

Movie Studio by Microsoft Project

Insights for Movie Production Studio Wing

Summary

This project's primary goal is to assist Microsoft in entering the movie industry by establishing a new movie studio. Microsoft currently lacks the necessary knowledge and insights regarding successful film genres that attract both substantial box office earnings and positive audience feedback. The aim is to analyze movie data comprehensively to deliver valuable insights that will help Microsoft make informed decisions about the types of films they should produce.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

The real-world problem this project aims to solve is Microsoft's entry into the movie industry.

Microsoft intends to establish a new movie studio but lacks knowledge and insights into the types of films that perform well in terms of box office revenue and audience reception. The objective is to explore the movie data to provide actionable insights that can guide Microsoft in deciding what type of films to create.

Specifying the Data Analytic Question

- To create insight on the types of films that perform well in terms of numvotes
- To create insight on audience reception on different types of movies
- To create insight on which directors have highest numvotes and average ratings
- To provide any other actionable insights to guide Microsoft on which films to create

Data

The provided dataset contains information related to movies and individuals involved in the movie industry. Here are the relevant characteristics of the data:

1. **movie_id:** A unique identifier for each movie, likely based on IMDb database.
2. **primary_title:** The title of the movie.
3. **genres:** The genres associated with the movie, indicating the types of themes and content it includes.
4. **start_year:** The year when the movie was released or produced.
5. **runtime_minutes:** The duration of the movie in minutes.
6. **averagerating:** The average rating or score received by the movie.
7. **numvotes:** The number of votes or ratings given to the movie, which is an indicator of its popularity or reception.
8. **primary_name:** The name of an individual associated with the movie (e.g., director, writer, actor).
9. **primary_profession:** The primary role or profession of the individual mentioned, such as director, writer, or actor.
10. **person_id:** A unique identifier for each individual in the dataset, based on IMDb or a similar platform.

This data appears to be useful for analyzing various aspects of the movie industry, such as genre popularity, movie ratings, and the involvement of specific individuals in film production. It is valuable for making informed decisions about movie production strategies and understanding audience preferences.

Methods

To clean, prepare, and present the data, several steps were taken to ensure the data is usable and insightful:

1. Data Cleaning:

- **Missing Data Handling(Null values):** Any missing values in essential columns, such as 'averagerating' or 'runtime_minutes,' were addressed. Options include imputing missing values, removing rows with missing data, or flagging missing values for further analysis. We also replaced some missing values with mode, which is a common industry standard for dealing with missing data.
- **Duplicate Removal:** Duplicate rows in the dataset were identified and removed (by dropping rows) to avoid counting the same information multiple times.

2. Data Exploration:

- **Descriptive Statistics:** Basic statistics, such as mean, median, and standard deviation, were calculated for numeric columns like 'averagerating' and 'runtime_minutes.' This provides an overview of the central tendencies and variability in the data.
- **Data Visualization:** Charts and graphs, such as bar charts for genre distribution or scatter plots for rating versus votes, were created to visualize trends and patterns in the data. Visualization makes it easier to understand the data at a glance.

Results

5.2 Based on insight on audience reception on different types of movies

Here's a summary of the key statistics for the top 10 genres based on mean ratings:

1. **Comedy, Documentary, Fantasy:** This genre received the highest mean rating of 9.4, indicating exceptionally positive audience feedback. However, please note that there's only one rating available for this genre.
2. **History, Sport:** With a mean rating of 9.2, this genre also received very high ratings. It is another genre with a limited number of ratings
3. **Game-Show:** Game-Show genre has a mean rating of 9.0, which suggests strong audience appreciation. There are a total of 6 ratings for this genre.

5.3 Here are the summary key statistics for the top 10 directors based on the sum of numvotes and the mean of averagerating

1. **A Normale Jef:** Movies directed by A Normale Jef have received a total of 1426 votes, with an average rating of 7.2.
2. **A'Ali de Sousa:** This director's movies have received 55 votes on average, with a mean rating of 4.2.
3. **A Blaine Miller:** A Blaine Miller's movies have received 8 votes in total, with an average rating of 7.0

Conclusions

Based on the results obtained from the project, which include genre statistics, director statistics, and top directors by average ratings, we can draw several conclusions and make recommendations:

1. Genre Statistics (5.2):

- The analysis of genre statistics has revealed that certain genres have exceptionally high mean ratings, indicating strong audience reception.
- However, it's crucial to consider that genres with extremely high ratings often have limited data, which means these ratings may not be representative of broader trends.
- Recommendations: For Microsoft's entry into the movie industry, it's advisable to explore genres that consistently perform well and have a substantial sample size for more reliable insights.

Conclusion: Director Statistics

2. Director Statistics (5.3):

- The director statistics provide valuable insights into how each director's movies are rated and the number of votes they have received.
- Several directors stand out with exceptionally high average ratings, although some have relatively small sample sizes.
- Recommendations: Collaborating with directors who consistently produce highly rated films could be a strategic move for Microsoft's movie studio. However, it's essential to consider the director's track record in terms of both ratings and the number of votes to assess their broader appeal.

Conclusion: Top Directors by Average Ratings

3. Top Directors by Average Ratings (5.4):

- The list of directors with the highest average ratings showcases filmmakers who have achieved perfect ratings (10.0) with a limited number of votes.
- While these directors demonstrate exceptional quality, the small sample sizes may not represent widespread audience sentiment.
- Recommendations: Microsoft could explore partnerships with top-rated directors for niche or specialized projects, but it's crucial to balance high ratings with audience reach. Collaborations should be based on a comprehensive evaluation of both factors.

Conclusion: Project Limitations

Project Limitations:

- The project's conclusions are based on available data, and the quality of the insights depends on the dataset's completeness and representativeness.
- Some genres or directors may not have sufficient data, limiting the depth of analysis.
- The analysis does not account for temporal trends or changing audience preferences, which could impact future performance.

Conclusion: Future Improvement Ideas

Future Improvement Ideas:

- Continuously update and expand the dataset to include more recent movies and ratings.
- Consider sentiment analysis to understand audience sentiments and preferences in-depth.
- Explore factors beyond genres and directors, such as cast, budget, and marketing, to gain a comprehensive view of movie success.
- Develop predictive models to forecast the potential success of future movies based on various factors.
- Conduct market research to understand the competitive landscape and audience demands in the movie industry.

Conclusion: Summary

In summary, while the analysis provides valuable insights, it's important to approach decision-making in the movie industry with a comprehensive understanding of various factors, including audience trends and market dynamics. Collaboration with highly-rated directors and exploration of popular genres should be part of a broader strategy for Microsoft's entry into the movie industry.

Thank You!

Email: ragiramogusu88@gmail.com

GitHub: @ragiravictor

LinkedIn: <https://www.linkedin.com/in/victor-ragira-5811394b/>