

Microsoft Movie Venture

Sylvia Sarange Manono

Summary

- This project seeks to provide actionable insights from exploratory data analysis to inform Microsoft's decision and ensure the venture is successful and profitable.
- A movie's production budget directly correlates with its worldwide gross income.
- Preferred genres by viewers are primarily drama and comedy.
- Viewers mostly prefer films in English, French, Japanese, Spanish, and Russian.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- Entertainment and tech companies are increasingly relying on original content.
- This is due to increasing difficulty in licensing content from movie studios due to the upsurge of streaming/distribution platforms.
- Microsoft, with limited experience, seeks to establish a new movie studio division to compete with other entertainment industry stakeholders.
- There is need for data-driven insights on trends.

Data

Dataset 1

Name: tn.movie_budgets.csv

No. of datapoints: 5782

Data columns: 6

Column names:

Column names:

#	Column	Non-Null Count	Dtype
0	id	5782 non-null	int64
1	release_date	5782 non-null	object
2	movie	5782 non-null	object
3	production_budget	5782 non-null	float64
4	domestic_gross	5782 non-null	float64
5	worldwide_gross	5782 non-null	float64

Data

Dataset 2

Name: rt.movie_info.tsv

No. of datapoints: 1560

Data columns: 12

Column names:

#	Column	Non-Null Count	Dtype
0	id	1560 non-null	int64
1	synopsis	1498 non-null	object
2	rating	1557 non-null	object
3	genre	1552 non-null	object
4	director	1361 non-null	object
5	writer	1111 non-null	object
6	theater_date	1201 non-null	object
7	dvd_date	1201 non-null	object
8	currency	340 non-null	object
9	box_office	340 non-null	object
10	runtime	1530 non-null	object
11	studio	494 non-null	object

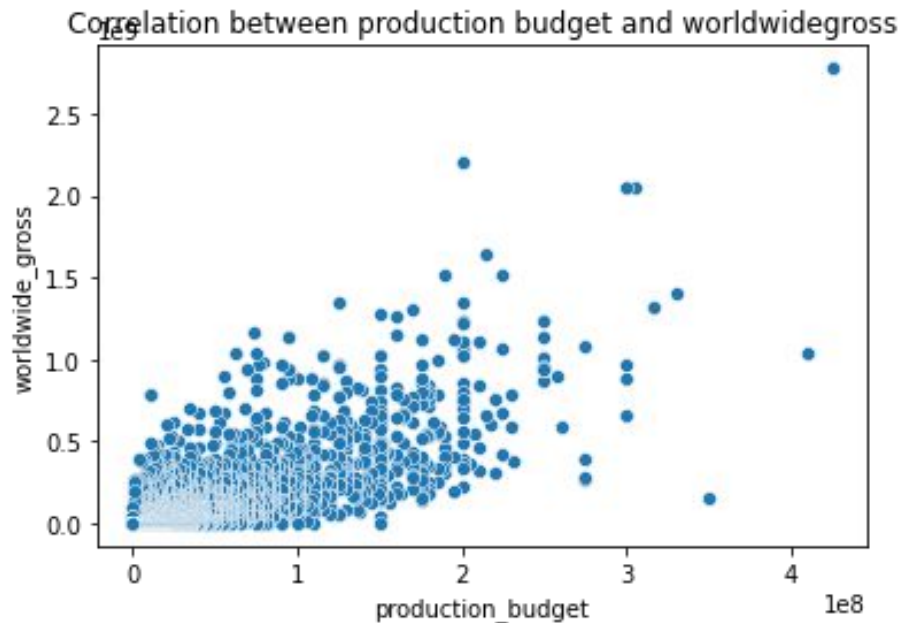
Methods

- I use exploratory data analysis to inform Microsoft's decision to invest in original content based on existing movie data from various data sources.
- The data preparation steps include:
 - Combining information on movie basics and ratings in the im.db database.
 - Confirming the absence of null values in the tn.movie_budgets.csv file.
 - Checking column data types.

Results

Production Budget and Worldwide Gross Income

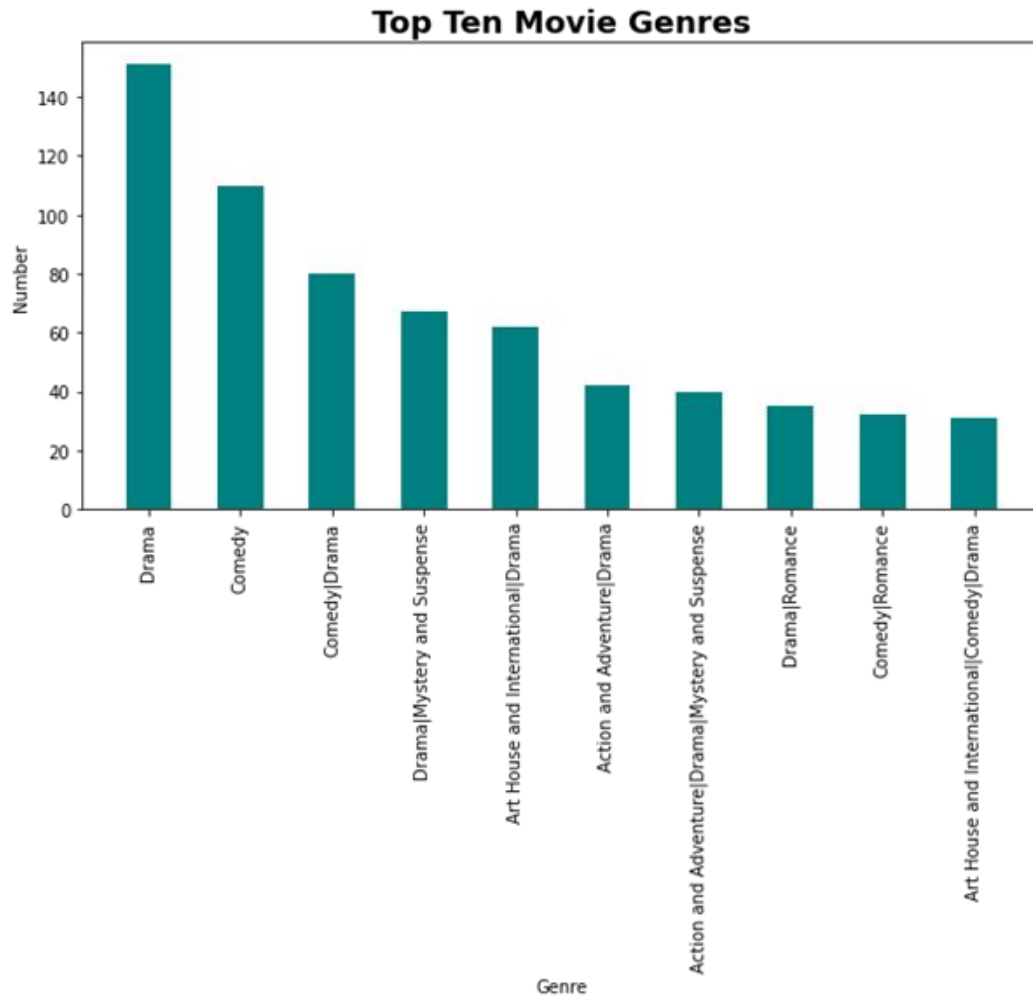
- There is a positive correlation between movie production budgets and worldwide gross income.
- Movies with the highest worldwide gross had production budgets ranging from US\$ 300 - 425 million.



Results

Top Movie Genres

- Drama is the top movie genre.
- This was followed by comedy, and varying combinations of genres.
- However, drama features in 4 out of the top 5 genres (and combinations).

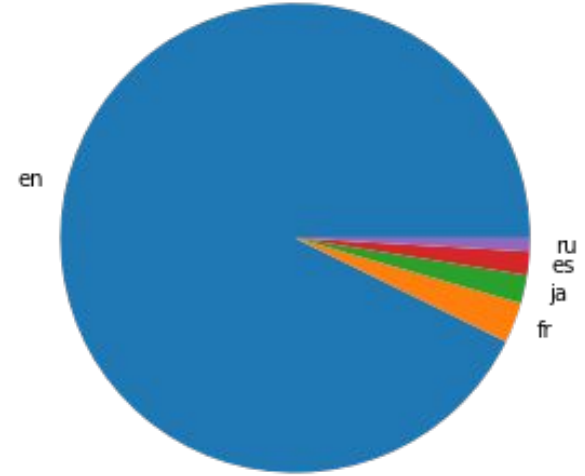


Results

Preferred Languages

- English is the most preferred language.

Most Preferred Languages



Key:

en - English

fr - French

ja - Japanese

es - Spanish

ru - Russian

Conclusions

From the findings, the following recommendations are deemed suitable for Microsoft's new venture:

#	Consideration	Recommendation
1	Budgetary allocation	Production budget approvals should be based on historic data with gross income considered. Additionally, nuanced decisions should be made based on the target market, i.e. domestic or international audiences.
2	Genre references	Investment priority should be drama and comedy genres.
3	Language preferences	Microsoft should prioritize investment in English language original content. However, the company should also make conservative investment in French and Japanese language content in an attempt to make inroads in those unsaturated markets.

Thank You!

Email: sarangemanono@gmail.com

GitHub: @SarangeManono

LinkedIn: [linkedin.com/in/sylviasmanono](https://www.linkedin.com/in/sylviasmanono)