**Project Phase I: Planning**

Pavan Kalyan Imadabathini

Sundeep Yalamanchii

Sri Swathi Nimmagadda

Sai Manikanta Maram

IFT 533: Data Visualization & Reporting for IT

Prof: Asmaa Elbadrawy

Nov 17th, 2023

**Section 1:**

**Movie Dataset Overview:**

This dataset consists of data obtained using the IMDb Open API. The metadata files included the whole of the 45,000 movies referenced in the complete movie collection. The purpose of generating this dataset was to develop diverse recommender systems that use sentiment analysis.

This information may be used to forecast the income and performance of a movie, depending on a certain measure. It can also determine which genre of movies tend to have higher average ratings on IMDb. Additionally, it can filter suggestions for the audience by considering content and collaborative comments.

This dataset includes the following files:

The file name is **movie\_metadata.csv**. This document serves as the metadata file for the movie. The **main\_data.csv** file contains information on many features such as length, language, producer country, IMDB rating, and color. The primary data file includes the movie title, genres, director's name, and actor's name.

**Movie Dataset Attributes:**

|  |  |
| --- | --- |
| main\_data.csv | |
| Column Name | **Attributes** |
| actr\_name\_1 | Nominal |
| actr\_name\_2 | Nominal |
| actr\_name\_3 | Nominal |
| gnres | Nominal |
| title\_of\_the\_movie | Nominal |
| dirter\_name | Nominal |

|  |  |
| --- | --- |
| movie\_metadata.csv | |
| Column Name | **Attributes** |
| clr | Nominal |
| dration | Ratio |
| gss | Ratio |
| lang | Nominal |
| cntry | Nominal |
| bdgt | Ratio |
| ttle\_year | Interval |
| imdb\_scre | Ratio |
| apec\_ratio | Ratio |
| fb\_likes | Ratio |

**Section 2:**

**Movie Dataset Users:**

1. **Film Producer:**

Film producers can use this dashboard to analyze the performance of actors and directors in their previous projects. This information can be valuable when considering casting decisions or selecting directors for upcoming films.

They can also use the dashboard to track trends in movie ratings and audience preferences to make informed decisions about the types of films to produce.

1. **Movie Critic/Blogger:**

Movie critics and bloggers can utilize the dashboard to compare their reviews and ratings with those of other critics. This helps them gauge their own performance and identify areas for improvement in their critiques.

They can also track the performance of actors, directors, and genres over time, allowing them to provide more insightful commentary on industry trends.

1. **Distribution Strategist:**

Distribution strategists can use this dashboard to assess the potential success of movies in different regions or markets based on historical ratings and audience preferences.

They can also analyze the performance of movies on various OTT platforms and make informed decisions about distribution deals and release strategies.

1. **Film Industry Analyst:**

Film industry analysts can leverage the dashboard to gather data on trends in movie ratings, genre popularity, and actor/director performance.

This data can be used to provide industry reports, market insights, and recommendations to stakeholders in the film industry.

1. **Data Scientist:**

Data scientists can work with the data behind the dashboard to perform in-depth analyses, such as predictive modeling for box office success or identifying hidden patterns in movie ratings.

They can also help improve the dashboard's algorithms and data processing to provide more accurate and valuable insights.

1. **Streaming Service Curator:**

Streaming service curators can use the dashboard to identify top-rated movies and trending genres. This helps them make content curation decisions for their platforms.

They can also assess the performance of movies on their platform compared to competitors and use this data for content acquisition and promotion.

1. **Movie Enthusiast:**

Movie enthusiasts can explore the dashboard to discover top-rated movies, trending actors, and popular genres. This information can guide their movie-watching choices.

They can also use the dashboard to stay informed about industry trends and gain insights into the world of cinema.

1. **Marketing Team Member:**

Marketing team members can use the dashboard to identify movies with high ratings and potential audience appeal. This can inform marketing campaigns and promotional strategies.

They can also analyze the performance of marketing efforts by comparing them to movie ratings and audience feedback.

**Section 3:**

**User Requirements**

1. Take the necessary time to fully understand the user's objectives and the specific insights they seek to extract from the dataset. Ensure that all the questions and information sought by the user can be effectively conveyed through the visualizations and remove any extraneous or irrelevant data that does not contribute to meeting these objectives.
2. Prioritize the efficient use of data visualization by maintaining a favorable data-ink ratio. Carefully select the most appropriate types of charts or graphics to represent the data accurately and comprehensibly. Utilize pre-attentive attributes effectively to emphasize and convey the desired information clearly to any user.

**Potential Questions:**

1. Which film genre is most favored by audiences?
2. What is the highest-rated film in each genre?
3. Which aspect ratio is most appealing to viewers in movies?
4. Is there a relationship between a film's budget and its commercial success?
5. What role does social media play in influencing movie ratings?
6. Which year experienced the highest volume of movie releases?
7. How much has each country invested in movie production over various periods?
8. Who has been recognized as the most acclaimed actor each year?
9. Is there a correlation between a movie's Facebook likes and its IMDb ratings?
10. How does the cast size relate to the budget required for a movie's production?
11. How have the content ratings for movies changed throughout the year
12. Which directors' films have the highest ratings among my preferred genres?**Section 4:**

**References:**

<https://app.mural.co/t/dvproject4830/m/dvproject4830/1699563634432/41164a00ee4d15bf041988db73e8d938b1dbbfb0?sender=u532456dadd8ca56710e31855>

<https://www.kaggle.com/datasets/bobirino/movie-metadata/data>