

Analysis of Movie Preferences from Rotten Tomatoes and IMDb Data

Kayla

November 22, 2024

Abstract

This project will examine the preferences of audiences and critics by merging data of Rotten Tomatoes and IMDb. Our goal is to identify trends and correlations between critic and audience ratings and to visualize these trends across a number of movie genres, release years, and other factors. Examining both critic and audience reviews could help us extract significant insights about movie reception and public opinion.

1 Introduction

The movie industry is an important part of entertainment worldwide. Reviews from audiences and critics help show how films are received. Rotten Tomatoes and IMDb are two popular platforms for movie ratings. Critics usually focus on things like acting, direction, and visuals, while audiences rate movies based on personal enjoyment. This project combines data from both sites to compare how critics and audiences rate movies and find trends or patterns.

2 Methodology

This analysis uses a combined dataset from Rotten Tomatoes and IMDb. It includes movie ratings, titles, genres, release years, and other details. The steps for the analysis are:

2.1 Data Collection

Data from the last 20 years was collected from Rotten Tomatoes and IMDb. Movies were matched by title to combine critic scores (Rotten Tomatoes) and audience scores (IMDb). The dataset also includes genres and release years.

2.2 Data Cleaning

The datasets were cleaned by removing duplicates and handling missing values. Only movies found in both datasets were kept. Outliers, like very high or low ratings, were flagged and removed.

2.3 Data Analysis

Descriptive statistics, such as the mean, median, and standard deviation, were used to summarize the ratings for both critics and audiences. A correlation analysis was performed to examine how critic and audience scores are related.

2.4 Visualization

Several plots were generated to better understand the trends in movie ratings:

- **Bar Charts:** Comparison of average runtime and the certificate for different movie genres.
- **Scatter Plots:** Relationship between runtime and audience ratings for each individual movie.

- **Time-Series Analysis:** Ratings trends over different release years.
- **Box Plots:** Distribution of critic and audience ratings across genres.

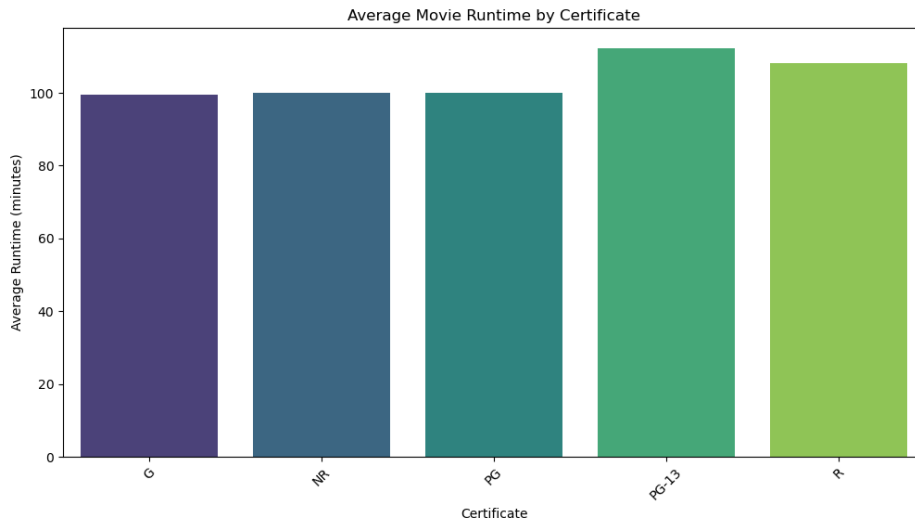


Figure 1: Runtime Correlation with Certificate

3 Analysis

3.0.1 Genre-based Analysis

It compared the average ratings that reviewers and audiences gave to different types of movies. The following results were seen:

- **Drama:** Critics generally rated drama movies higher, with an average rating of 66% compared to the audience's 60%.
- **Action:** Audience and critic ratings were about the same for action movies
- **Comedy:** The ratings were relatively similar, but critics tended to rate comedies slightly higher (around 62% on average) compared to audience scores (around 59%).
- **Thriller:** Audience ratings were generally higher for thriller movies, with an average rating of 75% compared to the critics' 70%.

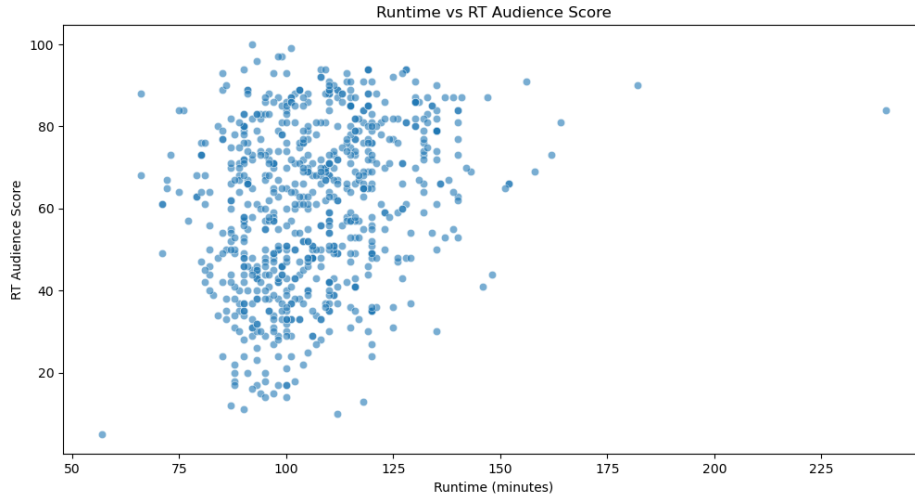


Figure 2: Runtime Correlation with Audience score

- **Romance:** Romantic movies showed a significant divergence in ratings, with critics giving an average rating of 71% while audiences rated them closer to 50%.
- **Crime:** The ratings for Crime movies were quite balanced, with critics rating them around 61% and audiences rating them around 59%.

These genre-based differences can be explained by how personal enjoyment plays a bigger role in some genres, with audiences often focusing more on enjoyment than technical details in their ratings.

3.0.2 Critic vs. Audience Correlation

A scatter plot comparing critic ratings to audience ratings showed that while the ratings are often similar, there are clear differences in specific genres, especially in action and horror films.

3.1 Trends Over Time

A time-series analysis revealed that critic and audience ratings have stayed fairly stable over the past 20 years. However, between 2010 and 2015, audience ratings saw a slight increase, likely due to social media's growing influence on public opinion about movies.

4 Conclusion

This analysis shows clear differences in how critics and audiences rate movies. Critics focus more on technical aspects, while audiences prioritize personal enjoyment. The project also highlights that genres like action have bigger gaps between critic and audience ratings. These insights help us better understand how movies are received and what influences the ratings.

5 Future Work

This analysis could be expanded upon in future studies by including elements like the director, cast, and budget of the film. Examining the effects of the audience and social media may also shed light on how ratings evolve over time. Furthermore, ratings depending on characteristics like cast, genre, and release year could be predicted using predictive modeling.

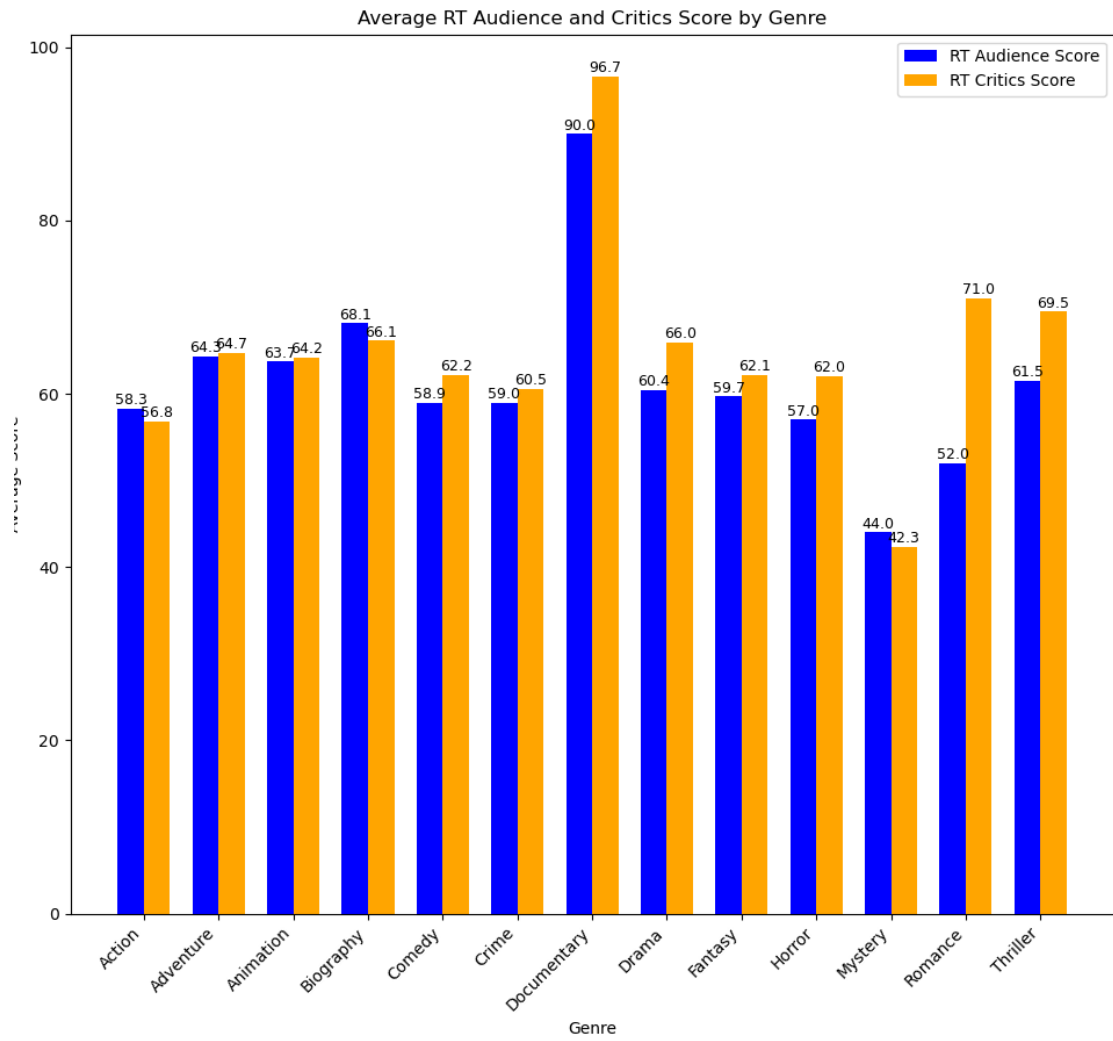


Figure 3: Average RT Audience and Critic Score Comparison

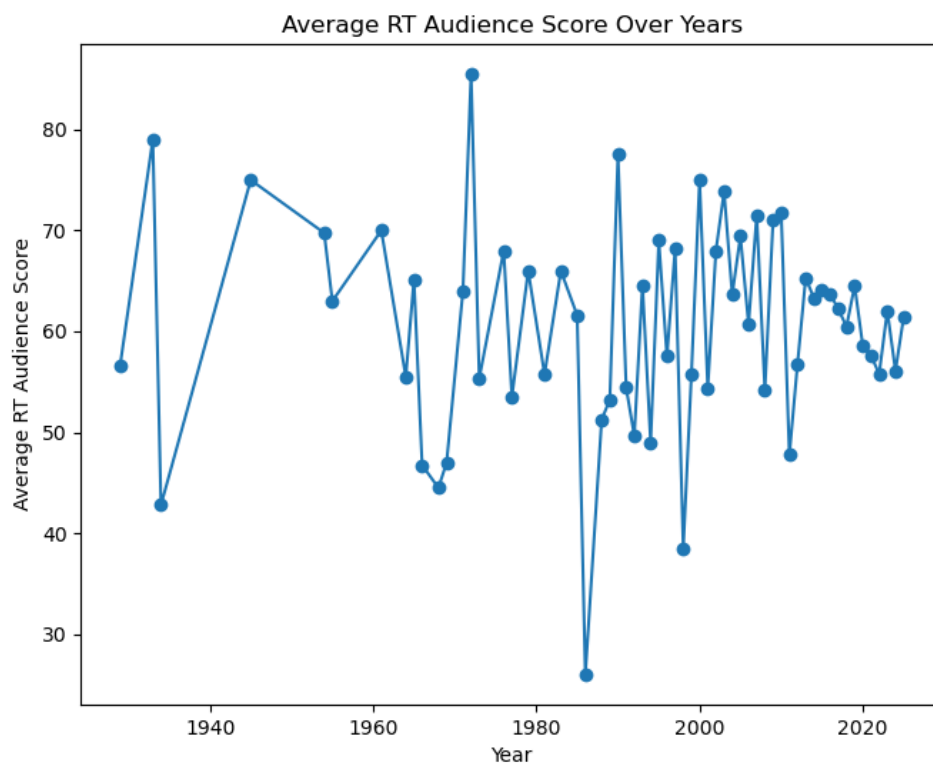


Figure 4: Trends Over Time