# Project Proposal

This project is for a Movies investing company that wants to know if a movie worth investing in at production stages, given that we will be studding the already released movies from IMDB which includes data regarding ratings, genres, release dates, whether or not each movie was directed by a winner of an Oscar and more, to predict the rating of movie before the production.

## Question/need:

Predicting movies ratings to help the Movies investing company make decision of what movie to invest on before production.

## Data Description:

The data we will be using is a set of 1000 movies scraped from IMBD using Beautifulsoup and Selenium, it include ratings, genres, movie release date, actors and movie director.

Variables included in initially processed data:

- Rating = refer to the rating of the movie.
- num-reviews = refer to the number of reviews in the movie.
- genres = refer to the genres of the movie.
- release-date = the date of the movie release .
- oscar-winner = boolean variable that will be true if the movie won an oscar.
- runtime = duration of the movie.
- MPAA Rating = rate a movie's suitability for certain audiences based on its content.
- Budget = movie budget.

## Tools:

Python, Pandas, SQL, Matplotlib, Jupyter notebook and NumPy, Beautifulsoup, Selenium,

Scikit-learn.