<> Code   ⊙ Issues   ⑆ Pull requests   1   ▷ Actions   ⊞ Projects   📖 Wiki   ⓘ Security   📈 Insights   ⚙ Se

Phase_2_Project   Public

⑆ 4 Branches   ⬥ 0 Tags

Go to file   t   Go to file   Add file ▾   Code

BrianRono7   Removed .DS_Store from tracking

229e0ee · 17 minutes ago

📁 assets	Updated README	1 hour ago
📄 .gitignore	Updated .gitignore t...	20 minutes ago
📄 Phase_2_Pro...	Key Insights and Re...	9 hours ago
📄 README.md	Updated README	1 hour ago
📄 index.html	Completed project.	1 hour ago
📄 notebook.pdf	Completed project.	1 hour ago
📄 presentation....	presentation slide a...	yesterday

About

No description, website, or topics provided.

📖 Readme
〰 Activity
☆ 0 stars
👁 2 watching
⑂ 0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors   3

BrianRono7 Brian Rono
Ivet566
chriskarta

Languages

● HTML 67.4%
● Jupyter Notebook 32.6%

Suggested workflows
Based on your tech stack

Jekyll using Docker image   Configure
Package a Jekyll site using the jekyll/builder Docker image.

Phase_2_Project

MOVIE STUDIO BOX OFFICE ANALYSIS.

OVERVIEW

This project entails a series of data cleaning, imputation, analysis, and visualization with an aim of generating insights for our company as we embark on creating a new movie studio.

BUSINESS UNDERSTANDING

The goal is to join the movie business through creation of films.The project therefore entails an analysis of the current films that are doing well at the box office with an aim of assisting the head of the company's new movie studio decide the type of films to create.

Key business questions

1 Which are the most frequently watched films?

2 Which films have the highest gross returns?

3 What factors have an impact on the frequency of film watching and the gross returns?

DATA UNDERSTANDING AND ANALYSIS

Source of Data

The dataset has been scraped from https://www.boxofficemojo.com/. It contains data from Domestic Box Office from 2024 to 2025.

Description of Data

The main data points analyzed are;

- Movie
- Total Gross
- Opening weekend gross
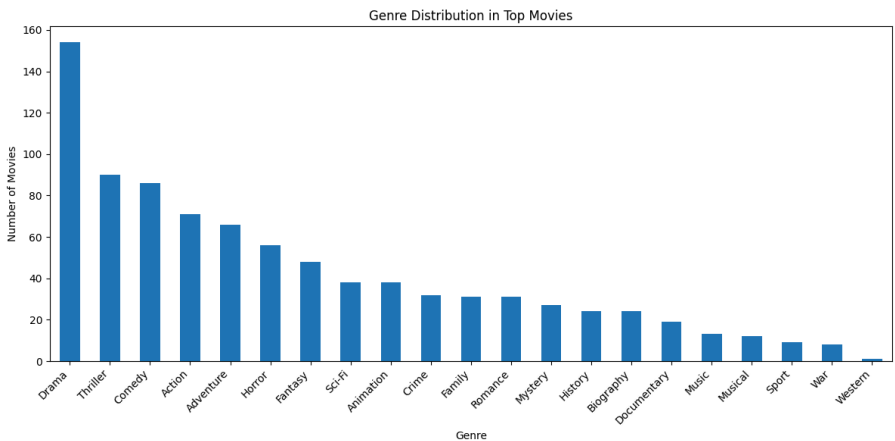- Opening weekend % of Total
- MPAA
- Running time
- Genres

📖 README   ✏️ ☰

values,i.e MPAA, Running time & opening weekend % of total.To cater for the missing values the MPAA missing points were replaced with Unknown values, while the other two had their missing values imputed by their median values.

VISUALIZATIONS

1. Genre Distribution Analysis – The number of movies associated with the different genres.

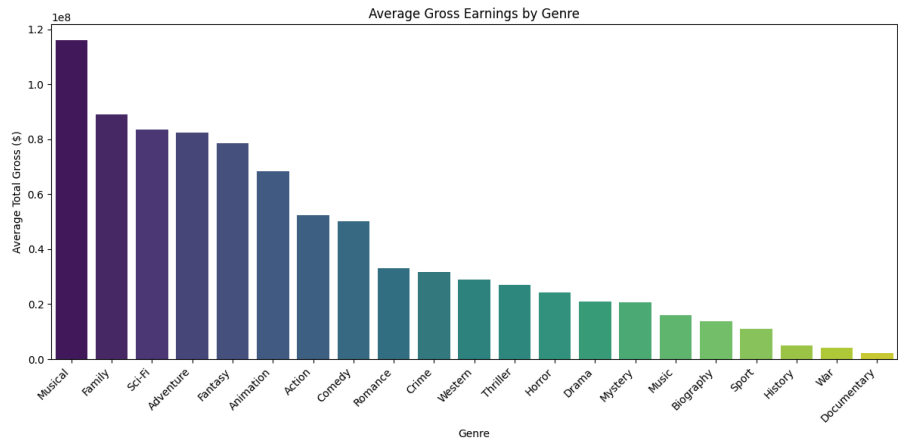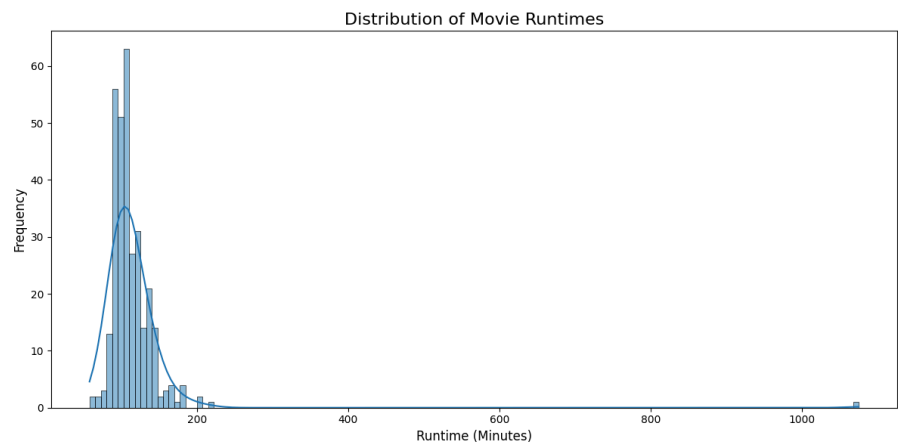2. Genre Perfomance Analysis – The average gross return per genre

Average Gross Earnings by Genre

3. Runtime Distribution

Distribution of Movie Runtimes

4. MPAA Rating Impact

Total Gross Distribution by MPAA Rating
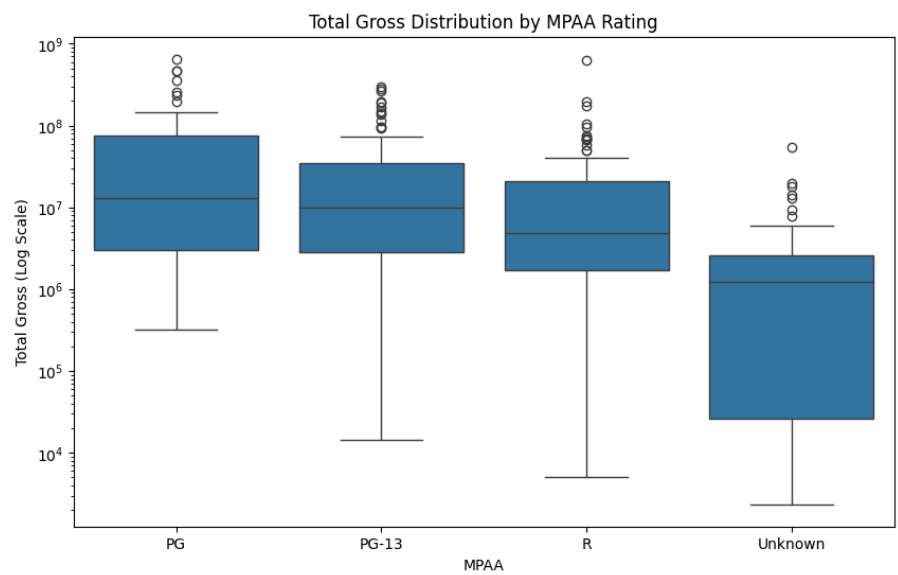
CONCLUSION

This analysis leads to the following insights and recommendations;

Genre Strategy

Analysis of 2024-2025 Domestic Box Office shows Drama, Thriller, and Comedy dominate in movie count while Musical, Family, and Sci Fi lead in Median Total Gross. Action Thriller hybrids are common, suggesting blending genres boosts popularity.

MPAA Rating Considerations

PG and PG 13 films generate the highest earnings (PG avg. 81.57 M, PG-13 42.59M), making them ideal for broad audiences. R rated films earn less (median 4.79 M), while G and Unknown ratings perform poorly. Focusing on PG/PG 13 maximizes reach and revenue.

Runtime Optimization

Optimal movie runtime for engagement and box office success is 100-130 minutes, balancing storytelling and retention. Most films run 70-200 minutes; extremely short or long runtimes are less common and risk lower performance.