

# Project: Visualizing Movie Data

## Data Cleanup and Attribute Selection

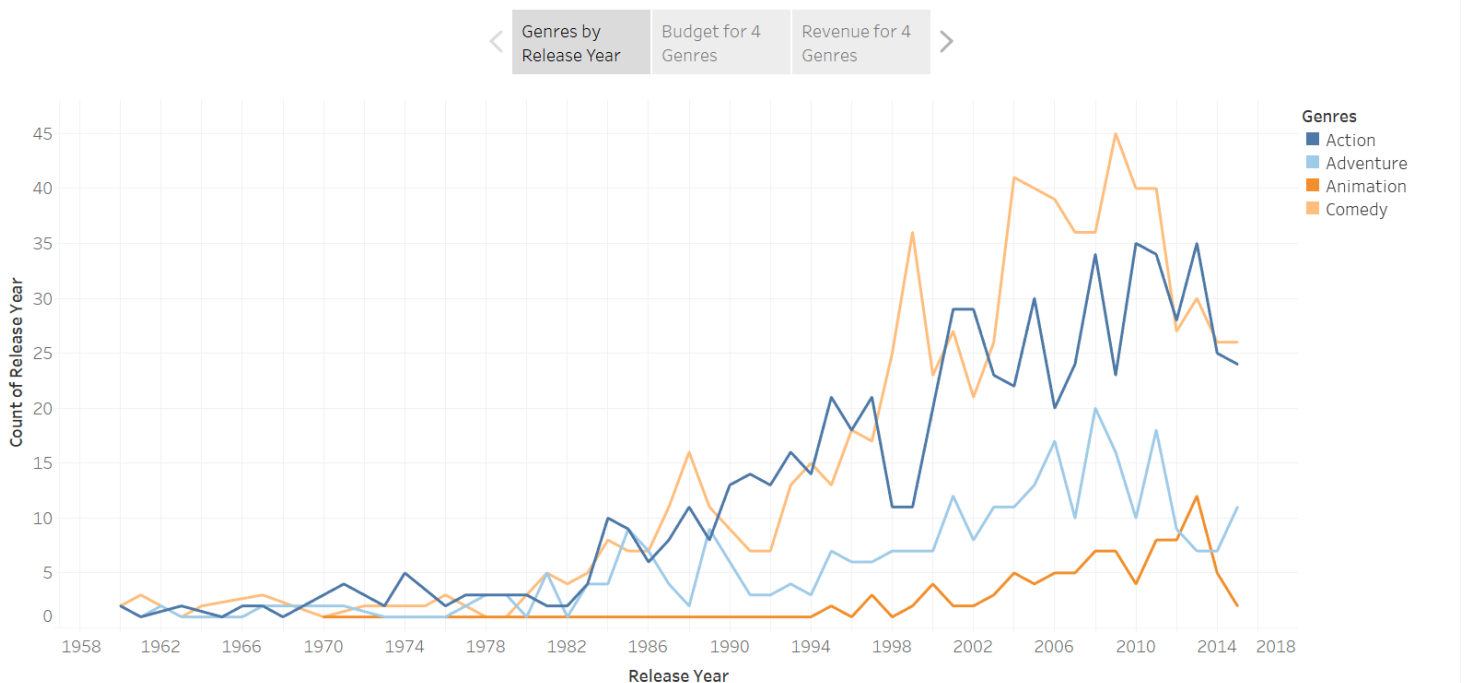
- Attributes that will be examined further are **Popularity, Genres, Release Year, Budget Adj, Revenue Adj, Gross Margin, Production Company** and **Keywords**.
- List of movies with 0 budget or revenue are removed. That leaves us with 3855 rows of data out of 10866 rows. Even though a large number of data are being excluded, the exclusion will help with the data investigation as budget and revenue are the main variables to be studied.
- When sorted by popularity, majority of movies with 0 budget or revenue are poorly ranked and should not skew the data, thus decision to remove them are justified.

## Tableau Visualizations

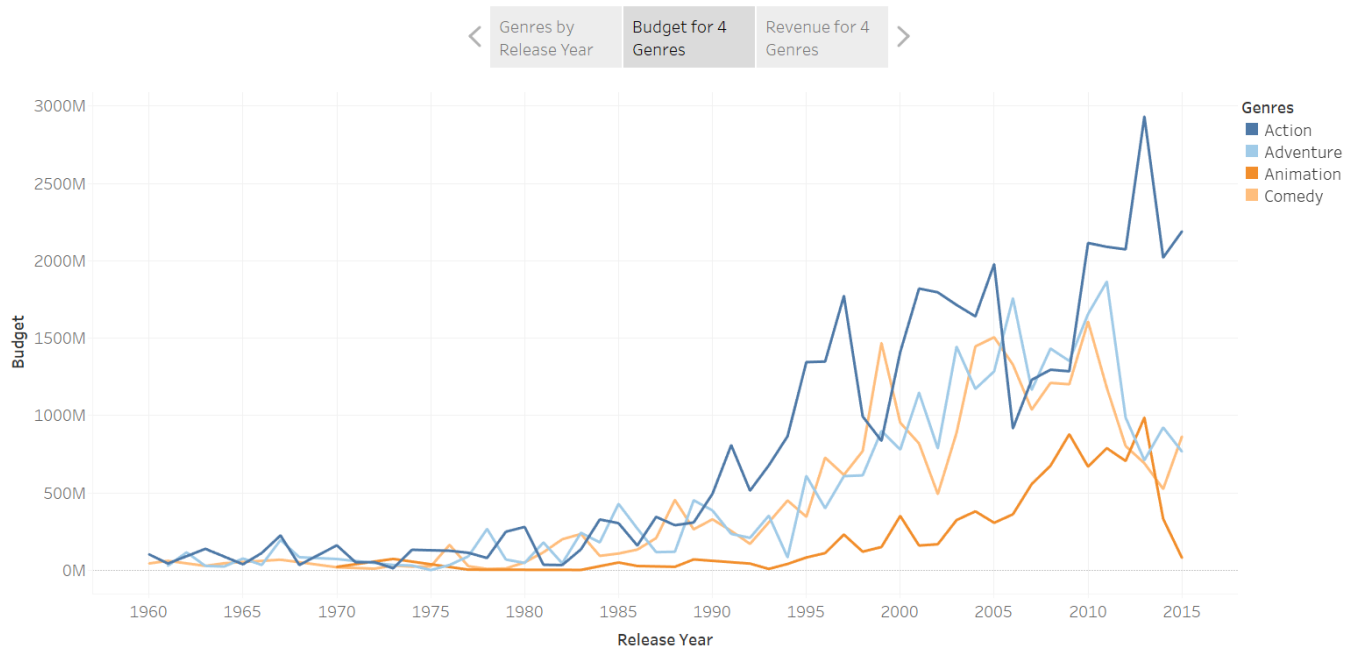
Public Profile: [https://public.tableau.com/views/p5\\_3/Q1?:embed=y&:display\\_count=yes](https://public.tableau.com/views/p5_3/Q1?:embed=y&:display_count=yes)

## Questions

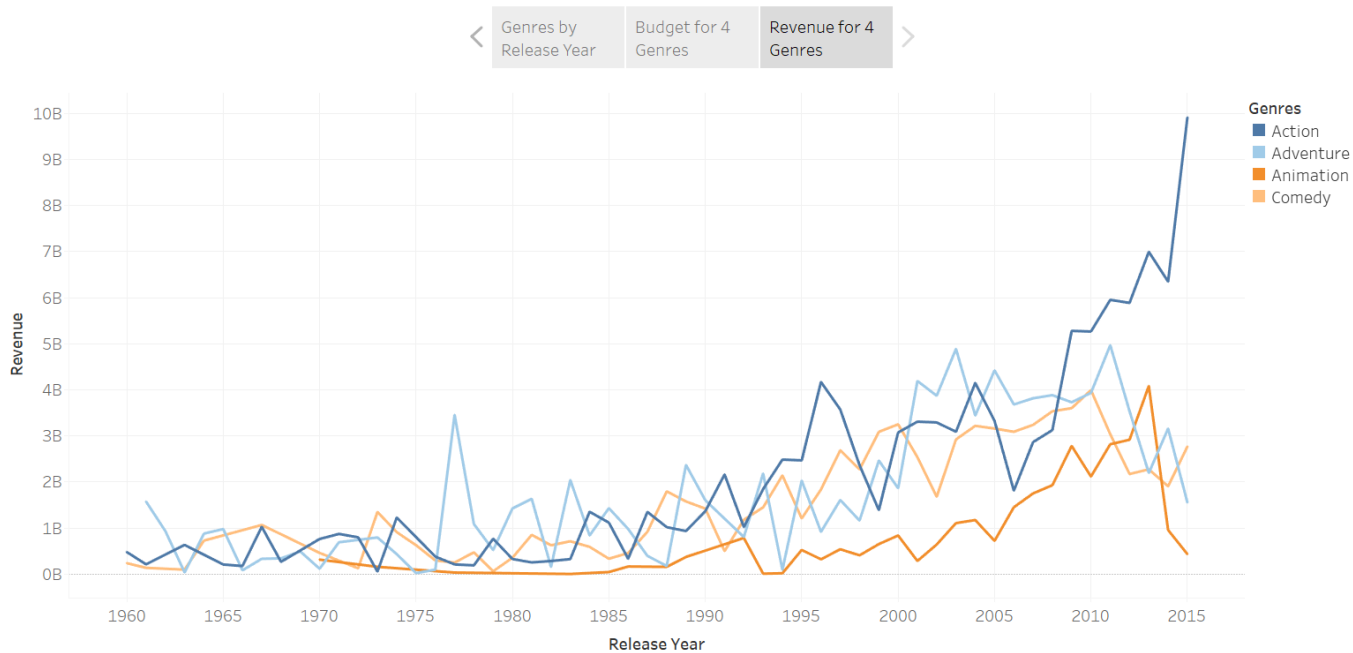
Question 1 - Story: How have movie genres changed over time?



4 characteristic genres were chosen to make the analysis simpler. In terms of number of movies released, comedy saw that sharpest increase, among the 4 categories.

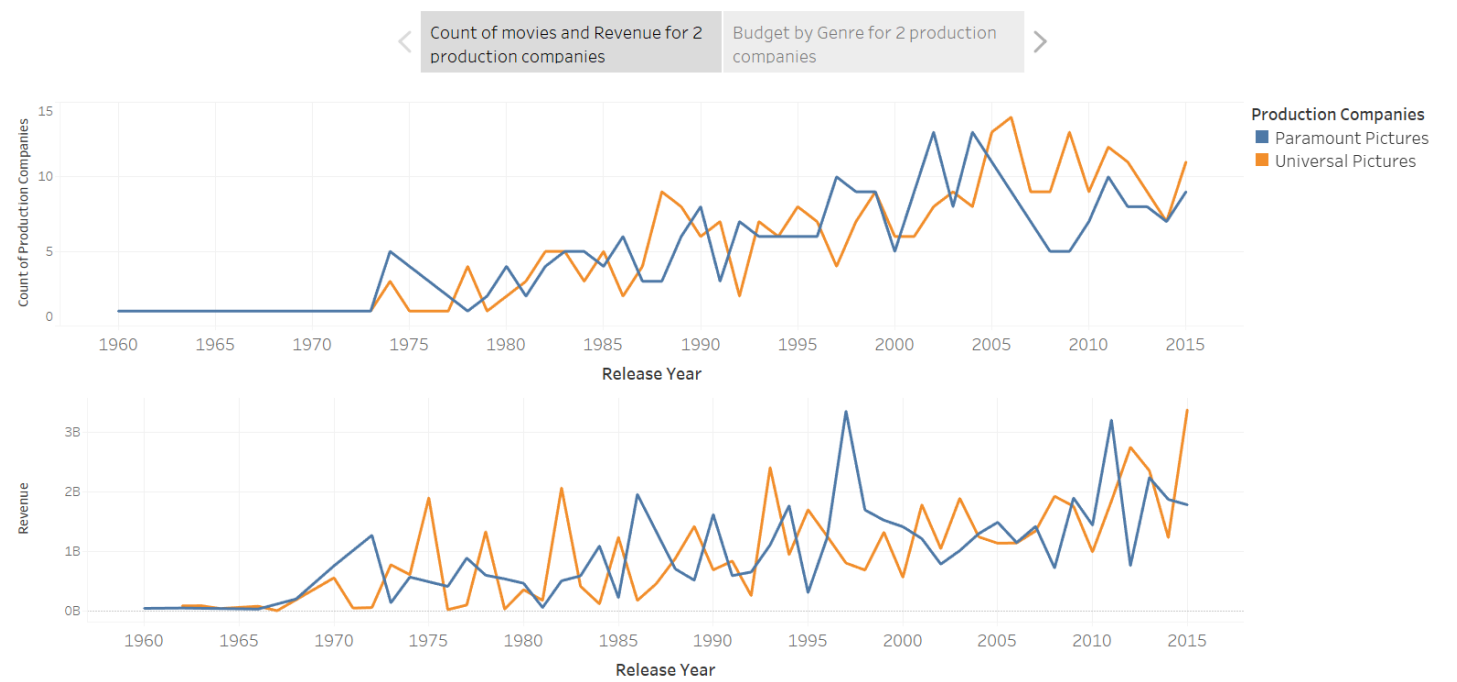


In terms of budget spent for the production of the movie, Action was the genre with highest increase from 1960 to 2015.

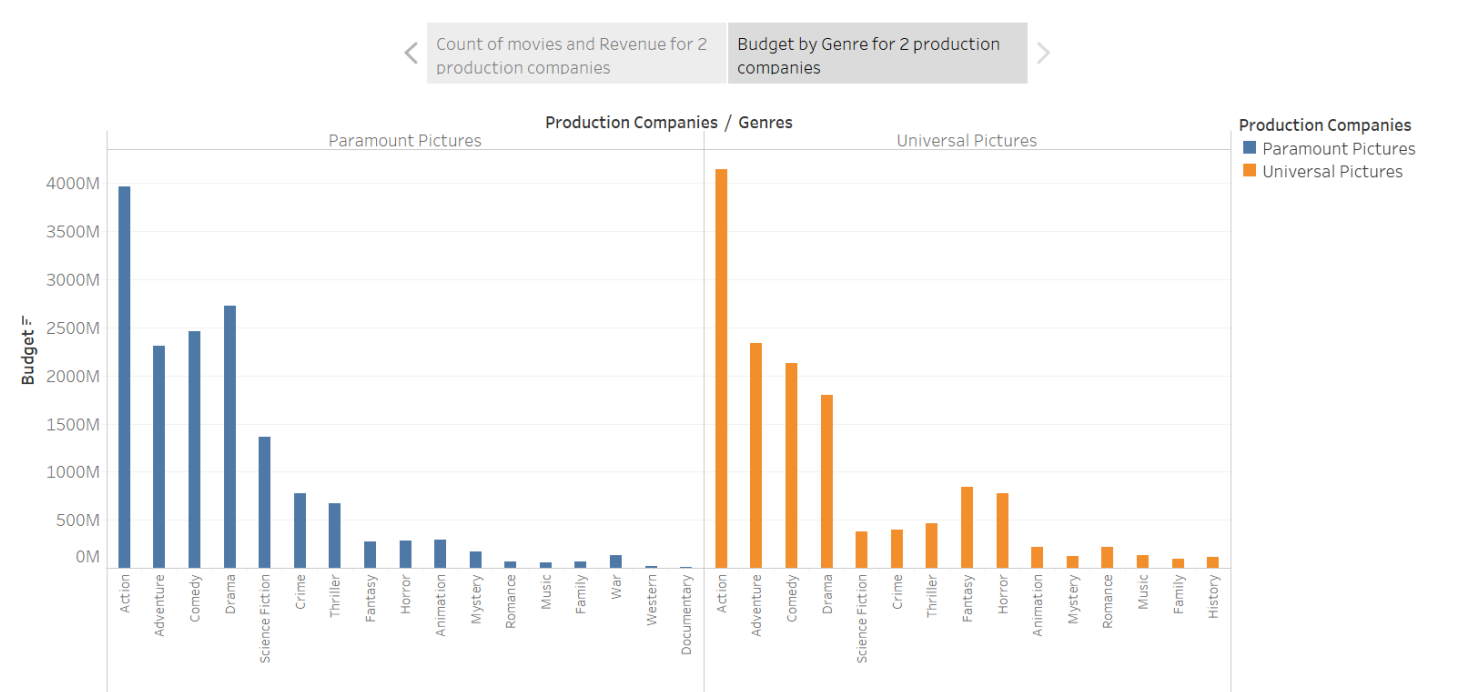


Regarding revenue generated from the movies, same as budget, action was the category with the highest increase.

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

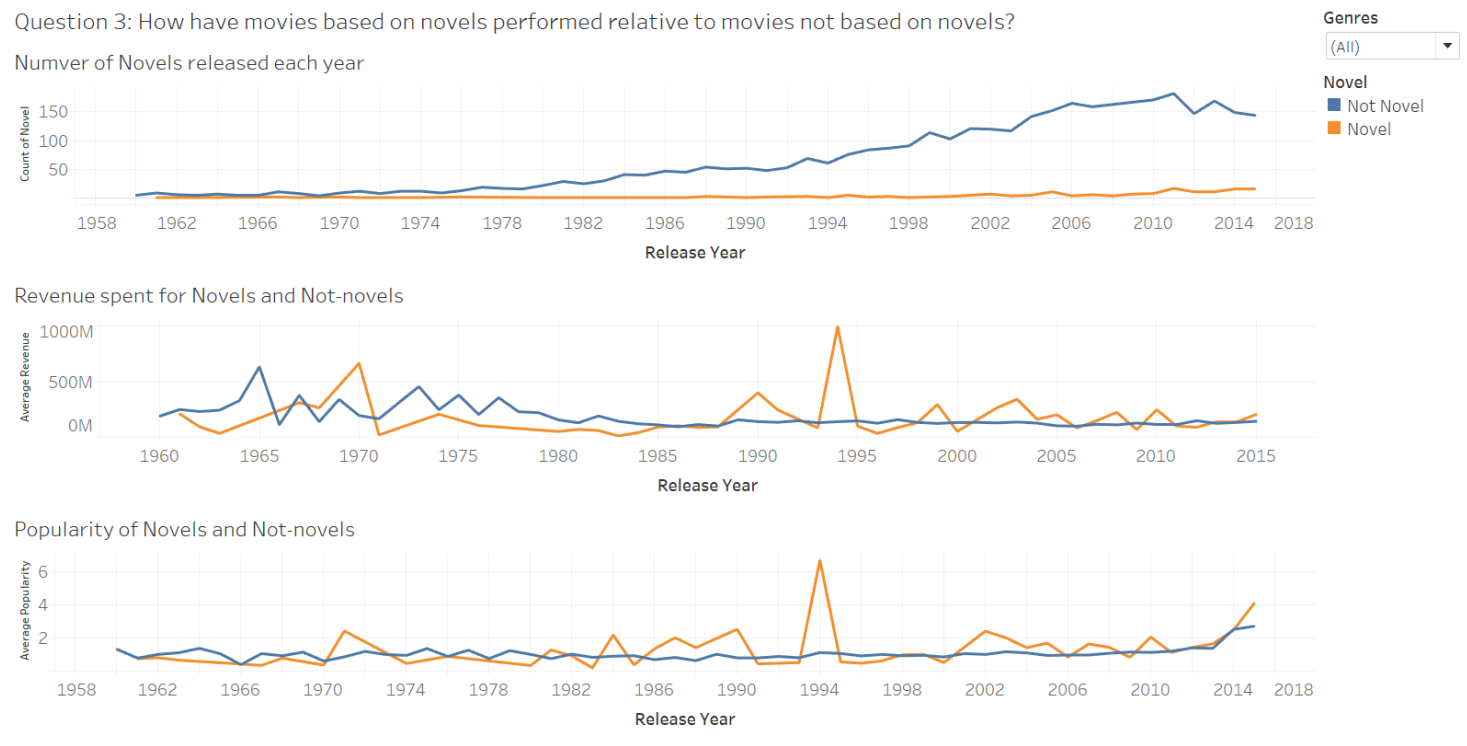


Universal Pictures tend to release more movies each year than Paramount Pictures. This fact affects the revenue the same way.



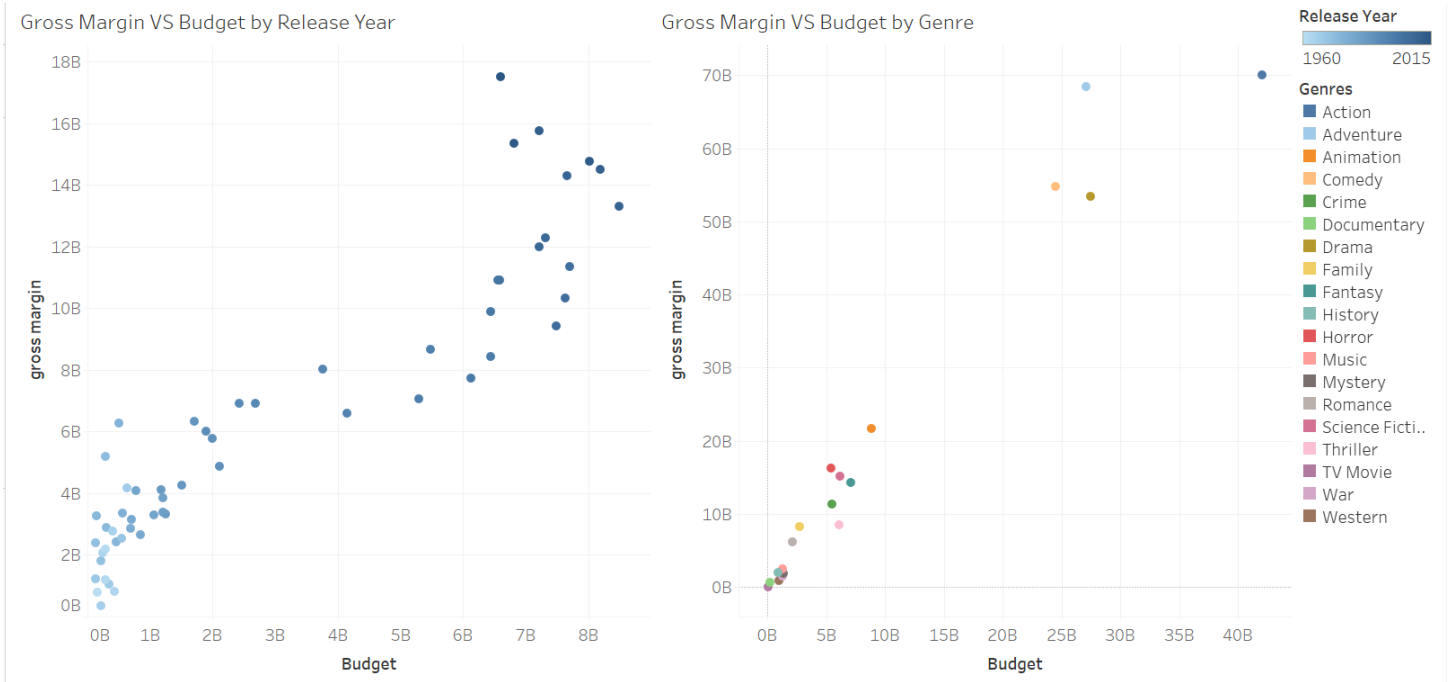
In terms of budget, the two companies seem to spend their money for movies of the same genre, while Paramount Pictures spend more for Comedy than Adventure whereas for Universal Pictures it's all the way around.

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



The graphs above compare the movies based on novels and those not, in terms of number of movies released, revenue and popularity. It's obvious that movies not based on novels have higher figures.

Question 4: What is the relationship between gross margin and budget of a movie?



It can be seen that as the years were passing by the relationship between budget and gross margin became negative. Also, Action, Adventure, Animation and Comedy were the genres with highest gross margin-budget proportion.