

INFO 5100 Project 2 - Interactive Data Visualization

Project Topic: Movie Recommendations!

Group Members: Ming Cheng, Nurpur Kale, Suganya Sankaran

The Story about Our Visualization:

“What movie to watch?” is such an important question that will directly affect the quality of your relaxing Saturday night and if you are one of those “who-cannot-make-up-your-mind” people, picking a movie can be a nightmare. And most importantly, movie preferences are personal. It’s more than what year it is published or which genre it belongs too. People care about more factors than the usual google results. We want to know how the movies are rated considering both box office success and the popularity of the lead actors.

So, unlike traditional movie rating, this data visualization is aimed to display data in a more dynamic way, which allows the users to choose the top 2 factors about a movie that matters most to them in addition to traditional filters. For example, if someone would like to see a movie that receives higher box office success and higher actor likes rather than just finding a movie with highest score rating, he can find just the right recommendations and customized ordering of the movies based on our visualization.

Now, get your popcorn ready, and see what our data visualization can surprise you with movie rankings that you’ve never thought about.

Data Description and Mapping:

We used two data source movie_metadata.csv and award.csv from [kaggle.com](https://www.kaggle.com). For both data sets, we extracted data only for movies that are published after 2000, in English, and have complete information for the variables we selected in order to keep the number of movies manageable for the purpose of this project. From movie_metadata.csv, we extracted variables of duration, genres, year, language, number of likes for lead actor on Facebook, budget spent on the movie, gross earning, movie score on imdb, number of likes for the movie on Facebook, title, plot keywords, and imdb link to the movie. Those data is combined with their Oscar nominations information from the award.csv file to give users more information on the movie. While combining the data, there are many issues we needed to watch out. For example, the data in award.csv is not labelled as desired. The name and filmName values are sometimes switched, or the year of the movie is incorrect. So while we combining the two data sets together, we made sure to check for the inconsistency in the data files.

On the left side of the website, the users are able to choose filters and factors for the data visualization. The data is filtered by the filters of user’s choice and mapped onto the scatter plot with 1st factor as x-axis and 2nd factor as y-axis in linear scale. On the right side, in stead of a standard scatterplot look of the data, we tried to mimic the star sky of the background. On the black background, the circles in yellow representing each movie and the shaded orange region indicates those that have both high x-value and y-value, which are the most recommended movies. The user can hover over each circle for more details on the movie, such as title, award nominations, year, duration, keywords of the plot, and also a clickable imdb link if they decide to watch it.

To improve usability, we had incorporated some features for the interaction. If none of the filter or the factor is chosen, defaults to the filters are “all”, and default factors are budget and gross. This prevents the program from crushing. Also, it will give a statement to ask the users change filter selection if no movie satisfies the selected filters. And we also added a switch button for the

users to choose movies either only show movies with Oscar award nominations or all movies regardless of awards.

Team Contribution:

Name/netid: Ming Cheng(mc2637)

1. Worked on data management and combination
2. worked on scatterplot and data mapping
3. worked on styling of title
4. worked on final report

Name/netid: Nupur Kale(nak85)

1. Worked on creating the HTML page: layout, styling and elements
2. Worked on jquery multiselect dropdown
3. Worked on sending accurate parameters from filters and x-y axis factors to showMovies function
4. Worked on oscar nominated movies switch interaction to show movies accordingly on the plot

Name/netid: Suganya Sankaran(ss3629)

1. Helped on initial ideation and brainstorming.

Data and Code Resource

movie_metadata.csv: <https://www.kaggle.com/deepmatrix/imdb-5000-movie-dataset>

award.csv: <https://www.kaggle.com/theacademy/academy-awards>

select dropdown styling: <https://codepen.io/ericrasch/pen/zjDBx>

checkbox styling: <http://www.csscheckbox.com/checkbox/42432/padded-orange-check/>

d3 tooltip: <http://bl.ocks.org/d3noob/a22c42db65eb00d4e369>

title animation: <https://cssanimation.rocks/starwars/>