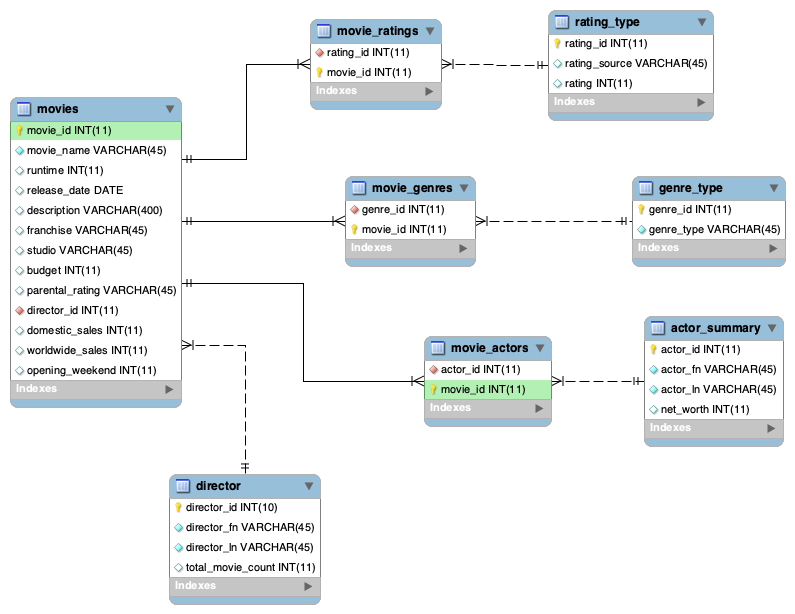
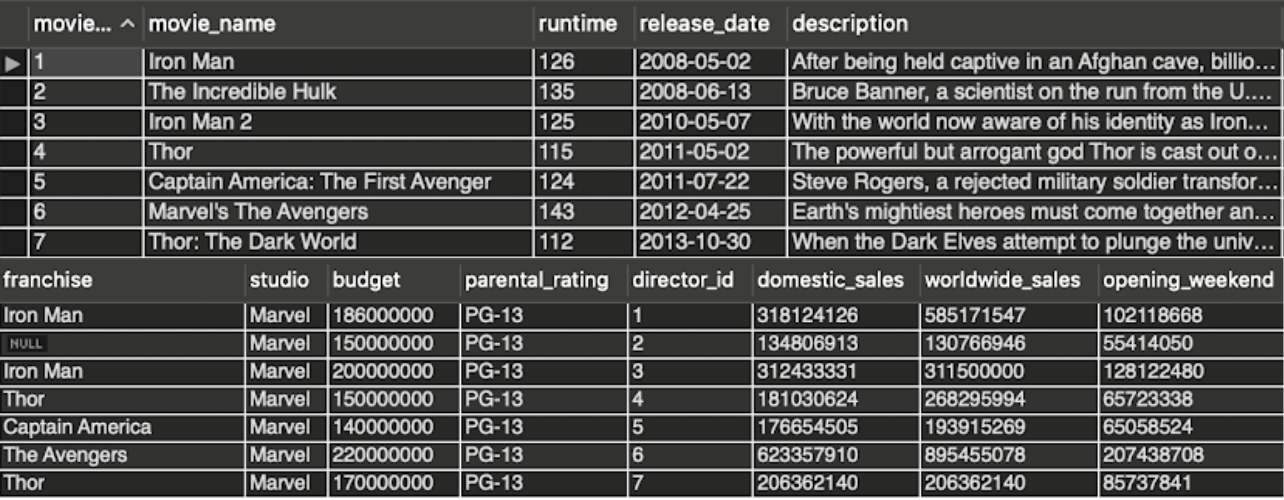
**SQL Movie Project**

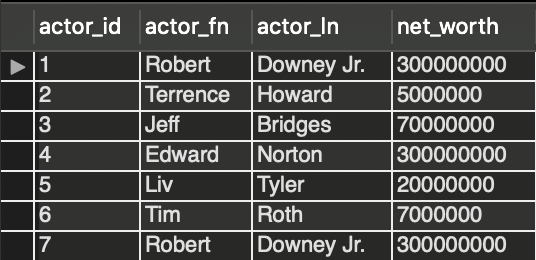
1. **Introduction**

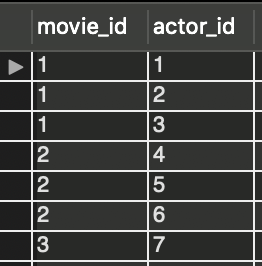
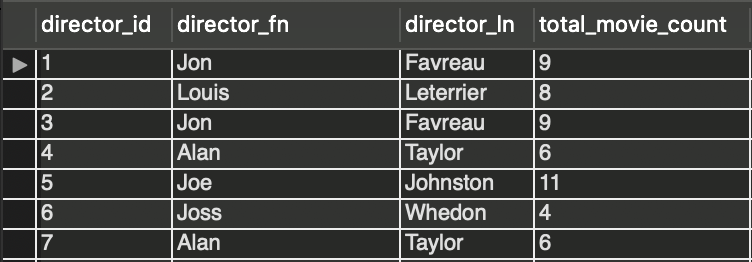
This database consists of information about superhero movies, including movies from the Marvel Cinematic Universe, DC Universe, and a select few non-Marvel or DC superhero movies. The database stores a wide variety of data including information on the movies themselves, movie runtime, critical reception, film studios, financial information, revenues, budgets, ticket sales, actor information, reviews, and more. Users can utilize this database to visualize how different movies, actors, directors compare to each other. The data can hopefully offer insight into interesting trends, statistics, or patterns about the superhero and comic book superhero film categories, or help stakeholders in the industry collect, find, and visualize information relevant to their needs.

1. **Database Description: Logical Design**
   1. The database consists of eight tables, including three linking tables that connect various specific informative tables to a table containing information on specific movies. The linking tables use the “movie\_id” from the movies table as primary keys, representing a one to many relationship in that one individual movie can contain multiple ratings. Additionally, the linking tables utilize foreign keys such as “rating\_id,” “genre\_id,” and “actor\_id,” respectively, to connect with other tables that contain information on movies’ ratings, genres, and actors with the movies table.



1. **Database Description: Physical Database**
   1. The physical database took multiple iterations to bring it to its current structure. Originally, I had additional tables that held various other forms of information. However, because these tables, including an Awards table, proved to be too complicated to format properly, I decided to drop them from the database.
2. **Database Description: Sample Data**
   1. Getting sample data into the database was a difficult but good experience. Originally, I attempted to manually input each row of data for each table, but quickly realized that was not the most efficient or effective method to input data. I later learned how to upload data in the form of a CSV. Once I learned this, I then collected, organized, edited, and formatted data in Google Sheets, which I then downloaded as CSV files and uploaded as individual tables. Almost all of the data collected is not only realistic in format, but actual, genuine data. This includes movies’ technical and monetary information, directors information, actor information, and genre information. The only data that is not literally accurate is the review information.
   2. *Sample data from the movies table:*
   3. *Sample data from the actors table:*



* 1. *Sample data from the directors table:*
  2. *Sample data from the movies\_actors linking table:*

1. **Database Description: Views and Queries**
   1. After creating the database and inputting sample data, I got to work on creating queries and views that engagee with various interesting questions users may want to see. The queries I wrote attempt to answer the following questions:
      1. *Which actors stars in the most super hero movies?*
      2. *What movies had the biggest opening weekend sales that had over $750,000,000 in gross worldwide sales?*
      3. *What are each directors’ average movie ratings and average worldwide sales?*
      4. *What is the most popular genre within the superhero movie scene?*
      5. *What is the minimum box office movie revenue of any superhero movie?*
      6. *Which directors had the highest grossing movies?*