

EXPLORATORY DATA ANALYSIS IN THE MOVIE INDUSTRY

Creating films that make an impact.

SUMMARY

Exploratory data analysis in the movie industry to generate insights for a business stakeholder (Microsoft) about what type of films are doing well. Data from IMDB and Box Office Mojo show that drama, action, adventure and Sci-Fi have a higher rating.

OUTLINE

- Business Problem
- Data
- Methods
- Results
- Conclusions



BUSINESS PROBLEM

Microsoft is looking to enter the filmmaking industry. My project's aim is to assist the head of it's new movie studio with insights that would help them decide what type of films to create.



DATA

In this project I'll be using movie datasets from IMDB and Box Office Mojo. The data files provide the start year, ratings, movie titles and other features like the genres and running time of the movies.



METHODS

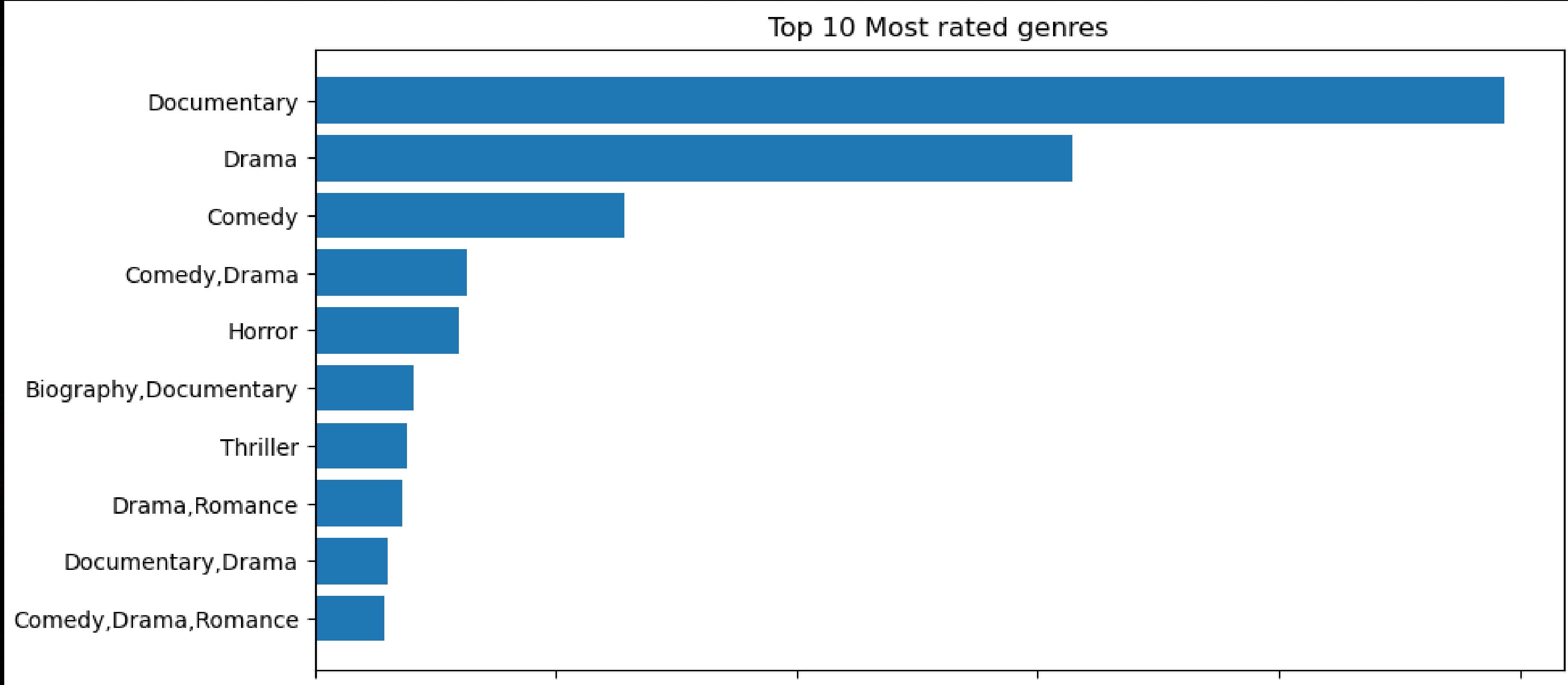
I used Exploratory Data Analysis (EDA) to explore the various fields and aspects of filmmaking. I also used data preparation and data modelling.



RESULTS

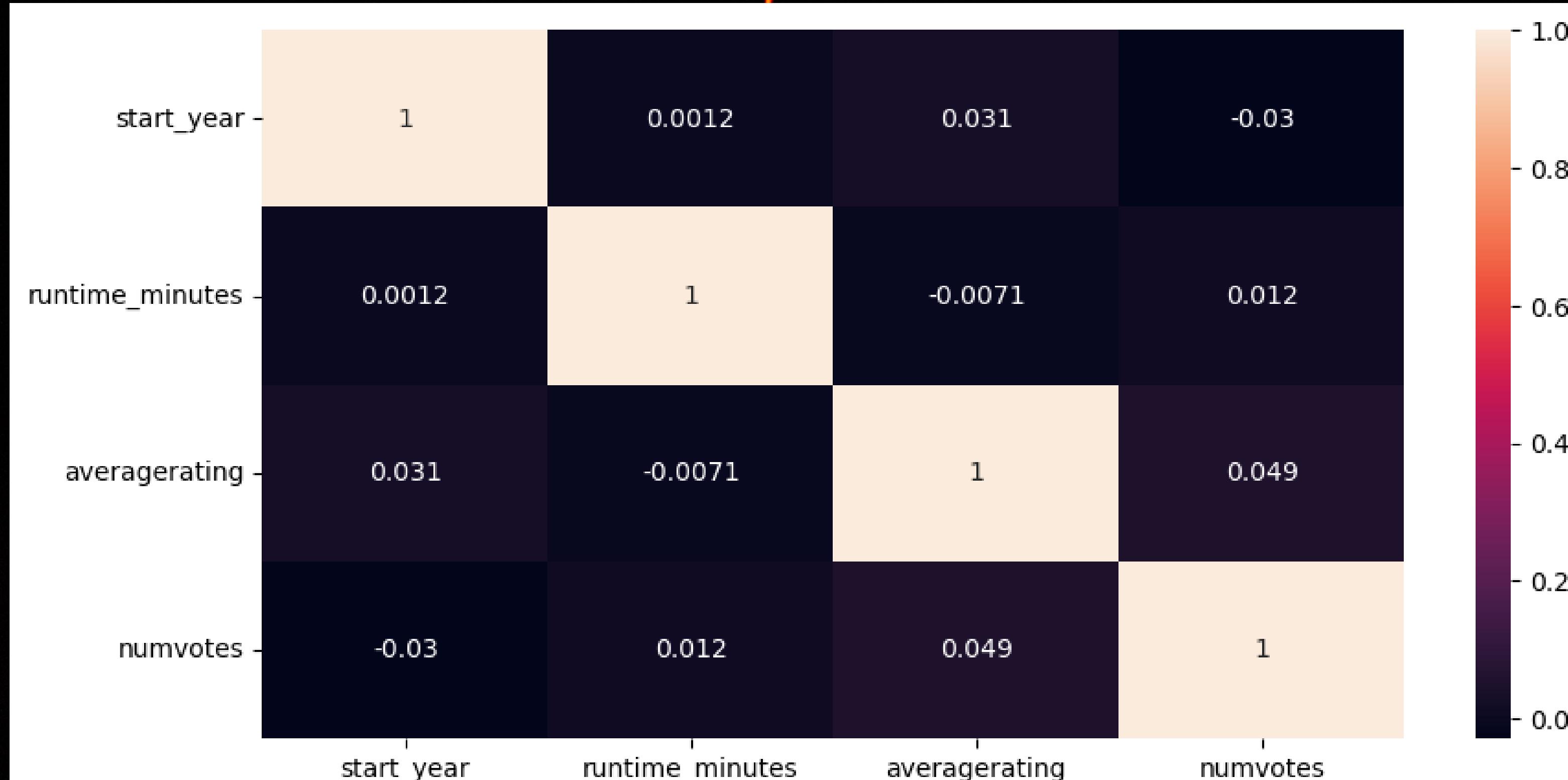
Bar Graph For Top 10 Most Voted Genres

Top 10 Most rated genres



Documentaries have the highest number of votes

Heatmap for visualization relationship between variables



There's a weak positive correlation between average rating (averagerating) and number of votes (numvotes).



CONCLUSIONS

- **Invest in making movie genres that reach a bigger audience e.g Drama.**
- **Producing movies with an average runtime of 2 hours.**
- **Examine revenue and profit margins for various types of movies**

NEXT STEPS

- Audience engagement through social media.
- Content production
- Assessing risks and challenges



“IF IT’S A GOOD MOVIE, THE SOUND COULD GO OFF AND
THE AUDIENCE WOULD STILL HAVE A PRETTY CLEAR
IDEA OF WHAT WAS GOING ON.”

- Alfred Hitchcock



Thank You!

Email : faithgitau04@gmail.com
GitHub : [FaithGitau](https://github.com/FaithGitau)
LinkedIn : [linkedin.com/in/faith-gitau-344b07106/](https://www.linkedin.com/in/faith-gitau-344b07106/)