Microsoft Movie Analysis

Lynn Rotich

Summary

The project uses exploratory data analysis to help guide Microsoft's entry into the movie industry by identifying key success factors. The analysis revealed:

- **Top-performing genres**: Combination genres like Comedy, Documentary, and Fantasy receive the highest audience ratings

- **Optimal release periods**: Movies released in **May, June, and July** generate the highest revenue

- **Ideal runtime**: Films between **90–120 minutes** lead to higher viewer satisfaction

These insights will enable Microsoft to make data-driven decisions as they launch their new movie studio.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- Identify movie genres that generate the highest audience ratings

- Determine the most profitable months for movie releases

- Find the optimal movie runtime that maximizes both audience satisfaction and revenue

Data

Datasets from various movie industry sources, including The Numbers, IMDb and TheMovieDB (TMDb)
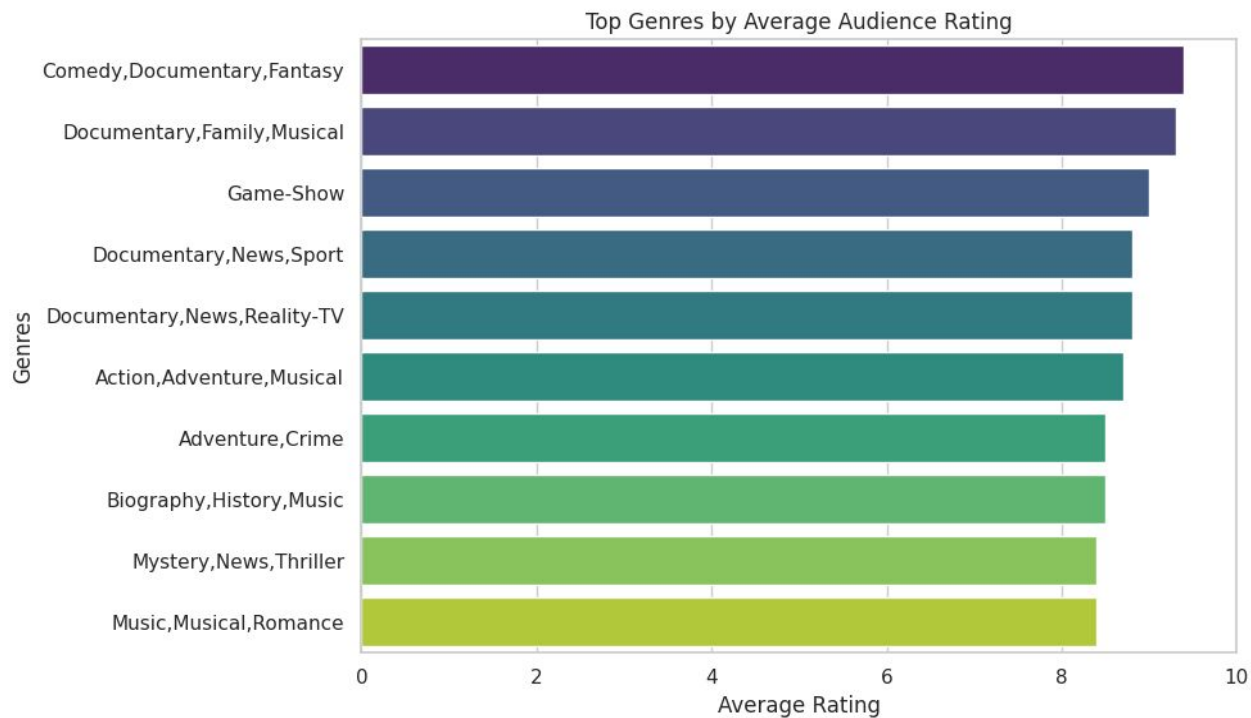
Includes movie runtimes, audience ratings, box office revenues, genres, production budgets, titles and release dates.

Methods

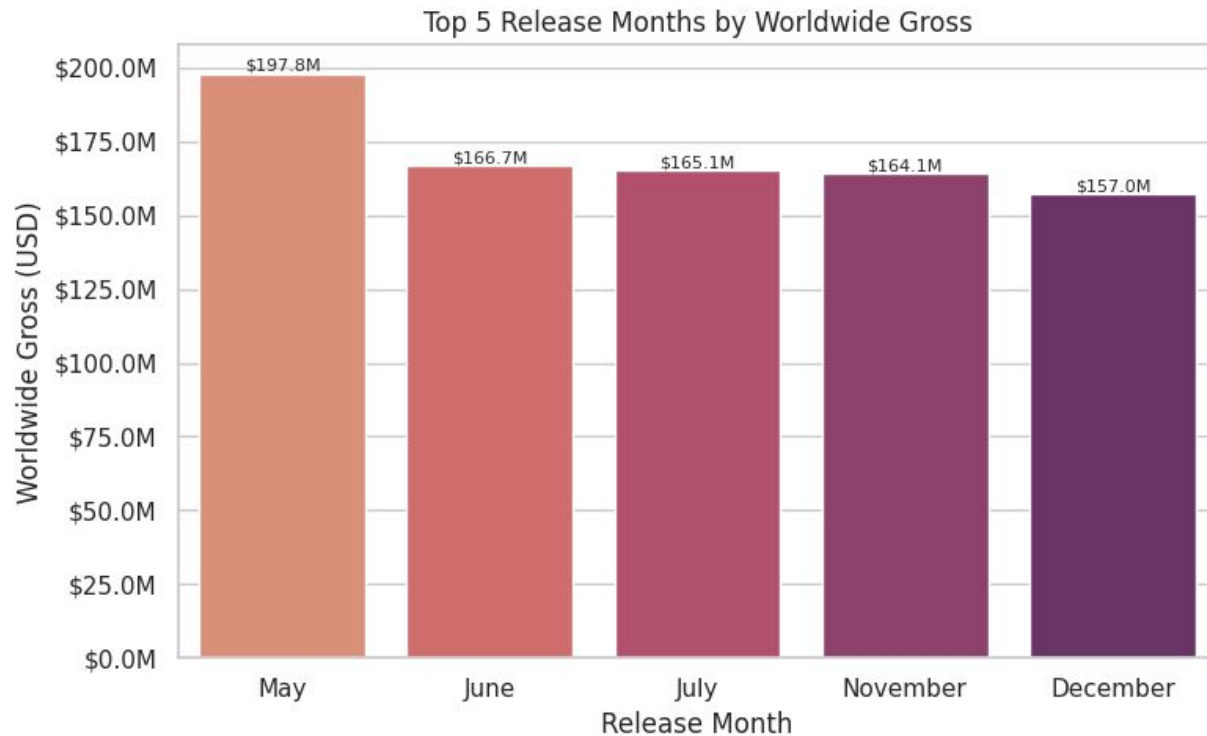Methods used to prepare the data for analysis include:

- Dropping rows with missing values
- Creating new columns for better analysis
- Handling outliers in the data

Results

Top Genres by Average Audience Rating

- Top Movie Genre is Comedy, Documentary, Fantasy

Results

Top 5 Release Months by Worldwide Gross
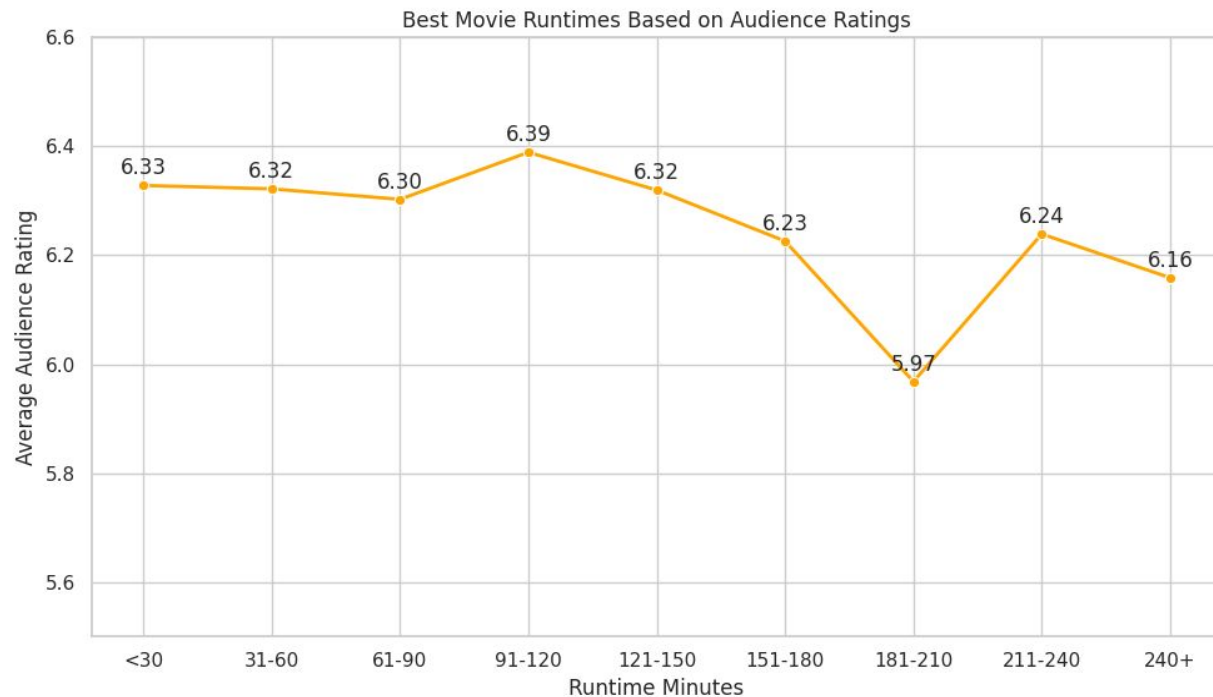
- Best months to release movies is in May, June and July based on the Worldwide gross

Results

Best Movie Runtimes Based on Audience Ratings

- Optimal movie runtimes is between 90-120 minutes based on Average Audience Ratings

Conclusions

Key Recommendations

- Prioritize top-rated genre combinations: Comedy-documentary-fantasy, Documentary-musical-family

- Target peak months: May, June, July

- Aim for 90–120 minutes movie runtime to boost ratings and satisfaction

Conclusions

Limitations

- Missing data may reduce accuracy and generalizability

- Production costs not fully accounted for

Future Recommendations

- Expand data to include **streaming vs. cinema revenue**

- Analyze **actor popularity by genre** for strategic casting

Thank You!

Email: lynn.rotich@student.moringaschool.com
GitHub: @lynn-rotich