Project 2: Interactive Data Visualization

Alisha Sojar (ahs264), Lisa Nuccitelli (lan48) , Ryan Dong

Data *Source:*
http://www.informationisbeautiful.net/2012/hollywood-budgets-a-data-viz-challenge

Overview

The entire dataset contains information on every Hollywood movie release from the years: 2013 – 2007. Each entry contains useful information about the movie's financial and critical success: RottenTomatoes ratings, number of opening theaters, profitability, budget, gross revenue, and a wealth of other numbers. The main focus of our visualization however was seeing profitability compared with budget and ratings. The data defines profitability as the ratio of revenue to budget, and thus is probably the most important factor in determining the financial success of a movie. Plenty of movies draw large audiences and generate enormous amounts of revenue, however that does not mean anything if the movie cost an enormous amount of money to film. So now the question to answer: "Does spending more money on a movie mean it is more likely to be a critical success? And does that correlate at all with its financial success, its profitability?"

Editing the Data

The original data was separated in an Excel spreadsheet, with separate sheets for each year. Most of our interactivity is in the ability to filter the data by various parameters – genre, budget, years, etc. – so it was important to compile the various sheets into a single .csv file. Once compiled, we added in a year tag to each object so that the code can differentiate between every movie. Besides that change the rest of the data remained unaltered.

Mapping to Visual Elements

The data is mapped onto a scatterplot, thus the important visual elements to note are the axes and the points themselves.

Axes

The x and y axes represent a movie's budget (in millions of US dollars) and the movie's rating, respectively. On the page are sliders, two of which specifically interact with the two axes. These sliders allow the user to set the min and max of each scale and by result

allows the user to filter out data not within these bounds. The axes themselves otherwise are normal linear axes created with the d3.axis library using a linear scale.

Data Points

The data points themselves will also be representative of information through its size and color. The first thing to note is mousing-over a data point brings up a tooltip with the movie's full name. Secondly, the area of the circles are linearly scaled with profitability, or the % of budget recovered. This is done by linearly scaling with the radius squared, rather than just the radius which would make for a counter-intuitive visualization. On the screen is also a slider that allows for the user to filter out any data point not within the selected range. Last but not least is the color and year of the data point. Visually the colors are representative of genre a specific genre while the year information is not mapped. However this information can be filtered through the radio buttons present on screen, allowing another form of filtering for the user.

Pie Charts

We also added two pie charts to visualize the profit breakdown for individual movies when the movie is selected from the scatterplot. The first pie chart shows the relationship between the domestic and foreign gross profit that the movie made, while the second pie chart shows the relationship between the profit the movie made on its opening weekend and the profit made on all days following its opening weekend.

The Story

This visualization shows that high-budget films are not necessarily the smartest decision when trying to maximize the profitability percentage of a movie. From our visualization, we see that the biggest circles are often towards the left, or the lower budget films. This means that these films are the ones that provided the most 'bang for their buck', and were the most successful economically. These films also tend to fall under the "Drama" genre.

We also see that highly rated films don't necessarily correlate to a higher profit. Films that got very low Rotten Tomato ratings still have some of the highest profitability percentages. In other words, an audience's opinion of a "good movie" is not necessarily an indicator of how financially successful it will be!