

Movie Data: An Exploratory Data Analysis

CSCI 5513.50 Data & Information Visualization

Shaquitha Maruni

Dataset Introduction

The data set chosen was the movie dataset. This dataset contains information from a large sample of movies compiled from various sources. It includes metrics about the movie industry such as production budget, genre, directors, gross sales, and ratings from the platforms RottenTomatoes, The Numbers and IMDB. The dataset was explored in Excel and made into visualizations with the software tool Tableau.

Project Contributors

This project was done individually by myself with contributions from no other classmates.

Analysis #1

Question : Which movie genres receive the largest production budgets?

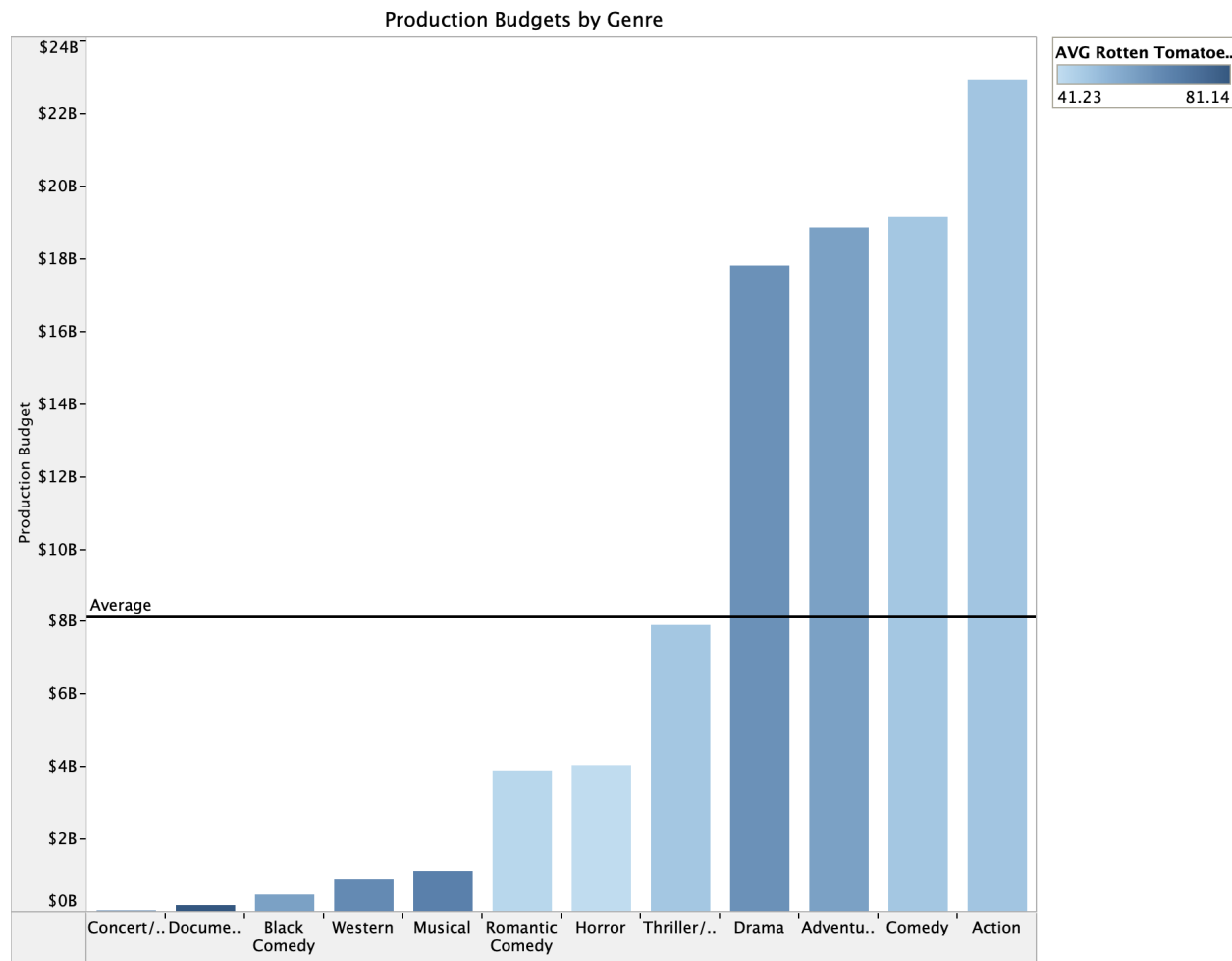


Fig. 1 Bar Chart of movie genre against total production budget

Explanation:

Figure 1 shows that Action movies have the largest total production budget of all the movie genres at over \$22 billion. Action movies are followed by Comedy, Adventure, Drama and then Thriller. The genres that have the lowest total production budgets were Concert, Documentary and Black Comedy. You can also see the average production budget is \$8 billion. After discovering Action movies get the largest production budget I wanted to see how that translated into RottenTomato ratings. RottenTomato is a trusted review website on the entertainment industry. Applying the RottenTomato ratings to the bar chart using gradient to indicate score (lighter blue hue for lower score and darker blue hue for higher scores) the Rotten Tomatoes rating for the corresponding genres actually show that the larger production budgets had reduced average tomato ratings. The best RottenTomato score actually was in the Documentary genre which actually ranked in the bottom total production budget. This is interesting in that it brings up the question if higher budgets equate better ratings. Intuitively you would expect better ratings for a larger production budget but from the visualization you can see that is not the case.

Analysis #2

Question : What genres do directors spend their total production budgets on?

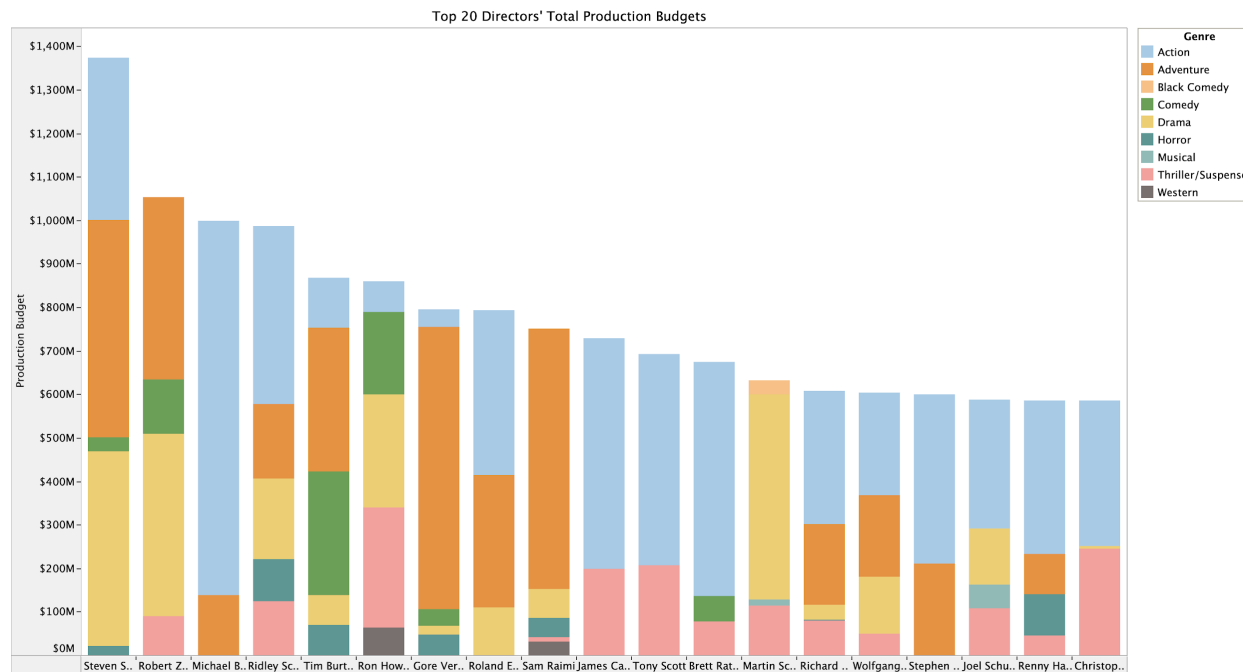


Fig. 2 Stacked Bar Chart of Top 20 Directors

Explanation:

Figure 2. is a stacked bar chart differentiated by color to represent the different genres the top 20 directors from the data set spent on their total production budgets. This visualization shows that the top three genres directors spend their production budget on are action, adventure and thrillers. Steven Spielberg comes in as the director with the largest production budget. You can also visually see that the largest total production budget is just over \$1.3 billion and the lowest production budgets are at approximately \$600 million.

Analysis #3

Question : How has Steven Spielberg's Production budget changed over time?

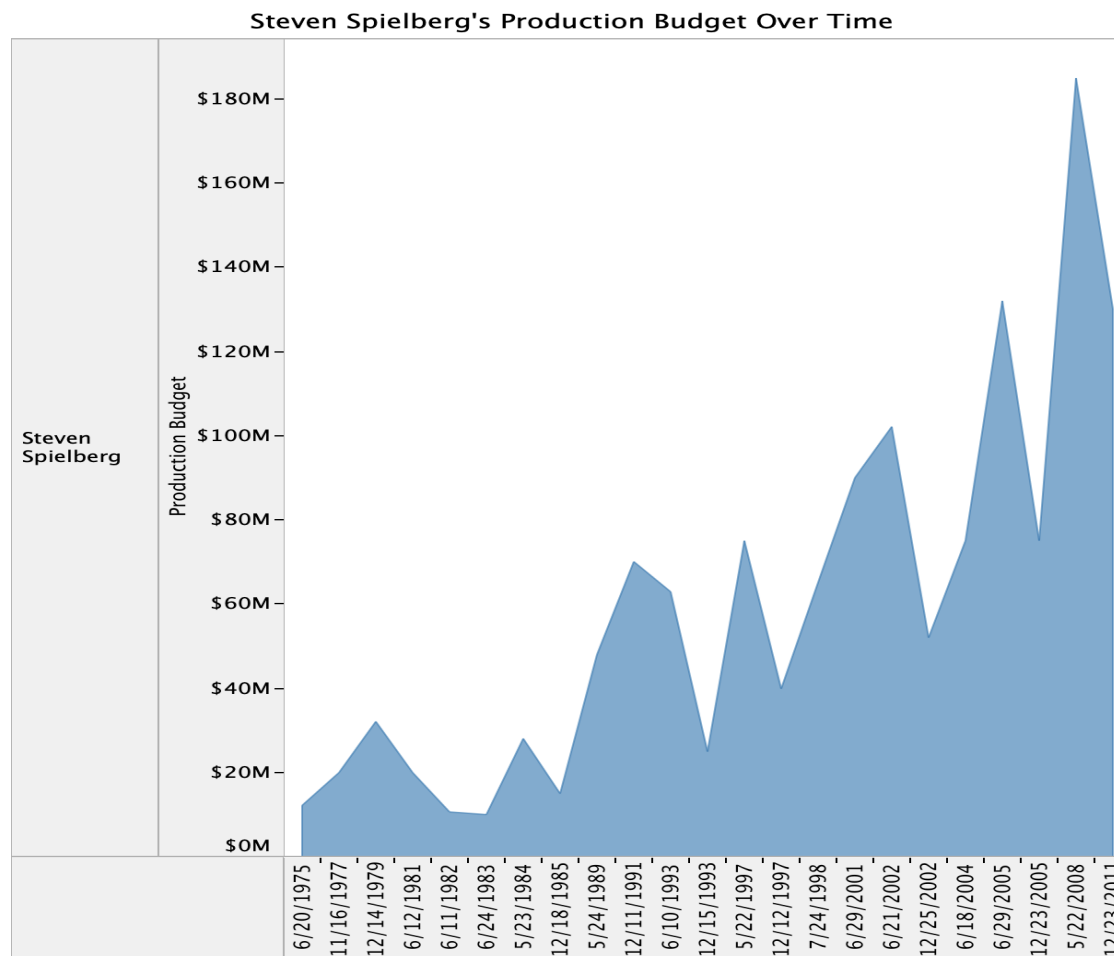


Fig. 3 Area Chart of Steven Spielberg's Production Budget Over Time

Explanation:

This visualization is an area chart that depicts Steven Spielberg's total production budget growth over time. The idea is to drill down into a single director's budget and view how they have progressed over the years. From the above we can see in 1975 his yearly budget was under \$20 million and by 2011 his production had increased leading up to his latest total production budget at approximately \$130 million. 1975 to 2011 spans a total career time of 36 years and you can see his greatest total production budget in his career was over \$180 million and his lowest total production budget was in the years 1982 and 1983 when he came in at under \$10 million dollars. The visualization also shows an upward trend in his total production budget as his career evolves so it would be safe to project that in the future his production budget will continue to grow. An interesting follow up question would be to explore what variables (genre, political climate, economy, etc.) resulted in the downward trends you periodically see in budgets over his career.

Analysis #4

Question : What is the average production budget of certain movie ratings?

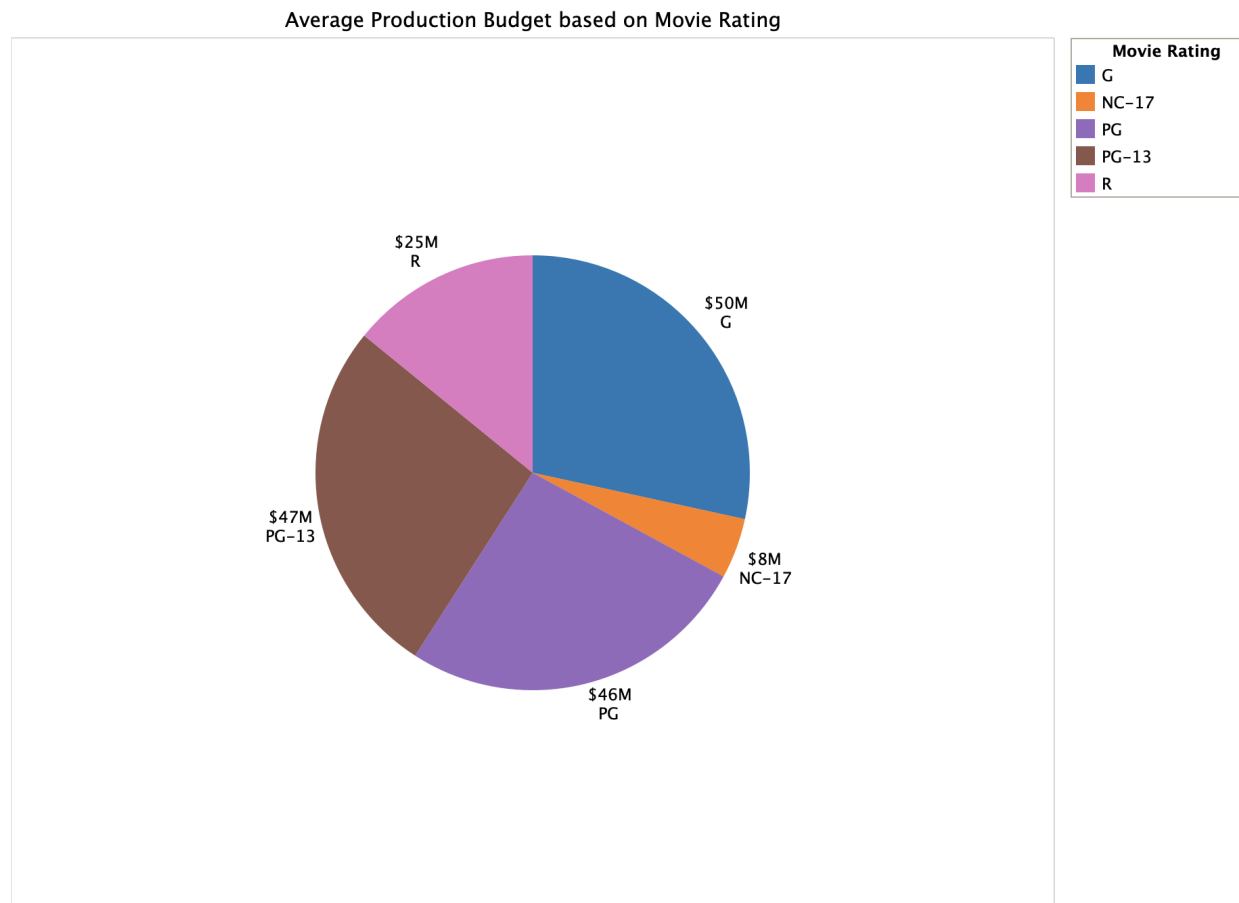


Fig. 4 Pie Chart of Movie Rating Average Production Budget

Explanation:

Figure 4 is a pie chart showing the average production budget of movies based on their ratings. This image illustrates how PG-13 movies have the largest average production budget at \$47 million while NC-17 rated movies have the lowest at \$8 million. Overall the majority of money spent is diverted to productions that cater to younger crowds ranging from G to PG-13 and ultimately this accounts for close to three quarters of all budgets.