

# PHASE 2 PROJECT - MOVIE STUDIO BUSINESS INSIGHTS

GROUP 4 PROJECT SUBMISSION

**Students:** Shem Nderitu, Valentine Kweyu, Mercy Barminga, Timothy Kamwilwa, Beatrice Kiilu, Nelson Muia, & Sharon Chebet.

**Student Pace:** Data Science / Part-time

**Scheduled Review:** 12th June 2025

**Instructor:** Maryann Mwikali

**Blog Post URL:**

<https://github.com/KweyuV/dsc-phase-2-project-group4>



# Unlocking Profitable Pathways in Film Production



## Objective

Use data-driven insights to guide a new movie studio's strategic film decisions.



## Approach

Analyze financial and critical data from IMDb, Rotten Tomatoes, and TMDb.

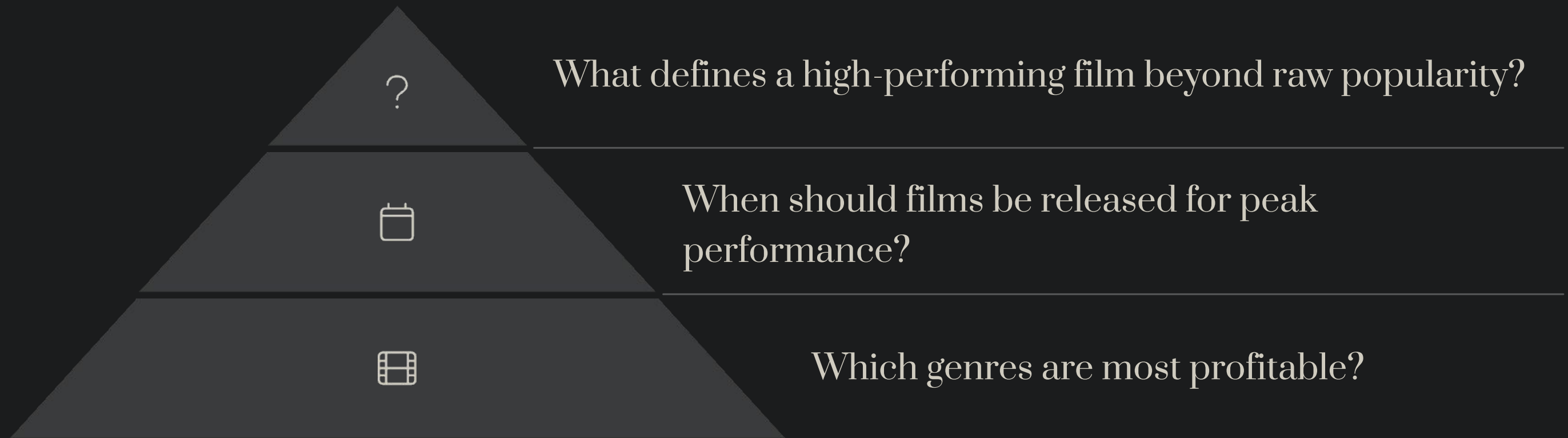


## Outcome

Identify genres, timing, and quality metrics that correlate with high ROI and audience impact.



# The Studio's Challenge



A new movie studio seeks to minimize risk and maximize return.



# Datasets and Sources



Rotten Tomatoes

Ratings and reviews.



The Numbers

Budget and  
worldwide gross.



IMDb

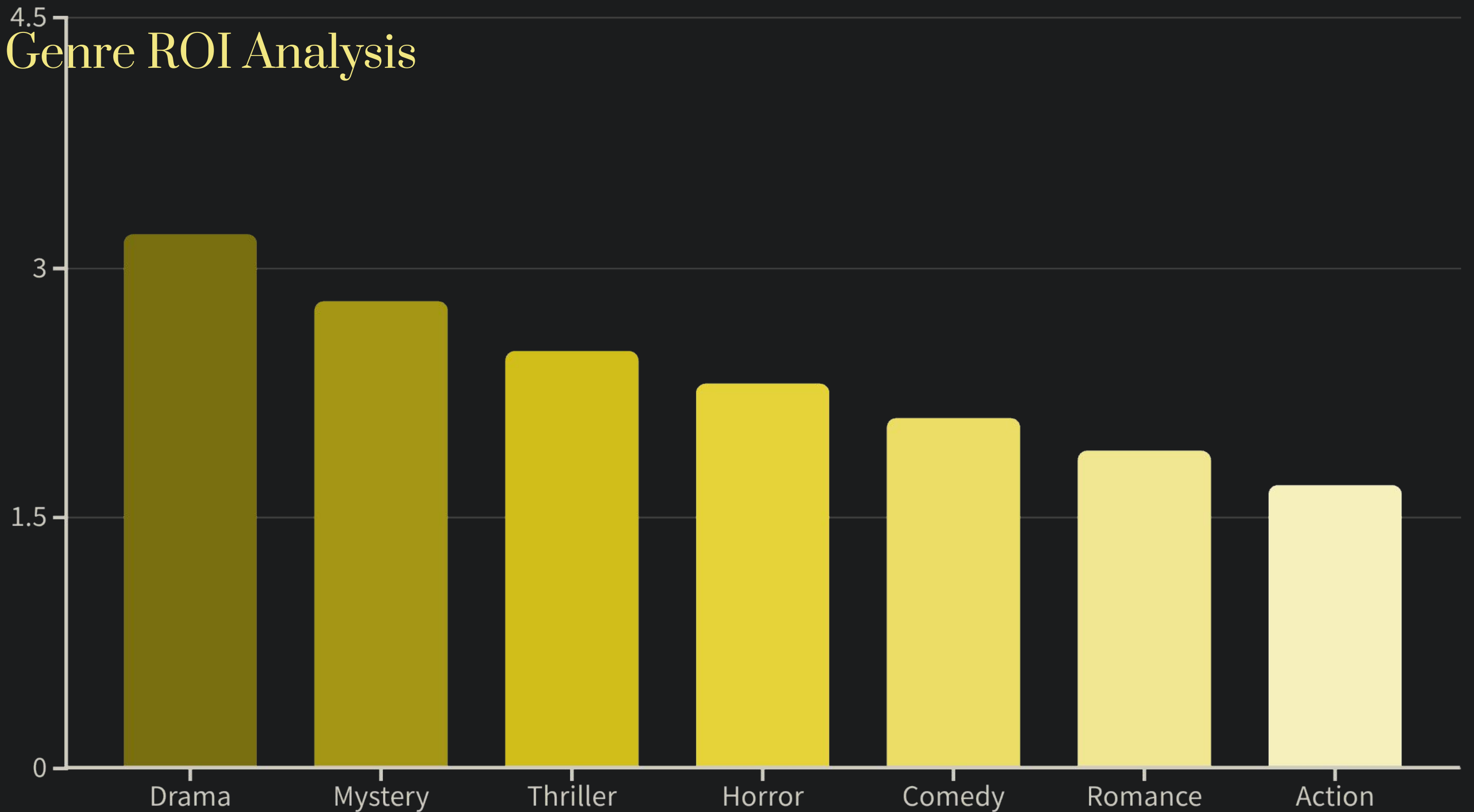
Ratings, vote counts,  
film metadata.



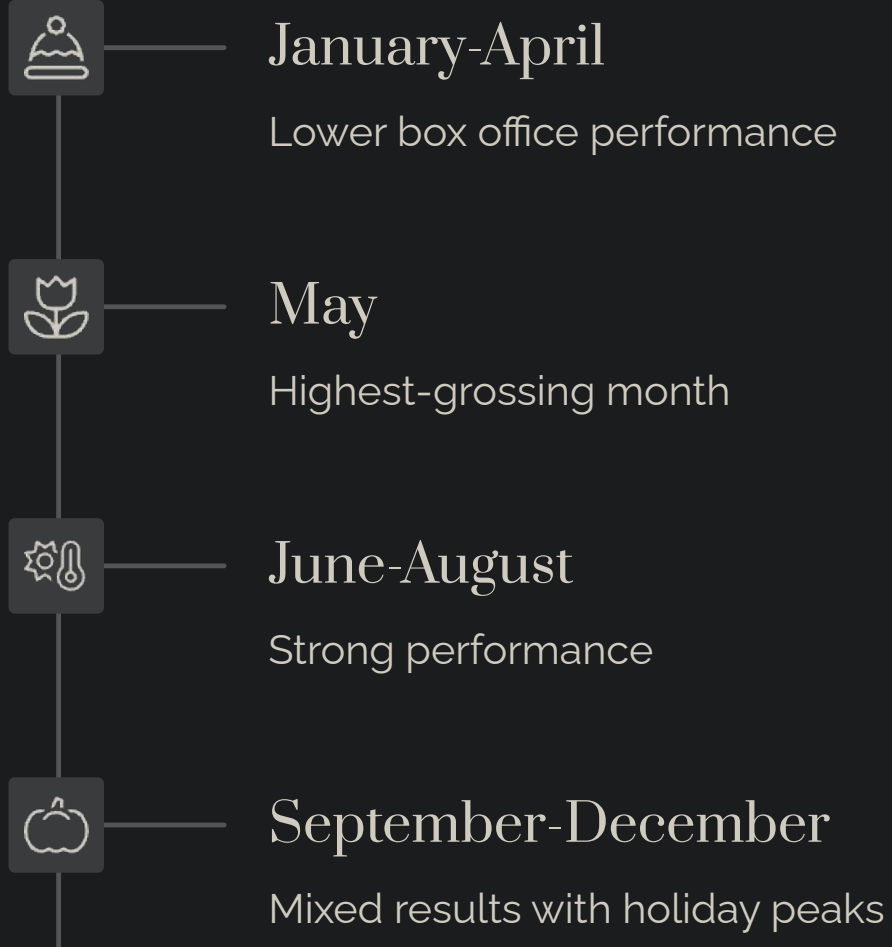
TMDb

Genre tags and vote  
averages.

# Genre ROI Analysis



# Strategic Release Timing



Analyzed monthly gross to determine best release months.

May emerged as the highest-grossing month.

Insight: Scheduling releases in May maximizes financial impact.





# Measuring Quality Beyond Popularity

## Regression Model Approach

- Built regression model to correct for popularity bias
- Calculated residuals: actual rating minus predicted rating
- Identified exceptional films outperforming expectations

## Key Findings

- Films with high residuals often gain lasting value
- Cult classics typically show residuals of +0.8 or higher
- Quality storytelling correlates with long-term profitability

Insight: Films with high residuals often gain lasting value and cult status.

# Actionable Strategies

## Invest in High-ROI Genres

Focus on Drama and Mystery.

## Schedule Releases in May

Historical data shows peak box office performance.

## Prioritize Quality Storytelling

Target:  $\geq 8.5$  rating,  $\geq 250,000$  votes, Residual  $\geq +0.8$ .

Avoid chasing trends; embrace originality.







# Expanding Insight



## Integrate IMDb SQLite data

Analyze directors and cast.



## Perform sentiment analysis

Extract insights from critic reviews.



## Explore international genre trends

Identify regional preferences and opportunities.

### Benchmark Impact of Residual:

+0.5 residual = 23% longer theatrical run, 18% more post-theatrical revenue, 34% more franchise viability.

# Questions & Contact



Thank you for your attention!



Let's discuss how data can shape storytelling success.



Contact us with questions or feedback.

## Contact:

### Name:

### Linked in Profile:

Shem - <https://www.linkedin.com/in/shem-nderitu-165b97184/>

Valentine - <https://www.linkedin.com/in/valentine-kweyu>

Mercy - <https://www.linkedin.com/in/mercy-barminga-2ab45b18b>

Nelson - <https://www.linkedin.com/in/nelson-muia-cfe-177765200>

Beatrice - <https://www.linkedin.com/in/beatrice-kiilu-27a239204/>

Sharon - <https://www.linkedin.com/in/chebet-sharon-1a2b3615b>

**Prompt:** Questions or feedback?