (PK), budget, homepage, original_language, original_title, overview, popularity, release_date, revenue, runtime, status, tagline, title, vote_average, vote_count

genre – genre_id (PK), name_genre

movie_genre – movie_genre_id (PK), movie_id (FK), genre_id (FK)

keyword – keyword_id (PK), keyword_name

movie_keyword – movie_keyword_id (PK), movie_id (FK), keyword_id (FK)

production_companies – company_id (PK), company_name

movie_companies – movie_companies_id (PK), movie_id (FK), company_id (FK)

production_countries – country_id (PK), country_name

movie_countries – movie_country_id (PK), movie_id (FK), country_id (FK)

spoken_languages – language_id (PK), language_name

movie_languages – movie_lang_id (PK), movie_id (FK), language_id (FK)

Query 1 – Average Budget of All Movies

```
SELECT ROUND(AVG(budget),2) FROM movies
```

budget

29045039.88

Query 2 – Find Movies Only Produced in the US

```
SELECT movies.title, production_companies.company_name FROM mdb.movies
  NATURAL JOIN movie_companies NATURAL JOIN production_companies
  NATURAL JOIN movie_countries NATURAL JOIN production_countries
 WHERE country_id = 'US' ORDER BY movie_id LIMIT 5
```

title company_name

('Four Rooms', 'Miramax Films')

('Four Rooms', 'A Band Apart')

('Star Wars', 'Lucasfilm')

('Star Wars', 'Twentieth Century Fox Film Corporation')

('Finding Nemo', 'Pixar Animation Studios')

Query 3 – Find Top 5 Movies that made the Most Revenue

```
SELECT title, revenue FROM movies  
ORDER BY revenue DESC LIMIT 5
```

title	revenue
-------	---------

('Avatar', 2787965087)

('Titanic', 1845034188)

('The Avengers', 1519557910)

('Jurassic World', 1513528810)

('Furious 7', 1506249360)

Query 4 – Find Movies that have both the Genre Science Fiction and Mystery

```
SELECT a.title AS Movie_Title, concat(a.name_genre, ', ', b.name_genre) AS genre FROM  
( SELECT movie_id, title, name_genre FROM movies  
  NATURAL JOIN movie_genre NATURAL JOIN genre  
  WHERE name_genre = 'Science Fiction') AS a  
INNER JOIN  
( SELECT movie_id, title, name_genre FROM movies  
  NATURAL JOIN movie_genre NATURAL JOIN genre  
  WHERE name_genre = 'Mystery') as b  
ON a.movie_id = b.movie_id ORDER BY Movie_Title ASC LIMIT 5
```

Movie_Title	genre
-------------	-------

('2001: A Space Odyssey', 'Science Fiction, Mystery')

('Atlas Shrugged Part II', 'Science Fiction, Mystery')

('Atlas Shrugged Part III: Who is John Galt?', 'Science Fiction, Mystery')

('Beneath the Planet of the Apes', 'Science Fiction, Mystery')

('Blindness', 'Science Fiction, Mystery')

Query 5 – Find Movies that have a Greater then Average Popularity

```
SELECT title, popularity from movies  
WHERE popularity > (SELECT AVG(popularity) FROM movies)  
ORDER BY popularity DESC LIMIT 5
```

title	popularity
-------	------------

('Minions', 875.581)

('Interstellar', 724.248)

('Deadpool', 514.57)

('Guardians of the Galaxy', 481.099)

('Mad Max: Fury Road', 434.279)