

MICROSOFT MOVIES ANALYSIS

March 23, 2023

By Steve Githinji

Summary

- Microsoft is seeking to create a new movie studio.
- This is a descriptive analysis of films to assist Microsoft make decisions on what films to make.
- Data from IMDB, Box Office Mojo & Rotten Tomatoes.
- 'Kids and Family|Science Fiction and Fantasy' and 'Science Fiction and Fantasy' genres of films earned the most at the box office.
- Films earned more internationally than in the USA.
- Films released in June, Nov and Dec earned more..
- 'Comedy, Documentary, Fantasy' and 'Documentary, Family, Musical' genres rated highest.

For more info:

[LINKEDIN](#) | [EMAIL](#) | [GITHUB](#)

Outline

01. | Business Problem

Description of Business Problem

02. | Data

Description of relevant data characteristics

03. | Methods

Methods for data preparation, analysis, modelling

04. | Results

Results of analysis, modelling

05. | Conclusions

Recommendations, limitations, future ideas

Business Problem

Microsoft plans to dive into the lucrative multi-billion dollar film industry.

It wants to set up a new movie studio.

It has no prior experience in the industry.

An analysis is needed to explore types of films are doing the best at the box office.

Findings should translate into actionable insights for Microsoft executives.



Data

Data from 3 online movie information databases:

- **IMDB** - Data in form of SQLite database; Contains 8 related tables such as movie ratings, movie akas, director, etc.; data for over 73,000 films.
- **Rotten Tomatoes** - Data in a TSV format; Contains 11 columns and 1560 rows of individual films; Columns are rating, genre, director & box office.
- **Box Office Mojo** - Data in a CSV format with each film's studio, domestic & foreign gross & release year; Data for 3387 films but no genre category.



Methods

Methods used:

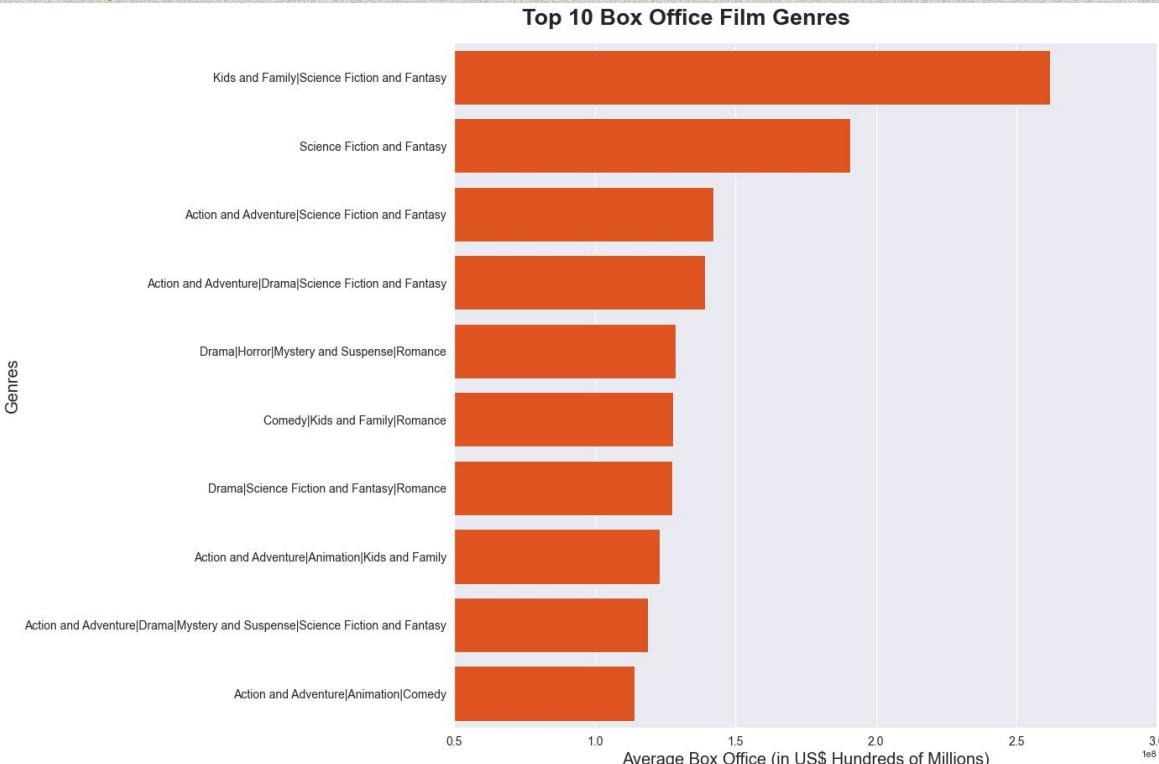
- Descriptive Analysis
- Data Visualization

Technologies:

- Python programming language
- Jupyter Notebook (uses Python and Markdown to present analysis)
- Pandas, Seaborn, Matplotlib libraries
- GitHub repository



Results

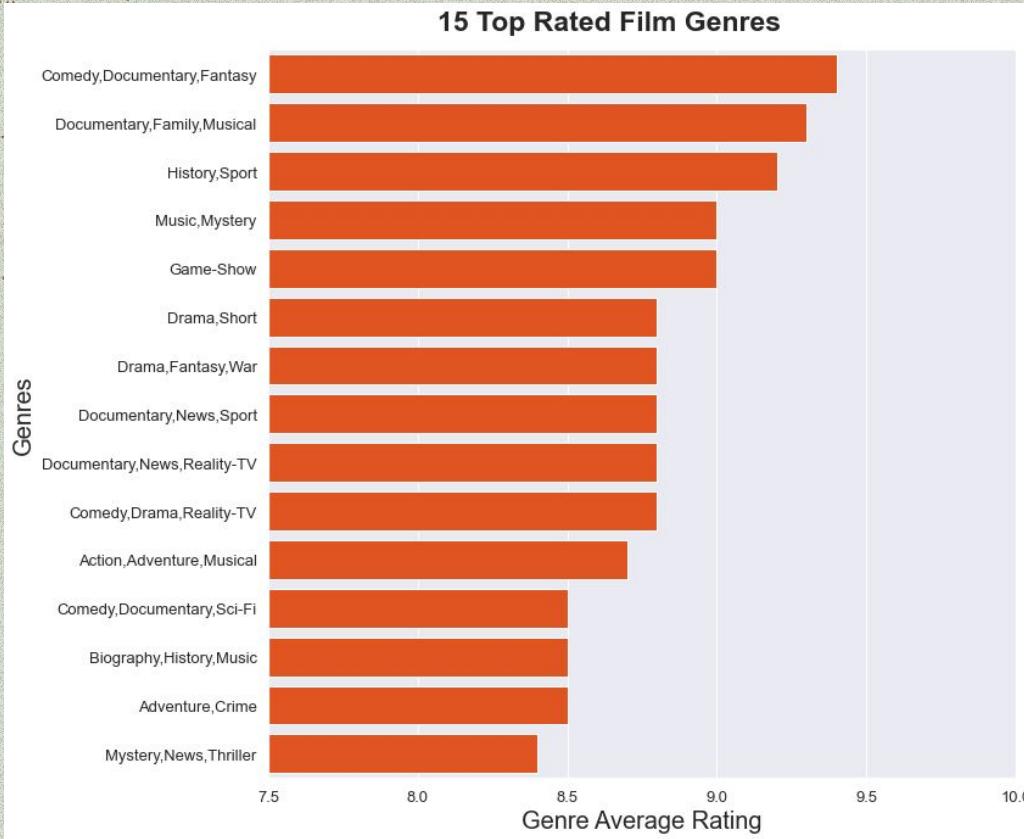


Highest earners are:

- Kids and Family|Science Fiction and Fantasy,
- Science Fiction and Fantasy, Action and
- Adventure|Science Fiction and Fantasy

All these genres have
Science Fiction

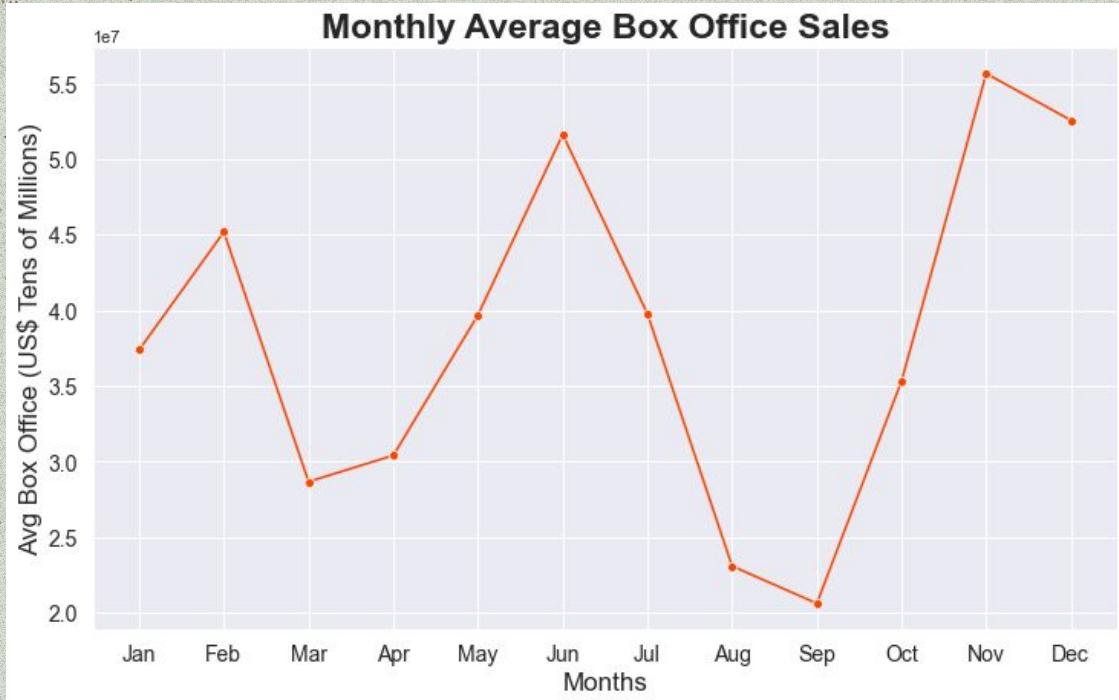
Results



Top rated genres are:

- Comedy,
- Documentary,
- Fantasy,
- Documentary, Family,
- Musical and
- History, Sport

Results



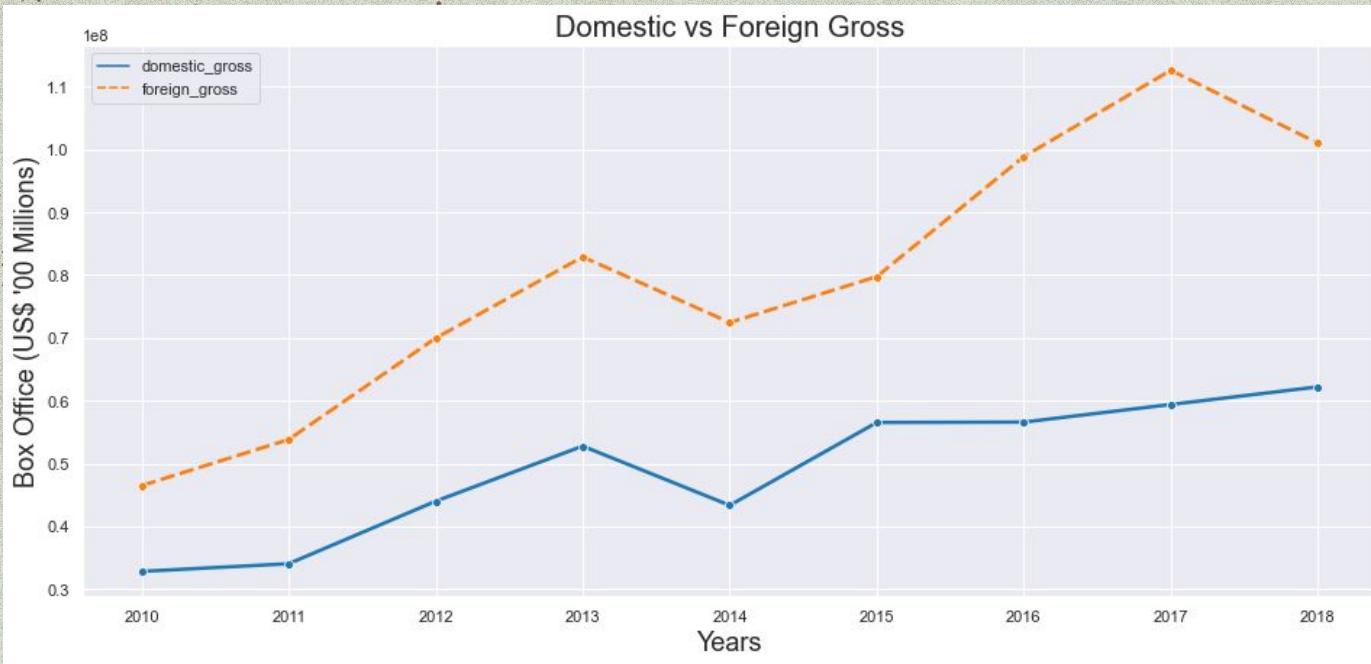
Top movie release months are:

- November
- December
- June

Worst months are:

- August
- September

Results



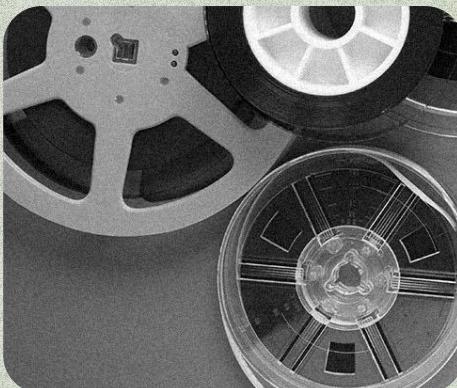
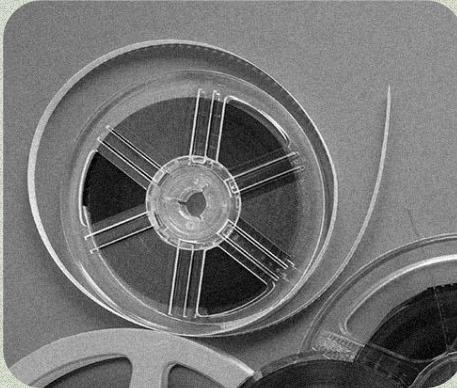
Foreign markets consistently perform better



Conclusions

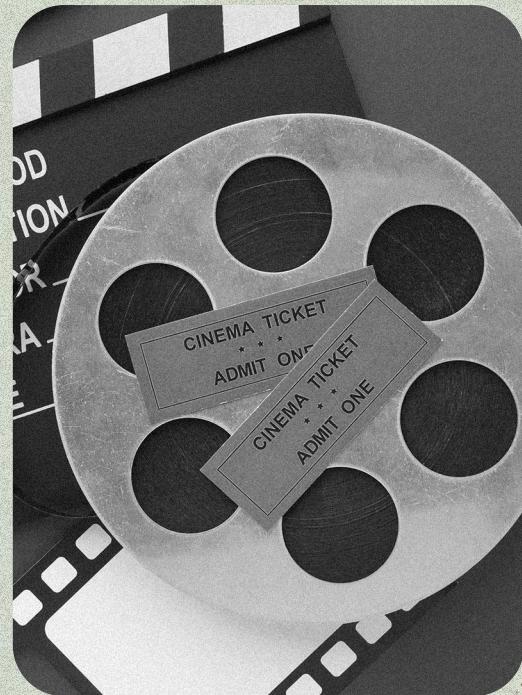
Recommendations:

- Create films under Comedy, Documentary, Fantasy, Documentary, Family, Musical and History, Sport genres for positive reviews.
- Make Science Fiction movies for high box office sales.
- Release in the months of November , December and June.
- Marketing budget should prioritize the foreign market



Further Study

- Correlation between production budget and box office gross
- Comparison between big screen releases and online subscription services.
- Analysis of best performing markets outside the United States.



- Thanks -

Do you have any
questions?

Email: githinjisteve96@gmail..com

GitHub: <https://github.com/stevegithinji>

LinkedIn: linkedin.com/in/steve-githinji-10ba0114a