We have several responses for movies from users on different platforms like IMDB, youtube videos, facebook, twitter, etc. and there is a lot of data available to judge people reaction for a particular type of movie. I believe if in-depth data analysis is done, it might be possible to find a good correlation between certain features and success of the movie. Once I am able to determine these key indicators, I can see their trends before the movie and predict if the movie will be a success. For instance, if a certain genre, certain star cast correlates strongly with a success of failure of the movie we can predict the outcome of the movie. We can study people's reaction to the movie trailers.

In the current submission, a single notebook is provided with some preliminary exploratory analysis. Please refer to this notebook for all the plots.

In the analysis, we have analyzed the impact of the budget on the ratings, Average rating for low and high budget movies over the years, trends of the proportion of movies with bigger star cast vs year. The results are shown and explained in the ipython notebook.

Basically, we observed that high budget movies have a high average imdb rating. We also found that the average rating of low budget movies is decreasing over the years. Also, average ratings per year are declining as well. We also observed that proportion of movies with bigger star cast is declining. These are some interesting results and suggests a future path to explore that is there a tendency to produce smaller start cast movies which might explain the low budget. Smaller start cast is not good enough to attract people when they have movies with bigger star cast available to watch and hence the average imdb rates are declining over the years. These are only preliminary results and further analysis is required to prove this hypothesis.