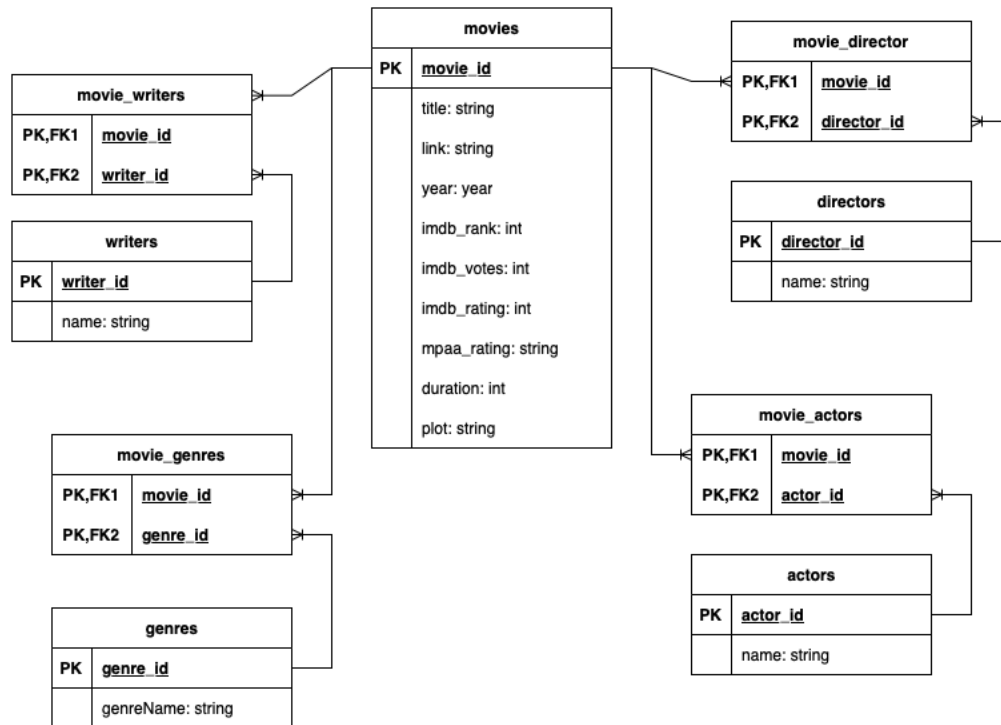


Exploring the Top 250 Movies on IMDb with R Shiny

Mackenzie Ross

Database Schema



- Tables:
 - Movies: contains the title, webpage link, release year, IMDb rank, number of IMDb votes, IMDb rating, MPAA rating (ex: PG-13), duration, and plot of the movie; primary key is movie_id
 - Movie_Writers: contains the movie_id and writer_id primary keys to identify the writer(s) of a movie in the database
 - Writers: contains the name of the unique writers in the dataset; primary key is writer_id
 - Movie_Genre: contains the movie_id and genre_id primary keys to identify the genre(s) of a movie in the database.
 - Genres: contains the unique genres names in dataset; primary key is genre_id
 - Movie_Director: contains the movie_id and director_id primary keys to identify the director(s) of a movie in the database.

- Directors: contains the first and last name of the unique directors in the dataset; primary key is director_id
- Movie_Actors: contains the movie_id and actor_id primary keys to identify what actors are in a movie in the database.
- Actors: contains the name of the unique actors in the dataset; primary key is actor_id
- Relationships
 - Actors and Movie_Actors have a many-to-one relationship because more than one actor can be in a movie.
 - Directors and Movie_Director have a many-to-one relationship because more than one director can direct a movie.
 - Writers and Movie_Writers have a many-to-one relationship because more than one writer can write a movie.
 - Genre and Movie_Genres have a many-to-one relationship because a movie can fall into more than one genre categories.
 - The many-to-one relationships between the movie_actors, movie_director, movie_writers, and movie_genres tables and the movies table identify what actors, director(s), and writer(s) worked on the movie and the genre(s) of the movie.

Database

- Constraints
 - Initially, the names in the writers, directors, and actors columns were going to be split into first and last names. However, there were too many irregularities with people having middle names/initials, suffixes, or multi-word last names to easily split the data.
- Planned Views
 - View #1: Top 10 Movies in the Database Based on IMDb rating.
 - View #2: Directors or Writers Who Appear in the Database More than Once

Individual Assessment

I would rate my task completion as a 7. This week, I decided to change to an individual final project, so I was pressed for time when building my database. In my previous group, I was doing a majority of the work and could not establish a good working relationship with my team members. I think now that I am working on my own I will be able to better manage my time on the project without feeling additional stress. My time commitment to the project was very high this week due to the switch. I am proud of the work I accomplished in a short amount of time. I think I was able to build a strong foundation in a short amount of time.