

# Movies Dataset Project

## Team:

Isis Ramirez

Siri Manjunath

Xiaoye Zhang

## ERD ENTITY DESCRIPTIONS:

- MOVIE(MOVIE\_ID, TITLE, POPULARITY, BUDGET, REVENUE, RUNTIME, ...)  
The MOVIE entity is central entity of our ERD, holds various basic information about movies to help analysis.
  - MOVIE\_ID - A unique number to identify different movies
  - TITLE - The original title of a movie, could be in different languages
  - POPULARITY - A numeric popularity score
  - BUDGET - The budget of a movie with US dollar as unit
  - REVENUE - The revenue of a movie with US dollar as unit
  - RUNTIME - The runtime of a movie with minute as unit
  - STATUS - The current status of a movie, could be released, planned or cancelled
  - TAGLINE - The tagline of a movie
  - OVERVIEW - The overview of a movie
  - VOTER\_AVERAGE - TMDB vote averages
  - VOTER\_COUNT - TMDB vote counts
  - PRODUCTION\_COMPANY - The production company of a movie
  - ORIGINAL\_LANGUAGE - The original language of a movie
- ACTOR(ACTOR\_ID, GENDER, NAME),  
The ACTOR entity holds information such as gender and name about actors and actresses that may appear in a movie.
- CAST(ACTOR\_ID, MOVIE\_ID, CHARACTER, ORDER)  
The CAST associative entity relates the ACTOR table to the MOVIE table by mapping the actresses and actors cast in a particular movie. It lists the character played by the actor and the opening credit order of the actor.
- CREW\_MEMBER(CREW\_ID, GENDER, NAME)  
The CREW\_MEMBER entity holds the gender and name about crew workers such as possible directors or writers a movie may have.
- PRODUCTION\_TEAM(CREW\_ID, MOVIE\_ID, JOB, DEPARTMENT)  
The CREW\_MEMBER table is associated to the MOVIE table using the PRODUCTION\_TEAM associative entity. It maps crew members to a particular movie and lists the job and department of the crew member.

- COUNTRY(ISO, NAME)  
The COUNTRY table holds information about countries where production for a movie might have taken place in. ISO is a country code as identified by the International Organization for Standardization.
- PRODUCTION\_COUNTRY(ISO, MOVIE\_ID)  
The COUNTRY table is associated to the MOVIE table using the PRODUCTION\_COUNTRY relationship. It maps movies to their production country or countries.
- REVIEWER(USER\_ID)  
The REVIEWER table has records pertaining to IMDB users who have reviewed movies.
- RATING(USER\_ID, MOVIE\_ID, RATING\_SCORE, TIMESTAMP)  
The REVIEWER table is associated to the MOVIE table using the RATING associative entity. It maps a reviewer, their rating, and a timestamp of when the rating was submitted to a particular movie.
- KEYWORD(KEYWORD\_ID, KEYWORD)  
The Keyword entity is an associative entity and is related to the Movies entity. It represents keywords used to identify or search for a certain movie(s).
- CONTAINS\_KEYWORD(MOVIE\_ID, KEYWORD\_ID)  
This entity maps out any keyword associated to a movie. It relates the KEYWORD and MOVIES tables.
- GENRE(GENRE\_ID, GENRE\_NAME)  
The GENRE associates with the MOVIE entity and represents which genre the movie is.
- BELONGS\_TO\_GENRE(MOVIE\_ID, GENRE\_ID)  
This relates the GENRE entity to the MOVIE entity. It maps movies to their associated genre or genres.
- COLLECTION(COLLECTION\_ID, COLLECTION\_NAME)  
This entity refers to sets of movies such as trilogies and sequels and is related to MOVIES.
- BELONGS\_TO\_COLLECTION(MOVIE\_ID, COLLECTION\_ID)  
This relates the COLLECTION entity to the MOVIE entity. It maps movies to collections they are part of.