



Instructor : Mark Tiba

Date: 20-November-2022

## **BUSINESS UNDERSTANDING**

Microsoft Company is planning to join the booming movie industry, since it is a lucrative business for most big tech companies

However, Microsoft has no knowledge about movie production or what to recommend to its costumers

My objective of this project was to give recommendations to Microsoft Company on how to enter the movie industry. I, therefore, will reach this goal by conducting proper data analysis from the data I have been provided

# Understanding The Problem

The movie industry is a multi-billion-dollar industry and is getting bigger every other day, especially with the young generation and with most people having access to the internet and devices

Investing in the movie industry sounds like a good business opportunity for a large company like Microsoft.

However, joining the industry without proper knowledge may lead to huge losses to the company despite it being a large company.

To come up with the best recommendations to give to Microsoft, I based my analysis on the following factors:

- Movie genre vs ratings and number of votes: Which is the best movie genre to invest in depending on ratings and votes?
- Movie budget vs movie income (gross): Does high budget translate to high income?
- Movie income vs genre: Does the movie genre determine its income?

## Technologies Used

- GitHub
- Jupyter Notebook
- Pandas Library
- Matplotlib
- Seaborn
- Sqlite3

## Data Understanding

The [tn.movie\\_budgets.csv](#): Provides data on production budget of a movie and the domestic and foreign gross of that movie.

The [im.db](#) is a relational database with tables containing information such as movie directors.

Using different technologies these files were read and prepared data accordingly for analysis.

Data was then cleaned through dropping rows with NaN values and duplicates to remain with the relevant data.

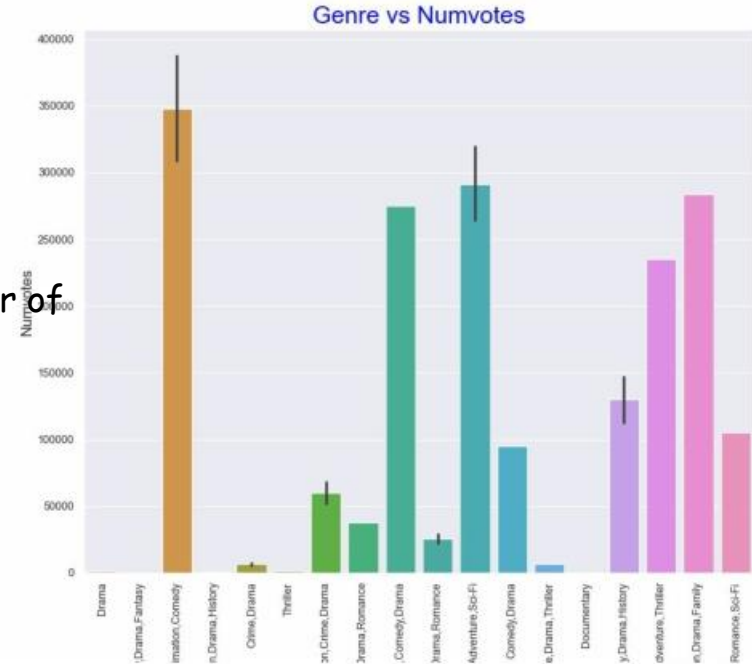
# Data Analysis

Which movie genres have higher ratings with high number of votes:

I plotted a bar graph of genres vs. numvotes, for movies with ratings above 6 and numvotes above 100.

From the graph the genres have high ratings and number of votes are :

1. Adventure, Animation, Comedy
2. Action, Adventure, Sci-Fi
3. Action, Drama, Family



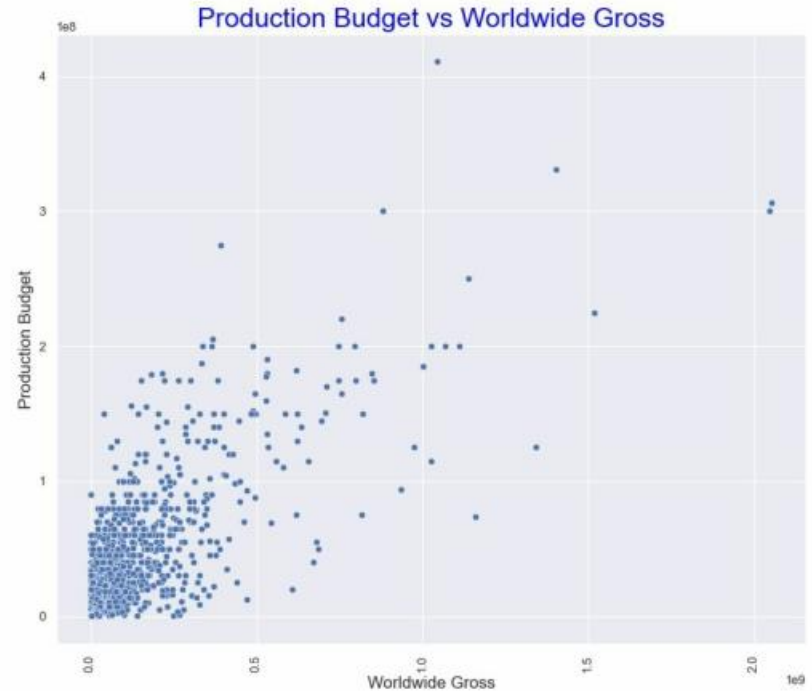
# Data Analysis

Does high production cost translates to high movie income?

The graph clearly shows that correlation exists between production budget and worldwide gross.

There are movies with high production budget that do not translate to high worldwide gross.

There are also movies with low production budget but high worldwide gross.

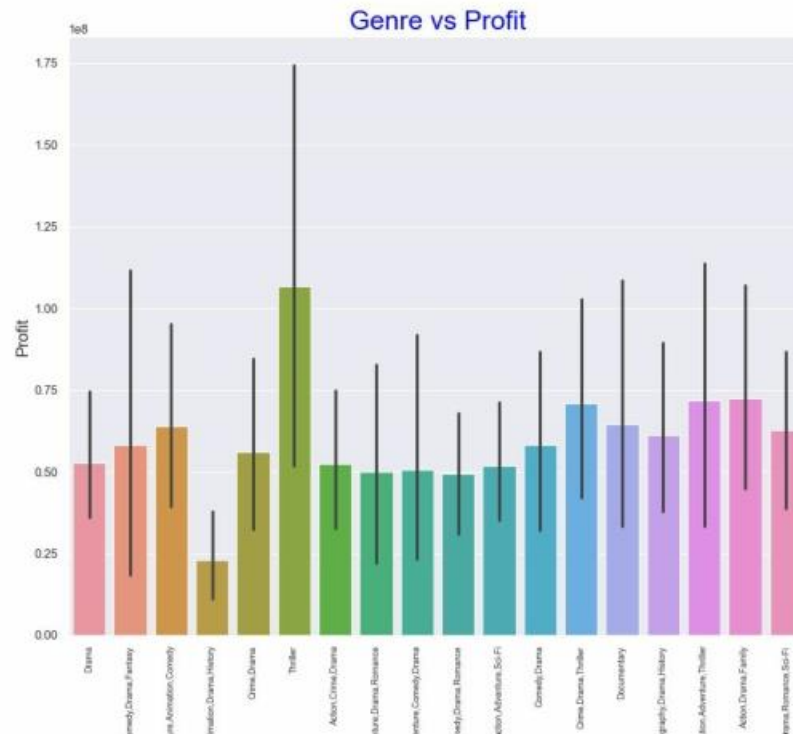


# Data Analysis

Does movie income depends with its genre?

The chart shows that the following genres make the highest profits:

1. Thriller
2. Action, Drama, Family
3. Action, Adventure, Thriller
4. Crime, Drama, Thriller





## Overall Analysis

- This data analysis proved wrong certain hypothesis about movie industry.
- For instance, it is assumed that high production budget translates to high movie income.
- From the analysis, it is proven that high production cost does not guarantee high income.
- Second, this analysis finds high movie rating does not translate into high income.
- For instance, Adventure, Animation, Comedy genre has a very high rating but low income compared to Thriller which has very low rating but high income. Action, Adventure, Sci-Fi is another example of genres with high ratings but average income.

This conclusion, however, may be affected by the number of movies in said genres.

## My Recommendations

- ★ First, using the graphs that I have presented, I would recommend that Microsoft's new movie studio kicks off with Thriller movies.
- ★ Thrillers have the potential of making more money for the studio and the company.
- ★ Thrillers generate the highest profits therefore making it the best to recommend
- ★ Second, the production budget should be kept at the possible minimum because high production cost does not translate into high income. However, this should not be done at the expense of the movie quality.