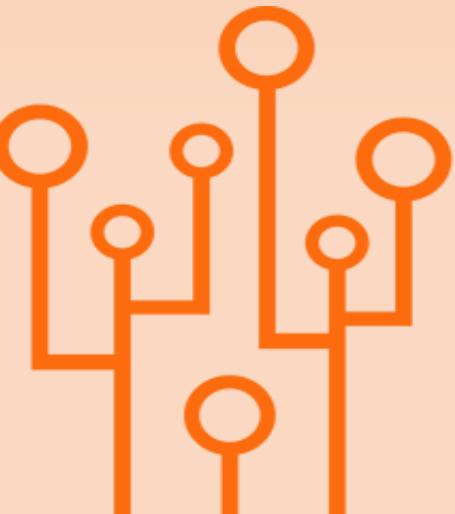


# Wakanda Studios – Entering the Movie Industry

1. Maryan Daud
2. Daniel Mutiso
3. Edwin Korir
4. Sarah Njoroge



**MORINGA**  
Discover · Grow · Transform

# Presentation outline



# Overview

- This project aims to analyze a movie dataset to identify the types of movies
- It aims to assess movies that are performing well in terms of
  - Production budget, worldwide gross, genre, runtime and rating
- We aim to identify trends, patterns and relationships
- We will use exploratory data analysis to generate insights for a business stakeholder

## Business Problem

### Background

- Wakanda Studios wants to compete with major companies in original video creation
- They plan to launch a new movie studio named Wakanda Movie Studios
- However, they lack experience in film production

### Key Business Questions

- Research top-performing box office films
- Identify key trends and successful genres
- Provide actionable insights for Wakanda Studios
- Help leadership decide on the best types of films to produce

# Objectives

- I. What is the most profitable movie genre
- II. What is the ideal production budget to produce a successful movie
- III. Which type of movie has the highest ratings
- IV. Does the movie runtime influence the success of a movie

# Data Management

## Data Understanding: Source

- Dataset was obtained from various locations;
- Tomatoes (<https://wwwrottentomatoes.com/>)
- IMDB (<https://wwwimdbcom/>)
- TheMovieDB (<https://wwwthemoviedb.org/>)
- Box Office Mojo (<https://wwwboxofficemojo.com/>)
- The Numbers Rotten (<https://wwwthenumberscom/>)

## Data Cleaning:

- Challenges encountered during data preparation included:
  - missing values, outliers and placeholders
- Data sets with Unique primary IDs were merged for ease of understanding and use

## Data Understanding: Formats

- Different files had different formats: compressed CSV (comma-separated values) or TSV (tab-separated values) files and contains information on over 30,000 movies
- Data was organized into several tables and several columns containing different information about the movies
  - The movie title, production budget, genre, runtime, and each movie having a unique ID

## Data Analysis:

- Data was analyzed using different libraries
- Graphs and Charts presented using seaborn and mat plot lib
- Data was analyzed using Python
- Data analysis was conducted in Three stages
  - Univariate analysis
  - Bivariate Analysis
  - Multivariate analysis

# Results: Summary Statistics descriptive statistics n=1,736

:

## Revenue (Domestic & Worldwide)

- Domestic Gross Mean: \$40.7M,
- Max: \$652M
- Worldwide Gross Mean: \$87.7M, Max: \$1.64B
- High variation in std indicating possible outliers showing movies perform differently

## Runtime

- Mean: 112 minutes
- Range: 82 to 165 minutes
- 75% movies of movies are Under 120 minutes

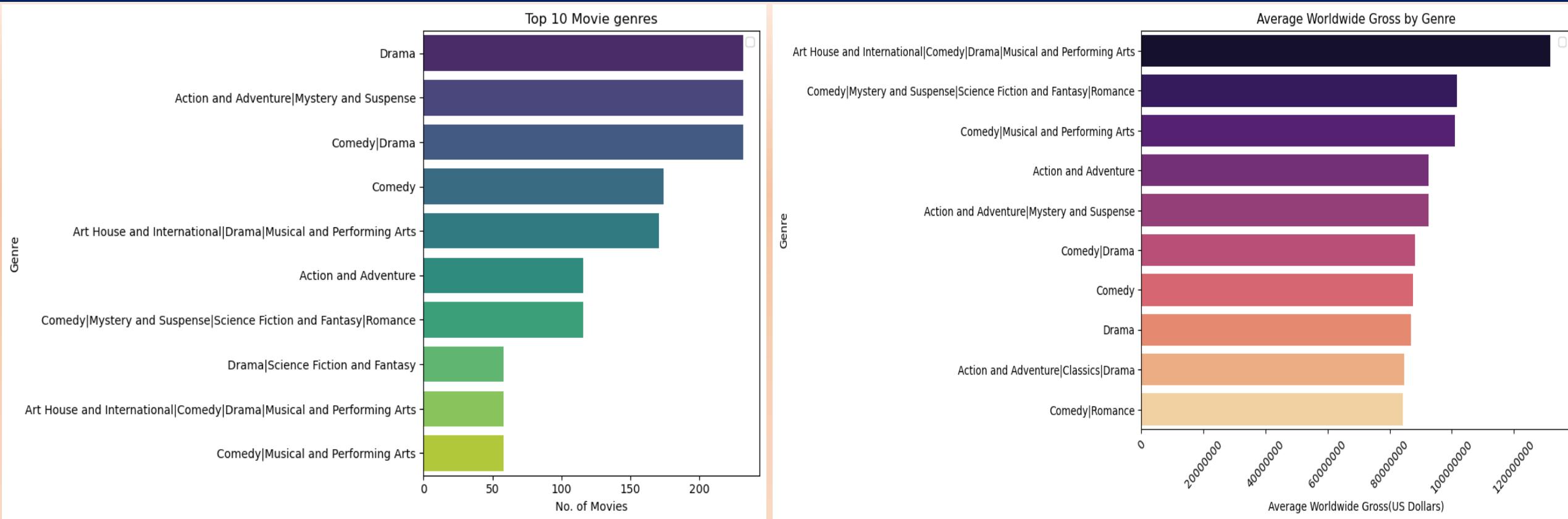
## Movie Ratings

- Mean: 4.71 (out of 10)
- Std Dev: 2.92 (high variation in ratings)
- Most movies tend to have average ratings

## Production Budget

- Mean: \$31.4M
  - Std Dev: \$41.2M
  - Wide range in budgets
- Summary statistics suggests most of the data is within range with possibility of few outliers in revenue

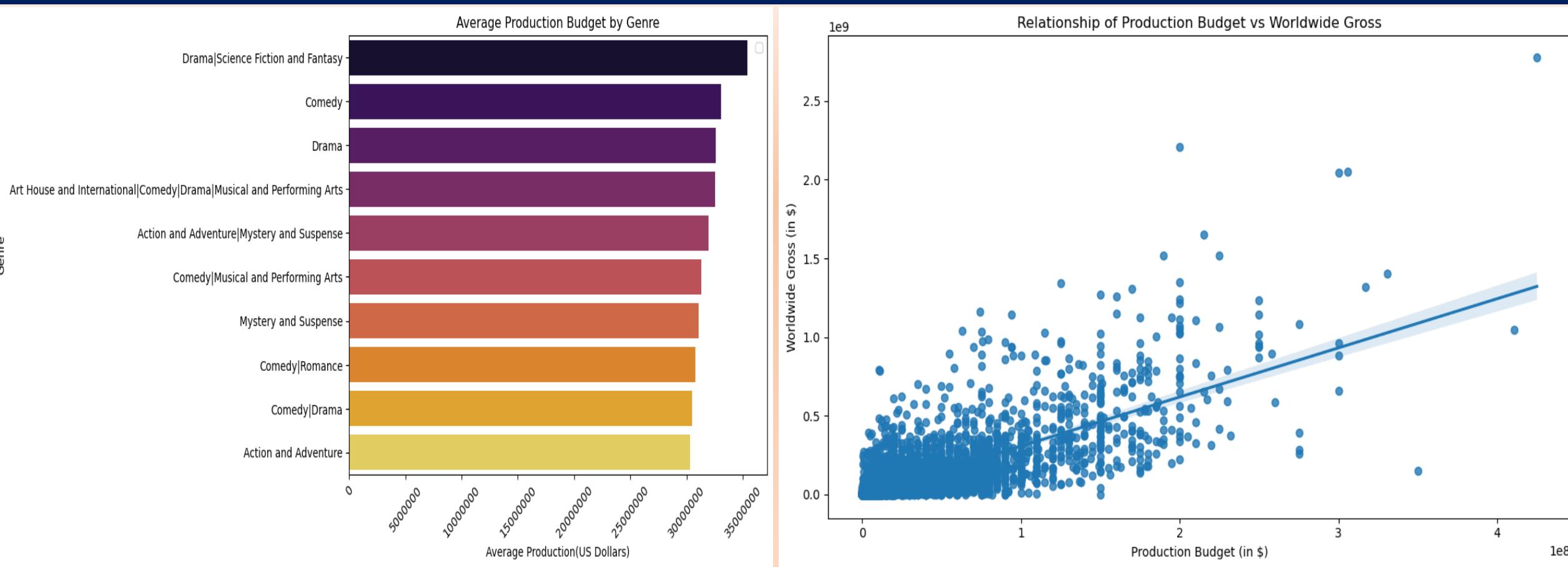
# Objective 1: Most profitable movie genre



- Most viewed movie genre: Drama followed by action and Adventure and Comedy/Drama
- Least watched is Art House and international and Comedy/Musical and performing Arts
- However this is inverse when you plot by Genre worldwide gross

- Genre that grossed highest in the worldwide scene: Art House and international and Comedy/Musical and performing Arts
- Genre that grossed Least:Drama followed by action and Adventure and Comedy/Romance

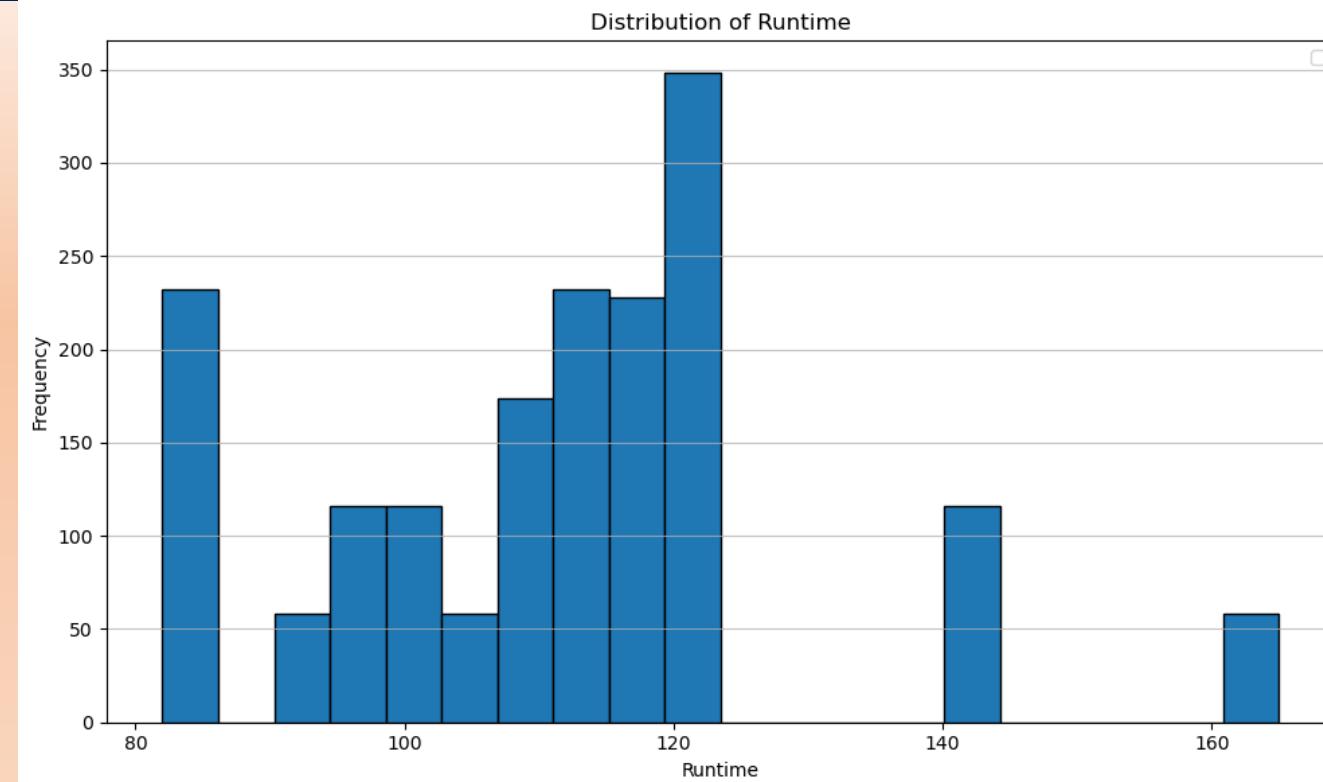
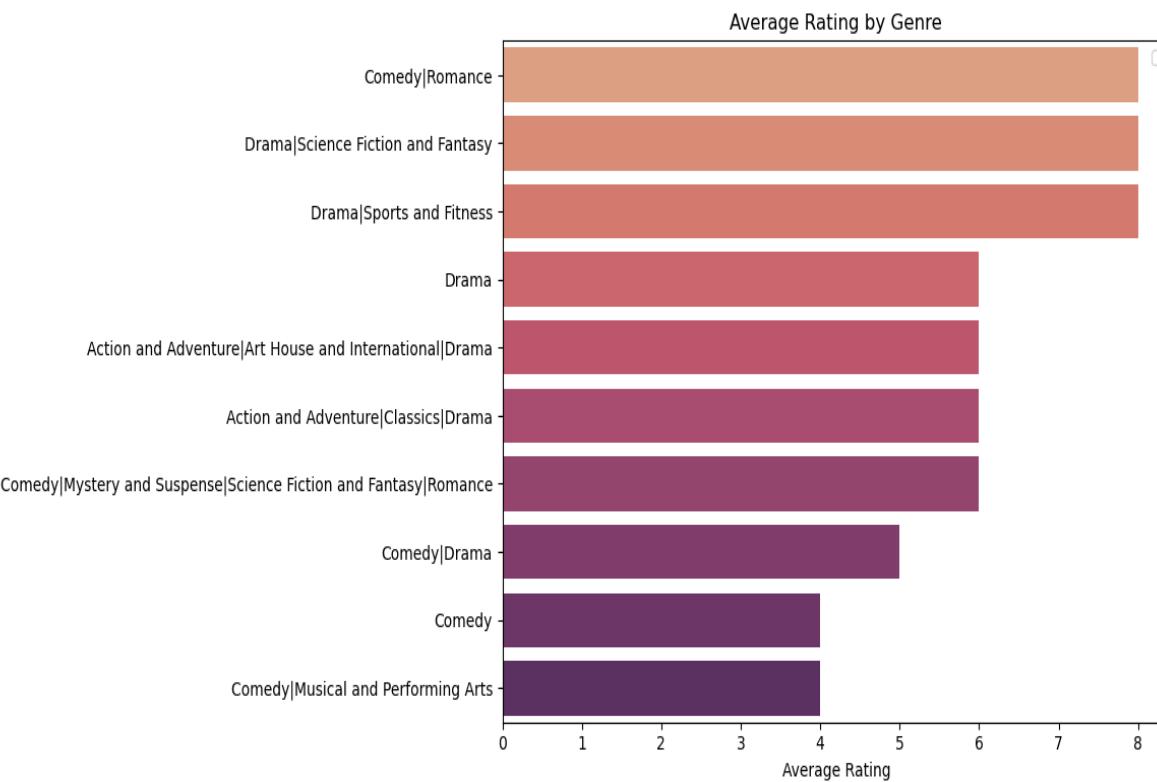
# Objective 2: ideal Production Budget to use



- Genre with highest production budget : Drama/science fiction and fantasy followed by Comedy and Drama
- Genre with least production budget: Comedy/Drama and Action Adventure (combined genre)

- This graph suggests a positive relationship between production budgets and world wide gross
- The higher the production budget the higher the worldwide gross
- Few Outliers also noted

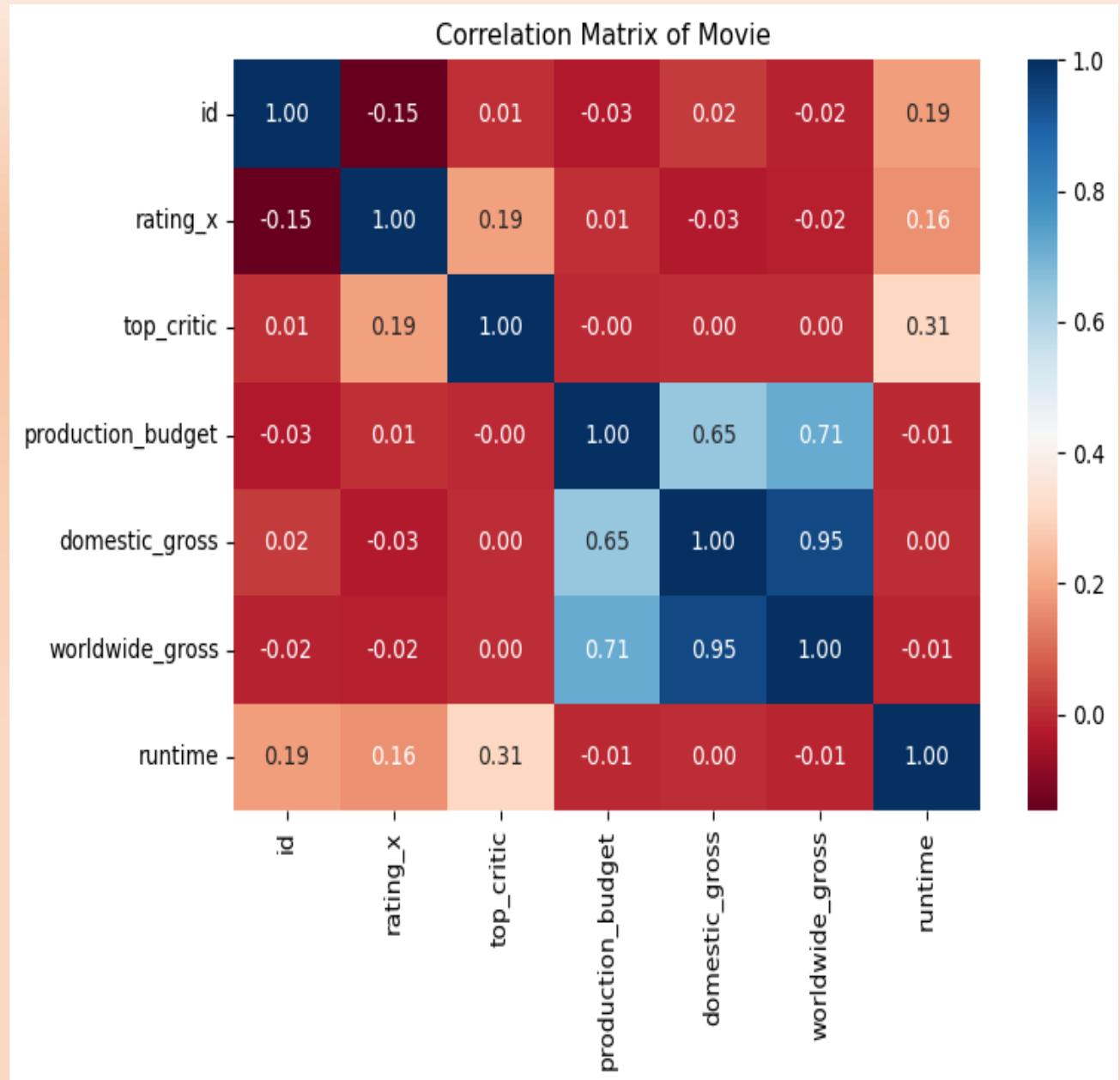
# Objective 3&4 : Highest ratings and optimal runtime



- Most movies had average ratings
- Best rated genre was comedy/Romance , drama/science fiction and fantasy and drama/sports and fitness
- Least rated genre included comedy and comedy/musical and performing Arts
- While Drama and action/adventures were moderately rated

- Seventy Five (75%) of the movies had a runtime of less than 120 minutes
- 350 of the movies analyzed had a runtime of 120 minutes
- Most movies averaged between 100 -120 minutes runtime
- Some outliers were present with runtime of 80 and 160 minutes

# Correlation matrix of Movie variables



- Strong positive correlation between Domestic gross and worldwide gross
  - Successful movies domestically are also successful globally
- Strong positive correlation between production budget and worldwide gross:
  - High budgets leads to high gross
- Weak negative correlation between rating and domestic and worldwide gross
  - Ratings don't strongly predict revenue
- Runtime has a weak correlation with worldwide and domestic gross:
  - Movie length doesn't impact revenue
- Runtime & Ratings have moderate positive correlations:
  - Longer movies receive better reviews

## Conclusion

## Recommendations

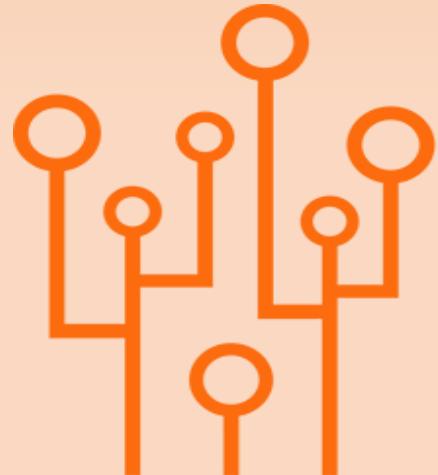
<ul style="list-style-type: none"><li>• Drama Genre has highest views and ratings however,</li><li>• Drama also has lowest revenue gross</li></ul>	<ul style="list-style-type: none"><li>• Produce movies from drama/ comedy/ action/adventure genre for better success</li><li>• Increase tickets costs for higher revenue gross</li><li>• Reduce movie piracy in this genres after production</li></ul>
<ul style="list-style-type: none"><li>• Combined movie genre require less production budgets than single genres</li></ul>	<ul style="list-style-type: none"><li>• Focus on production of Combined genres, specifically combined genre with highest views such as Comedy/Drama/Adventure</li></ul>
<ul style="list-style-type: none"><li>• Art House and international and Comedy/Musical had highest worldwide and domestic gross</li></ul>	<ul style="list-style-type: none"><li>• Produce Art House and international comedy/drama and Musical and preforming arts for higher revenue</li></ul>
<ul style="list-style-type: none"><li>• The High the Domestic gross the higher the worldwide gross</li></ul>	<ul style="list-style-type: none"><li>• Invest in adequate marketing and movie promotion locally i.e. movie premiers and trailers to maximize its success in the global scene</li></ul>
<ul style="list-style-type: none"><li>• The higher the production budgets higher gross both domestic and worldwide</li></ul>	<ul style="list-style-type: none"><li>• Allocate higher budgets to movie production while factoring in other aspects like genres, runtime and ratings for higher revenue</li></ul>
<ul style="list-style-type: none"><li>• Optimal movie runtime is &lt; 120 minutes</li></ul>	<ul style="list-style-type: none"><li>• Focus on Movies with a runtime of 100 -120 minutes</li></ul>

# Next Steps

Further analysis to improve prediction:

- Stay up to date with real time industry data and adjust to new concepts and ideas coming out
- Further research on movie data using other methodology's i.e qualitative data seeking opinions and observations from movie fans and industry professionals (Focus Group Discussions and key informant interviews)

# THANK YOU



# MORINGA

Discover · Grow · Transform