Data Set: Movies

Initial question: **Is there a relationship between reviews on websites such as rotten tomatoes and how much revenue a movie brought in?** To analyze this I made a simple scatter plot of both IMDB and Rotten Tomato ratings vs the World-wide Gross which you can see below.
It really didn’t seem like the worldwide gross and critical reviews had much if any correlation, which I found surprising. Just to make sure, I also tested it against the US gross and found a similar result as you can see below. However for the total US Gross, the values tend to be skewed more towards the left -- There are less high-grossing movies that have low ratings then there are in the World-wide gross. I think part of this may have to do with the fact that movies intended for a large scale international release tend to have lots of special effects but little in the way of plot or dialogue - probably because it makes it easier on producers to translate the film into many different languages. But this would also have the adverse affect of lowering critical review scores.

Still, I wanted to find **what has the largest effect on ratings? Is it production budget?** So again, I made another scatter plot with, this time with production budgets on the x-axis, but unfortunately the results weren’t any more promising. The graph below has different colored circles corresponding to different genres, but regardless of the genre, the relationship between ratings and production could not be established.
Then I thought that maybe I just wasn't displaying the data in a way where the relationships between the variables could easily be determined. So I tinkered with things a bit and using the same three variables of Rotten Tomato rankings, Avg. Production budget and genre, I created the graph that you can see on the right. Again you can see how little correlation there is between production budget and ratings, however what does
stand out from this are two things:

1) Documentaries have the lowest production costs but have the highest average rotten tomato ratings. **Why is this the case?**

2) Action and adventure movies have the highest production costs but also do pretty highly in rankings, so I considered these genres to be the most consistent with what one would expect.

I will be coming back to the documentary phenomenon, but I first wanted to analyze adventure movie and I got a new idea, I was thinking that **maybe inflation may have had something to do with why production budget and rankings weren’t matching up** quite as well as I thought they did. The same movie made in 1980 wouldn’t cost as much to make as it would in 2010 for instance, so comparing movies made in different eras directly could skew the data: A “cheaply” made film in 1980 could have done really well ranking-wise compared to an “expensive” film in 2010 if you compare the data directly, but adjusting for inflation it could be the fact that the 1980 film was actually more expensive and thus it would make sense why it did better. To test this theory, I created the following graph, which compares average Rotten Tomato ratings over time with average production budget over time, with different colors representing different genres.
Here is the same graph but with Adventure movies highlighted in orange:

And here is one more with Drama's highlighted in a salmon color:

Again, unfortunately, my hypothesis seems to be wrong and there doesn't seem to be any correlation whatsoever between how much money a studio pours into a movie and how well it
seems to do in the rankings. What I was expecting was to see a more or less linear relationship: the bars on both graphs would increase at a roughly equal rate, but what we see instead is that while action/adventure movies are pouring in more and more money into their budget, their ratings are going down and down. This could possibly tie into what I mentioned earlier about how in order to make movies more easily relatable to people of different countries, more special effects (ie more money) are used and less plot and dialogue are used (lower ratings). This is not to say that the foreign market is dumb - far from it - it’s just that a highly nuanced and intricate plot is 1) hard to translate accurately and 2) certain cultural references and assumptions might be lost on people unfamiliar with American culture.

But I liked the timeline concept, so I stuck with it for a bit and decided to change the color from representing genre to representing MCAA ratings. **What types of ratings would bring in the most money over time compared to production value?**

![MPAA Rating](image)

The above chart on the left is the result.

What surprised me the most was just how far and above PG-13 movies were above movies of other rankings. It didn’t surprise me that NC-17 movies for instance would barely make a presence simply because historically an NC-17 movie has been a sort of “kiss of death” as few theaters would be willing to show an NC-17 movie. But after some consideration, it started to make sense why PG-13 movies would be so popular and it’s because they can attract the widest audience. For instance, parents would be unwilling to bring their young children to see an R movie, but most
would be okay with bringing them to see a PG-13 movie, even if they were under 13 (Harry Potter for instance comes to mind). However, not many adults except those with children usually go to see PG and G rated movies but many adults would definitely go to see a PG-13 movie on their own. So PG-13 movies are that perfect balance that can appeal to everyone.

Going back to the documentary issue which I found interesting, I wanted to see just if I could showcase production budget, ratings and genre in an easy to see format and came up with the following chart where size of each ring represents production budget and color is genre. I liked the look of this very much and decided to do it a second time, but with MCAA ratings instead of genre.

These became my final visualization:

![Diagram](image-url)