

Project: Visualizing Movie Data

Complete each section. When you are ready, save your file as a PDF document and submit it [here](#).

Step 1: Data Cleanup and Attribute Selection

- Clean up any missing information and choose the most important attributes you will explore further in your visualizations.
- List out the attributes (or variables) you plan to dive further with your visualizations. You should explore no more than 8 attributes.

I have chosen the following attributes and variables to illustrate my plots:

- ✓ Revenue Adjusted
- ✓ Budget Adjusted
- ✓ Popularity
- ✓ Runtime
- ✓ Calculated field of "Profit"
- ✓ Calculated field of "Efficiency Ratio"

- Please refer back to the [Data Cleanup course](#) to help you clean up your data.

Step 2: Tableau Visualizations

- Please make sure you follow the [rubric](#) and include Tableau Dashboards, Stories, and the appropriate visualizations (small multiples, scatter plot, bar chart, etc..) your reviewer expects your visualizations to contain. Remember: You need one Dashboard for every question (Q1-Q4) and in addition, you also need one Story, pertaining to a question of your choosing.
- Attach your visualizations as Tableau Workbooks in a zip file along with this report.

IMPORTANT: Please upload the workbooks to **Tableau Public** to allow reviewers to access your workbooks. Note that simply saving your file as a ".twbx" is not enough to allow all reviewers to access. [Instructions on how to do this](#).

Step 3: Questions

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

To understand how movie genres have changed over time, we needed to prepare the given data and define several attributes to explore in depth.

First, I have looked at the number of movies produced each by genre. The small multiple plots represent the different genres and the number of movies in each genre over the period. Genres like Comedy, Drama, Horror, and Thriller have increased dramatically.

Second, we created a plot for Average Budget adjusted and Average Revenue adjusted over the period.

The calculated field of “Efficiency Ratio” is also used here. In almost every one of the genres, we see an undesired uptrend meaning the movies in that genre are not as profitable as they used to be.

Lastly, I have brought a plot of Average Popularity for each Genre over time ([link](#)).

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

To compare the two production companies of Paramount Pictures and Universal Pictures, I have used a few attributes we have created before.

First, I have used the Efficiency Ratio, this time as defined by the ratio of Revenue to Expense. Paramount Pictures has not been very efficient in reducing costs and increasing the revenue over the year as we can see from the obvious downtrend since late 1980s. Universal Pictures on the other hand, has had a good streak of positive years when they have reversed the downtrend into a rather strong uptrend. ([link](#)).

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

The profit as an attribute is calculated and then added to the scatter plot as the size of the dots. The bigger the dot the more profitable the movie is.

The movies based on novel have outperformed the rest of the movie in almost every attribute we presented here. They are about 80% more popular on average and they have produced about 80% more revenue on average compared to the ones with original scripts. However, the movies based on novels on average are about 60% more expensive to produce.

In summary, the movies based on novels have outperformed the rest of the movies by 13% on average in terms of Efficiency Ratio ([link](#)).

- What is your additional question that you proposed? What is the answer? How did you come up with this question?

The question I proposed was to know Steven Spielberg and his movies. It was out of my own curiosity and I tried to answer the question by providing 2 visuals that contain several great info about Spielberg and his movies.

First, the scatter plot shows all his movies on a plot of Popularity (on x-axis) and Revenue (on Y-Axis).

The second plot provides two bar charts one for popularity and one for revenue by genre. As we can see “Action” has been Spielberg’s most popular and has generated the most revenue for the companies. ([link](#)).