

# The movie database

## Introduction:

The movie database TMDB, which is biggest source of metadata is a community-built movie and tv show database created to answer many questions related to movie industry.

In our movie database domain, we have taken subset of this huge metadata (around 5000 movies) which contains information about Movie Names, Lead Actors Names, Genre, Budget of that Movie, Revenue, Release date, Rating given to that movie, Run Time of the movie etc.

For our project we are considering dataset from [www.kaggle.com](https://www.kaggle.com/tmdb/tmdb-movie-metadata), the link to dataset that we are using is <https://www.kaggle.com/tmdb/tmdb-movie-metadata>

## The ER Diagram for the domain (The Movie Database):

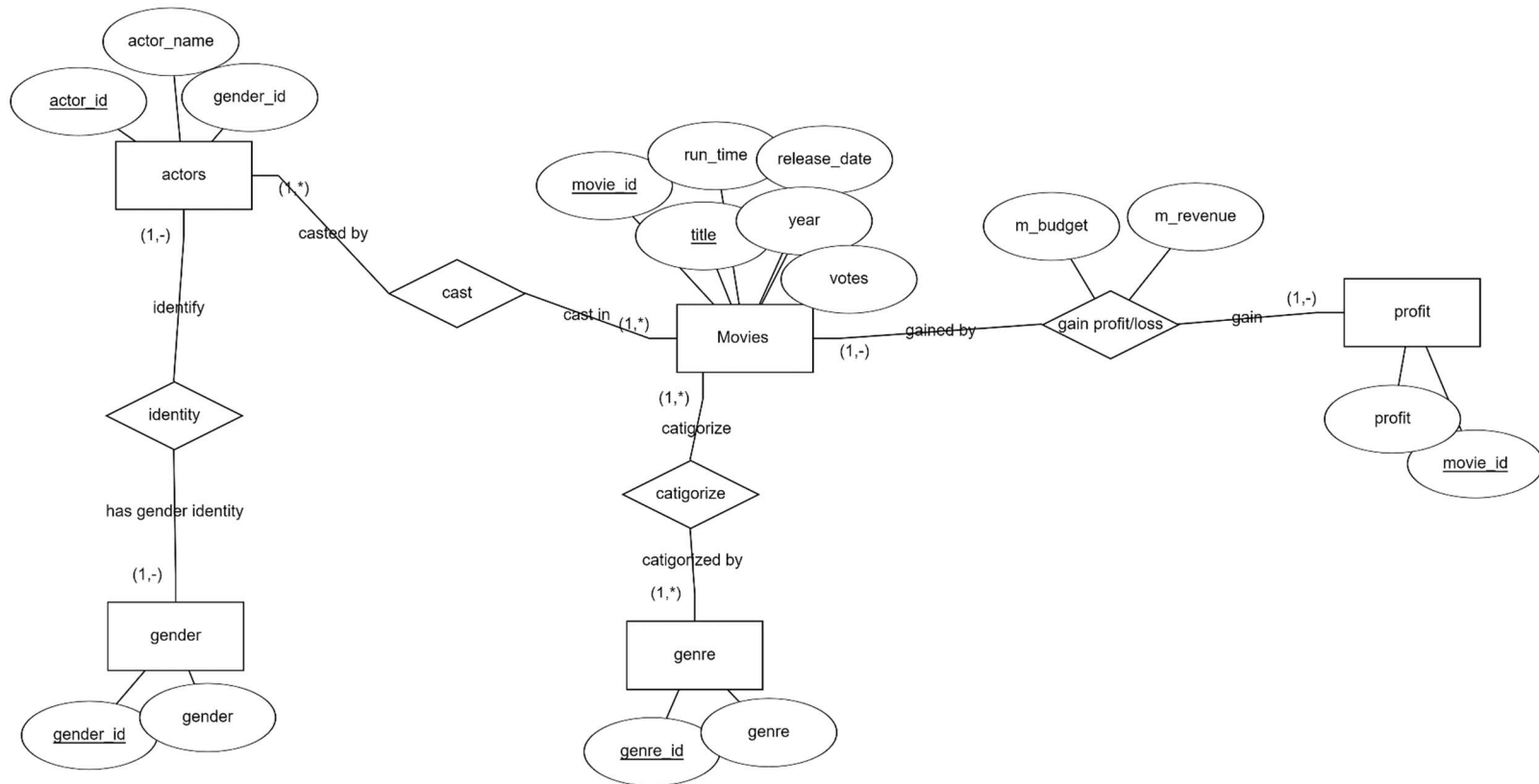


Fig 1: ER Diagram of "The Movie Database"

### ***Relational Schema:***

Using the “tmdb-movie-metadata” data source, we are able to generate 8 relational database tables containing different datatypes, attributes, relations and cardinalities. The relational schema is described below.

- i. **Movies (movie\_id, title, release\_date, year, run\_time, votes)**  
Movies.title has UNIQUE constraint on it
- ii. **Genre (genre\_id, genre\_name)**
- iii. **Movie\_Genre (movie\_id, genre\_id)**  
movie\_id is a foreign key referring to Movies.movie\_id  
genre\_id is a foreign key referring to Genre.genre\_id
- iv. **Budget\_Revenue (movie\_id, m\_budget, m\_revenue)**  
movie\_id is a foreign key referring to Movies.movie\_id
- v. **Profit (movie\_id, profit)**  
movie\_id is a foreign key referring to Movies.movie\_id  
*(Here movie\_id is both primary key and a foreign key; Profit entity has one-to-one relation with Movies entity)*
- vi. **Gender(gender\_id, gender)**
- vii. **Actors (actor\_id, actor\_name, gender\_id)**  
gender\_id is a foreign key references to Gender.gender\_id
- viii. **Movie\_Actors (movie\_id, actor\_id)**  
movie\_id is a foreign key references to Movies.movie\_id  
actor\_id is a foreign key references to Actors.actor\_id

***Questions that can be asked from this Domain:***

1. Which movies were release on same date as the movie “Harry Potter and the Goblet of Fire”
2. which movies have low budget but high revenue?
3. List the name and id of all movies that has a budget over 150000000, with female actors
4. Which actor has appeared in most movies and what is that count?
5. What is the highest budgeted movie per year?
6. list the name of the movies and its genre which has the highest votes
7. find the actor who has the most movies under the genre "Action"
8. find the Names of the male and female actors who has a total revenue over 1000000000
9. how many Number of movies were acted by each actor.
10. find the names of the actor who had no releases in the year 2008
11. find the movies with votes > 8 and budget less than 105000000 and had revenue greater than 105000000
12. What is the budget spent on each genre?
13. find the total revenue made by "James Bond" each year.
14. What are the number of movies made per each genre?
15. find which genre of movie is most popular(highest votes)
16. Which year saw the most releases?
17. How many Number of movies were released in each year?
18. find the genre's released in the year 2009
19. list the highest budgeted movie in each year
20. find the name of the actors who has worked in more than 2 genres