ChairoSolutions / **Movie-Studio-EDA-Project**

<> **Code** ⊙ Issues ⇡↓ Pull requests **1** ▷ Actions ⊞ Projects **1** 📖 Wiki ⚠ Security 📈 Insights ⚙ Settings

👁 Watch 0 ▾   ⑂   ☆ ▾

Phase 2 Project: Exploratory Data Analysis to recommend movie types for movie studio success

⚖ View license

🧑‍🤝‍🧑 Contributing

☆ **0** stars   ⑂ **4** forks   👁 **0** watching   ⑂ Branches   ∿ Activity
🏷 Tags

🌐 Public repository

⑂ ▾   ⑂ **14 Branches** 🏷 **0 Tags**   ⑂   🏷   🔍 Go to file   t   **Go to file**   **Add file** +   Code   ···

ChairoSolutions	Remove accidental nested repository (submodule)	54a08f8 · 54 minutes ago	🕐
📄 .gitignore	Add gitignore before adding files	5 days ago	
📄 CONTRIBUTING.md	Downloaded required dataset and sup…	last week	
📄 LICENSE.md	Downloaded required dataset and sup…	last week	
📄 Movie dashboard.twbx	Tableau Dashboard	yesterday	
📄 Presentation.pdf	Saved presentation	yesterday	
📄 README.md	Merge branch 'main' into feature/movi…	yesterday	
📄 README.pdf	Complete EDA, add recommendations, …	yesterday	
📄 github.pdf	add github.pdf	yesterday	
📄 index.ipynb	Merge branch 'main' into business-und…	yesterday	
📄 movie_data_erd.jpeg	Downloaded required dataset and sup…	last week	
📄 requirements	Reproducibility	yesterday	
📄 student.ipynb	Merge branch 'main' into feature/movi…	yesterday	
📄 student.pdf	Complete EDA, add recommendations, …	yesterday	
📄 ~Movie dashboard__58352.twbr	Tableau Dashboard	yesterday	

📖 README   🧑‍🤝‍🧑 Contributing   ⚖ License   ✎ ☰

Movie Studio Market Analysis

Stakeholder

Head of the New Movie Studio

Problem Statement

Our company is launching a new movie studio and needs to decide what types of films to produce in order to maximize box office success. However, the company currently lacks historical knowledge about which movie characteristics lead to strong financial performance.

BUSINESS UNDERSTANDING

KEY QUESTIONS FOR MOVIE STUDIO STRATEGY

1. **Which movie genres generate the highest revenue?**
   By analyzing the total earnings from different genres, we can determine which types of films are more profitable and more likely to attract large audiences.

2. **Does movie rating or critic score affect performance?**
   Using ratings from Rotten Tomatoes and IMDb, we can measure whether higher-rated movies tend to earn more revenue.

3. **Does a longer runtime affect ratings or revenue?**
   By analyzing movie runtimes, we can identify whether certain runtimes generate higher earnings or better ratings across different genres.

4. **Does release timing affect earnings?**
   By examining release months and their associated revenues, we can identify the optimal times of the year to launch films for maximum success. feature/movies-EDA

Data Understanding

By loading our datasets we are able to inspect the missing values, duplicates, outliers and also understand the different columns that we need to work with to achieve the required recommmendations for the stakeholder. We use the bom.csv.gz(for revenue), imdb.db(movie genres,runtime and their ratings),tmdb.csv.gz(release_time). This is the foundation of visualization and EDA.

Data Preparation

In data cleaning, we tune our data after inspecting the datasets and merging the 3 datasets above for analysis. we filter columns , hereby remaining with only necessary columns that we will use in this analysis. We also remove the null rows from the merged dataset. Save the cleaned data set to data/processed/movie_clean.csv for reproducibility.

Evaluation

Plot A : Average Box Office Revenue by Release Year ![PLOT A](Data/Images/Average Box Office Revenue by Release Year.png)

Plot B : Average Movie Rating by Runtime ![PLOT B](Data/Images/Average Movie Rating by Runtime.png)

Plot C : Movie Duration vs IMDB Rating ![PLOT C](Data/Images/Movie Duration vs IMDB Rating.png)

Plot D :Top Genres by Average Box Office Revenue ![PLOT D](Data/Images/Top Genres by Average Box Office Revenue.png)

Plot E : Movie Runtime vs Box Office Revenue ![PLOT E](Data/Images/Movie Runtime vs Box Office Revenue.png)

Reccomendation

- Investing in Sci-Fi, Animation and Adventure offers the strongest opprtunity to maximize box office performance while supporting long term growth.
- Strong ratings improve average performance slightly, but they are not a reliable predictor of box office success and should not be used in isolation for investment decisions.
- Despite short-term volatility, the long-term upward trend in average box office revenue supports sustained investment, with flexibility to adapt during market downturns.

Plan future investments with a long-term growth mindset while remaining agile to short-term market fluctuations. Average box office revenue per movie has increased over time, indicating sustained growth potential in the industry, but periodic declines highlight the need for flexibility.

- While longer films receive slightly higher ratings on average, the impact is small and inconsistent, and runtime alone does not meaningfully influence how audiences rate a movie. Avoid using movie length as a primary lever to improve audience ratings
- While longer films tend to receive slightly higher ratings on average, the effect is **small and inconsistent**. Movie runtime alone does **not meaningfully influence** how audiences rate a movie. Focus on other factors such as story quality, cast performance, and production values to improve ratings.Avoid using movie length as a primary lever to improve audience ratings

main

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors 5

Languages

- **Jupyter Notebook** 100.0%