Project: Visualizing Movie Data

Step 1: Data Cleanup and Attribute Selection

The attributes explore in this project are:

- original_title: Movie title;
- popularity: Relative number of page views on The Movie Database;
- keywords: list of keywords associated with the movie;
- genres: list of genres;
- production_companies: list of production companies;
- release_year: Release year;
- vote_average: average of votes;
- budget_adj: Budget adjusted for inflation, in 2010 US dollars;
- revenue_adj: Revenue adjusted for inflation, in 2010 US dollars.

We take the revenue adjusted, subtract the budget adjusted, and then divide by budget adjusted, to obtain the return on investment, a measure of profitability.
We remove movies with null value in the “genres” field. Moreover, from http://www.imdb.com/genre/, we obtain a list of movies genre, then the movies with value “TV Movie” and “Foreign” are removed.
The movies with value 0.0 in the “budget_adj” and “revenue_adj” removed, given that this value means that information is not available. We have considered a few records and we have investigated the correctness of values through googling. There are incorrect and inaccurate values, in particular, those low budgets adjusted or low revenues adjusted. We decide to remove movies with budgets adjusted and revenues adjusted lower than 100 000.

Step 2: Tableau Visualizations

https://public.tableau.com/profile/santino.farinella

Step 3: Questions

- **Question 1:** How have movie genres changed over time?

From 1960 to 2015, the most popular movies genres are four: drama, comedy, thriller and action. These genres grew more than other. For each genre, we observe trend lines of the adjusted budget, adjusted revenue and return on investment, and we derive, in general, that the average budget adjusted is increased, the average revenue adjusted and return on investment are decreased. On side adjusted budget, the genre with the greatest increase is animation. If we consider adjusted revenue, the genre with the greatest decrement is horror. The horror is also a genre with the greatest decrement of the return on investment. In the overall period, the genre with the highest average return on investment is horror instead of the genre with lower average return on investment is history.

- **Question 2:** How do the attributes differ between Universal Pictures and Paramount Pictures?

In the period of analysis, the production of movies increased for both Universal Pictures and Paramount Pictures. Universal Pictures has produced more movies. We observe, in the overall period, a productions similarity between two production companies. The top genres of film produced are the same. In general, the average budget adjusted, average revenues adjusted and return on investment are growing or shrinking over time. However, observing trend lines, the performance of companies is similar. In particular, their profitability decreased. This because the revenues reduced, but remains greater than average budget adjusted although over time the average budget adjusted increased. Paramount Pictures has slightly better performance.

- **Question 3:** How have movies based on novels performed relative to movies not based on novels?

From plots, the number of movies not based on novels release is smaller than movies based on novels. In particular, we observe a trade-off between adjusted budget and adjusted revenue. Movies with the lower adjusted budget have higher ROI. Movies based on novels have a higher return on investment than movies not based on novels. We consider the annual performance, only a year, 1970, the ROI of movies not based on novels is results higher than movies based on novels.

- **Question 4:** How performance movies: quantitative vs. qualitative measures

The Movie Database is an online database of information related to the film. If I do not know what to watch, how to understand whether a movie is good or not. The Movie Database is a place to look for a movie rating. In particular, we can consider performance movie, quantitative measures, as return on investment, or qualitative measures, as popularity and vote average. The return on investment explains financial performance, profitability; therefore, it depends on revenue and budget. The movie is considered as an investment. The question is the movies with the greatest profitability are those more popular and to like best. We observe that there is no clear relationship between quantitative measures and qualitative measures. Then, for each movie, the dashboard allows understanding his performance, quantitative or qualitative.