

Microsoft Film Analysis

Author: Alex Muturi

Summary

The goal of this project is to analyse a number of datasets and derive insights/recommendations to a company that is looking to join the steadily growing film industry.

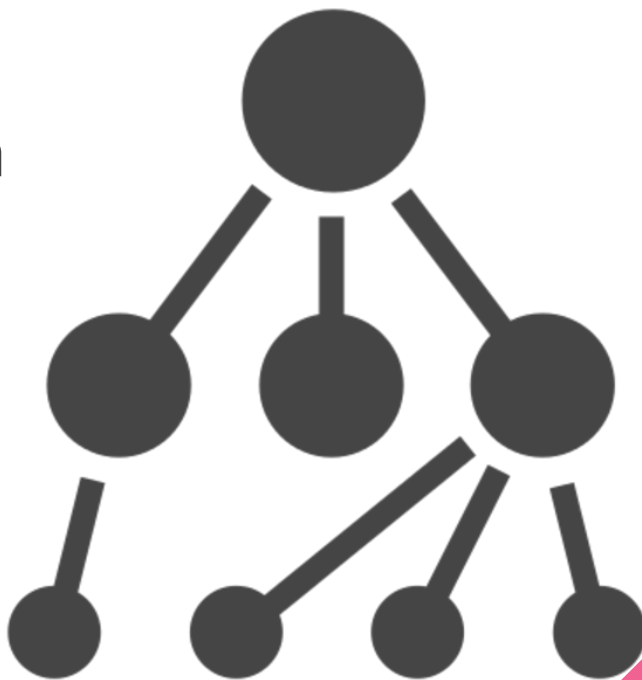
After the analysis, the results suggest that the company invests **generously** in the following genres:

- Adventure
- Action
- Comedy
- Drama



Outline

- **Business Problem**
- **Data**
- **Methods**
- **Results**
- **Conclusions**



Business Problem

The business problem was presented in the form of 3 questions:

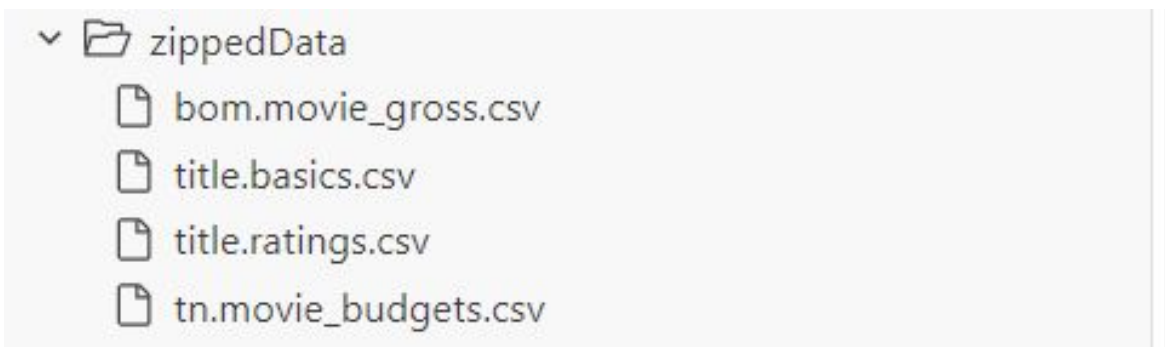
1. What are the most Popular Genres?
2. Which Genres have the highest monetary returns?
3. Does a higher production budget translate to higher returns?



Data

Relevant datasets were sourced from IMDb's developer non commercial datasets.

They are as shown below and are in the /zippedData folder:



Methods

Tools used:



Main tool used for
analysis and
manipulating
datasets



NumPy

Used when handling
outliers

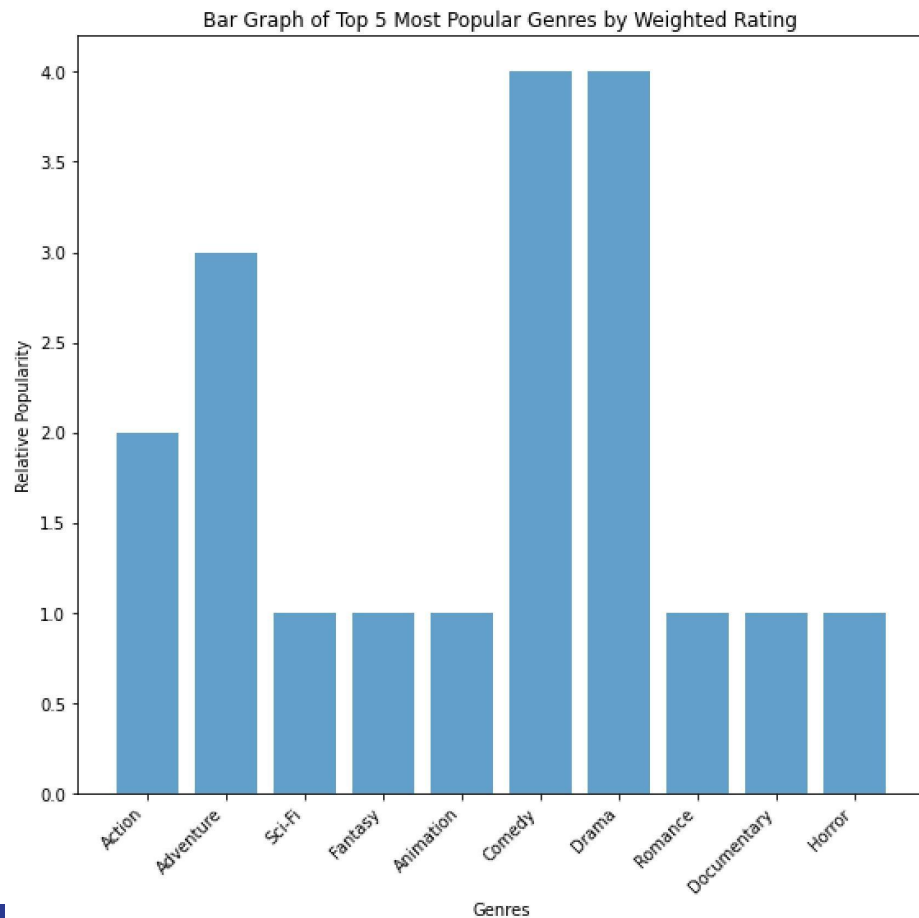


seaborn *matplotlib*

Used to generate histograms, boxplots
and other visualizations

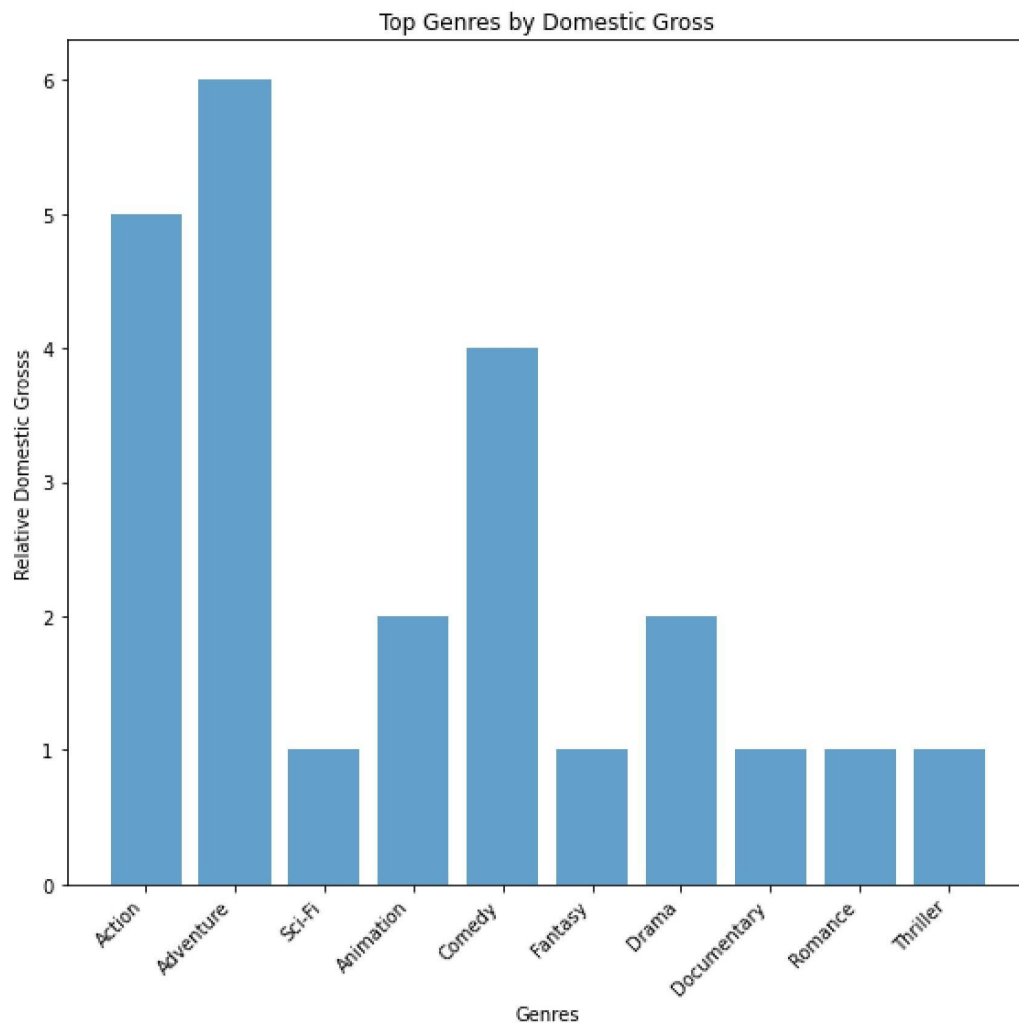
Results

**Most
popular genre
(by weighted rating)**



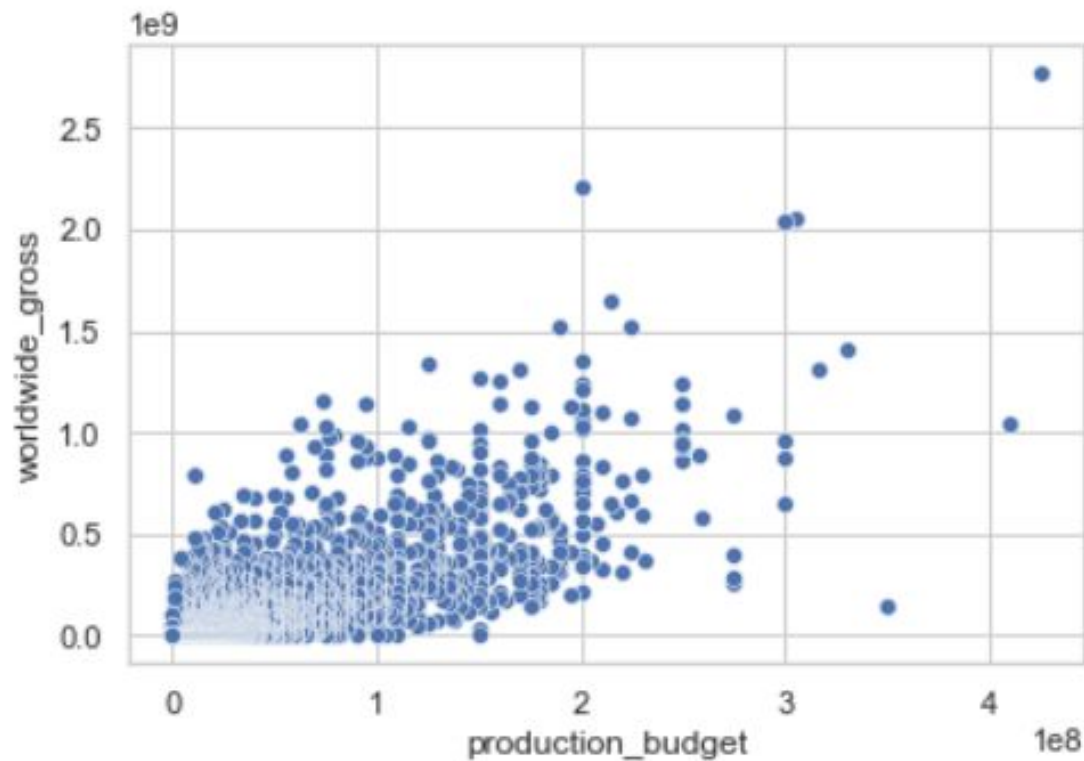
Results

**Genre with
Highest Returns
(domestic gross)**



Results

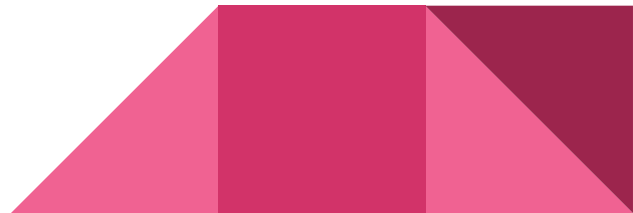
**Genre with
Highest Returns**
(domestic gross
Vs
worldwide_gross)



Conclusion

The analysis is considered a success, having answered all questions set out in the introduction.

From the results, I would recommend that the stakeholder considers creating films in Adventure, Action, Comedy and Drama genres.





Thank You!

Email: alexmuturi@gmail.com

GitHub: [@alexmuturi](https://github.com/alexmuturi)