

Executive Summary

While using SQL for doing this assignment many insights were gained about the movie industry. In Segment 1, we learnt about some of the broad information about the data we were given. Firstly, we learnt about the number of entries in each of the tables and the null values present. We were also able to find the kind of movies that were being produced (genre), as well as how popular each genre was. We also deduced the popularity of movies in different countries based on the number of movies produced. As well as how long movies are on average.

In segment 2, we started by looking for outliers in the data by checking the minimum and maximum a movie had been rated and the number of votes a movie got. We followed that up by getting insights on the movies. We found information such as the top 10 movies and the distribution of different ratings. It was discovered that a rating of '7' is the most common. We then ranked the production companies in order to find the best ones and found the most popular genre. We also did some key word searches to determine how popular the key word 'the' is in movie titles.

In segment 3 we found information on the popularity of actors based on both country as well as language. We also found information on directors and were able to determine the best directors to direct movies in the most popular genres.

In segment 4, we were given a broad understanding of the tables present such as the average duration of each of genres of movies. Top actors/actresses and the production house with the highest number of hits. Meaning most likely to be successful.