# Visual Analytics Group Work

A screen shot of a graph

Description automatically generated

The purpose of this scatter plot is to explore the relationship between production budget and US gross revenue of the film across various genres. A logarithmic transformation has been applied to both axes to better visualise the distribution of the dataset. The scatter plot uses colour to represent genre and sizing of the marks represents the IMDb rating. The aim of this analysis is the observe whether budget and genre correlate with a movie’s financial success in the US.

The visual effectively utilises position, scale, colour and size to represent 4 important aspects of the movies dataset. Displaying US gross revenue on the x-axis and production budget on the y-axis using a logarithmic scale is effective for financial data; by condensing the wide range of values across the dataset, the visualisation makes it easier to detect patterns and trends across genres with diverse budgets. Through colouring the marks, it allows clear visual distinction between the genres, allowing readers to see clustering patterns ain budget and revenue. Through sizing by IMDb rating, with larger circles indicating higher ratings, we have added a nuanced layer of information that allows the viewers to see relationships between audience reception and financial success of the movie.

While the visuals experience design trade-off, like overlapping data marks due to high density, particularly among genres with similar budgets and revenue ranges. However, the use of logarithmic scaling, genre colour coding, and size differentiation provide sufficient visual cues to identify general trends and patterns within the dataset, even in the more crowded areas of the visual.

The scatter plot uncovers several insights about genre-based patterns in budget, revenue, and IMDb ratings. Action, adventure, and comedy films, represented in green, light blue, and orange show a broad spectrum in both production budget and gross revenue. Many of the high grossing films fall within these genres, particularly those with high production budgets, suggesting their significant financial appeal within the wider film industry. However, drama and documentary films, shown as yellow and brown, tend to have smaller budgets and lower earnings. However, it is worth noting that there are outliers in the drama genre that illustrate substantial revenue can be achieved even with a modest budget. Highly rated films, based on their IMDb ratings and shown as larger circles in the scatter plot are distributed relatively evenly between genres, but more frequently around higher budget movies. This implies a trend that higher-budget movies gain improved audience ratings, though it cannot be inferred that there is a straightforward correlation between budget size and IMDb rating. The logarithmic transformation applied to the axes highlight the positive association between budget and gross revenue, yet shows instances when lower-budget films achieve impressive box office numbers, particularly in the horror genre. The findings suggest that many genres can thrive financially even with limited production budgets, highlighting individual’s unique preferences and contrasting audience demands.

A screenshot of a graph

Description automatically generated

The purpose of this boxplot is to understand the distribution of IMDb ratings across movie genres, while incorporating financial metrics by superimposing through colour and sizing. The colour of each mark behind the boxplot indicated US gross, using a red-amber-green scale with red representing low gross and green indicating higher, each mark is also sized based on the production budget of the movie. The design tries to reveal relationships between a movie’s genre, audience rating, and financial performance. The aim of this analysis then is to offer insights into how the genre and budget of a movie may impact the commercial reception and its financial success.

The choice of visualisation can be well justified for the purpose of this analysis. With genre on the x-axis, and IMDb rating on the y-axis, the visualisation forms a boxplot that displays the spread and central tendency of ratings within each major genre. The arrangement easily shows the comparison of median rating, interquartile ranges, and any outliers across the genres. With colour being used to show US gross, this provides an additional dimension to the visual, highlighting the financial performance across the genres, so that viewers can easily assess the revenue-based success of the films. Furthermore, production budget is represented by the size of the marks, with larger marks indicating larger budget, this allows a further layer of financial information that views can easily digest and gain further insight to budgetary trends across genres.

There are numerous trade-offs within the design, particularly as the plot can become clustered due to the high number of data points, and using sizing techniques within the data marks, which can obscure smaller-budget films or those with moderate ratings. The US gross variable also experiences negative skewness, which then visualise many of the marks as a red, which struggles to truly show the distribution of this across some of the genres. On the whole, the use of colour and size of marks does improve the ability to see more patterns within the data, despite the density of the data.

The analysis successfully visualised several trends within the dataset. Genres like action, adventure and comedy display a wide range of IMDb ratings with many outliers, yet their median ratings fall between 6 and 7. Black comedy, documentaries and musicals have comparatively higher median ratings, displaying a strong audience appreciation within these genres, even with generally lower budgets, shown through the smaller marks behind the box plot. Films with higher budgets are scattered across all genres, however genres with higher US gross (green and yellow marks) are more prevalent in adventure, action and comedy, which shows that they perform well financially, even though their ratings vary. In contrast, the horror genre shows a much lower average rating and smaller production budgets, with very few instances of high-grossing films, highlighting the fact that while the genre does not attract high ratings or revenue, they do not cost as much to produce.