Microsoft Movie Studios

—

by Mario Mocombe

Summary

Microsoft is creating a new movie studio and is inquiring about the types of films currently doing the best at the box office. The datasets used in answering the business problem are from movie websites and methods used were exploratory data analysis, data cleaning, and data manipulation.

Summary

-   Adventure, Action, Comedy are the top-grossing film genres

-   May is the best month of the year to release a movie

-   Strong correlation between high budget and higher profits

-   Christopher Nolan should be hired as director

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

Microsoft has created a new movie studio but don't know anything about creating movies. We will discuss the types of films currently doing the best at the box office and hopefully Microsoft Studios can use these findings to make informed business decisions.
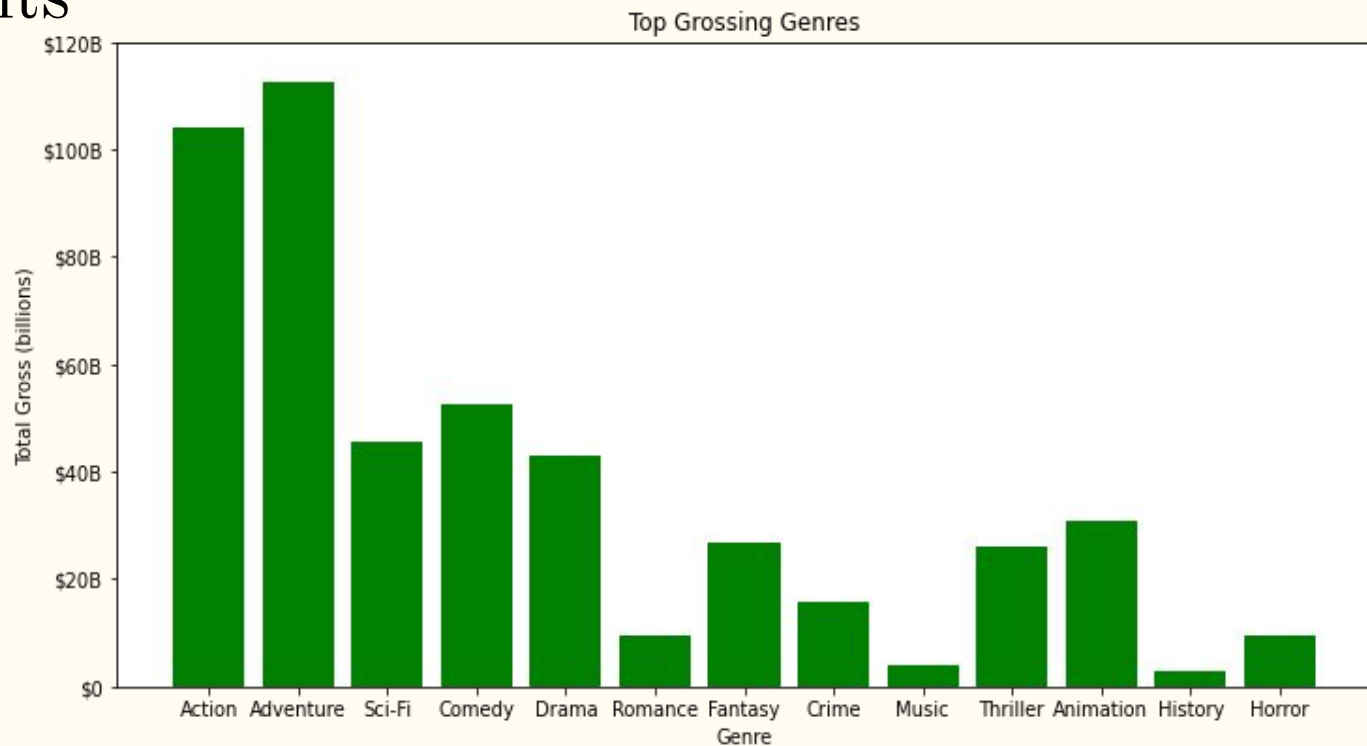
Data

IMDB Database: Contains information on genre, online user votes, average user rating, identity and roles of people involved in the films.

The Numbers Database: Contains box office data, release dates, and budget data.
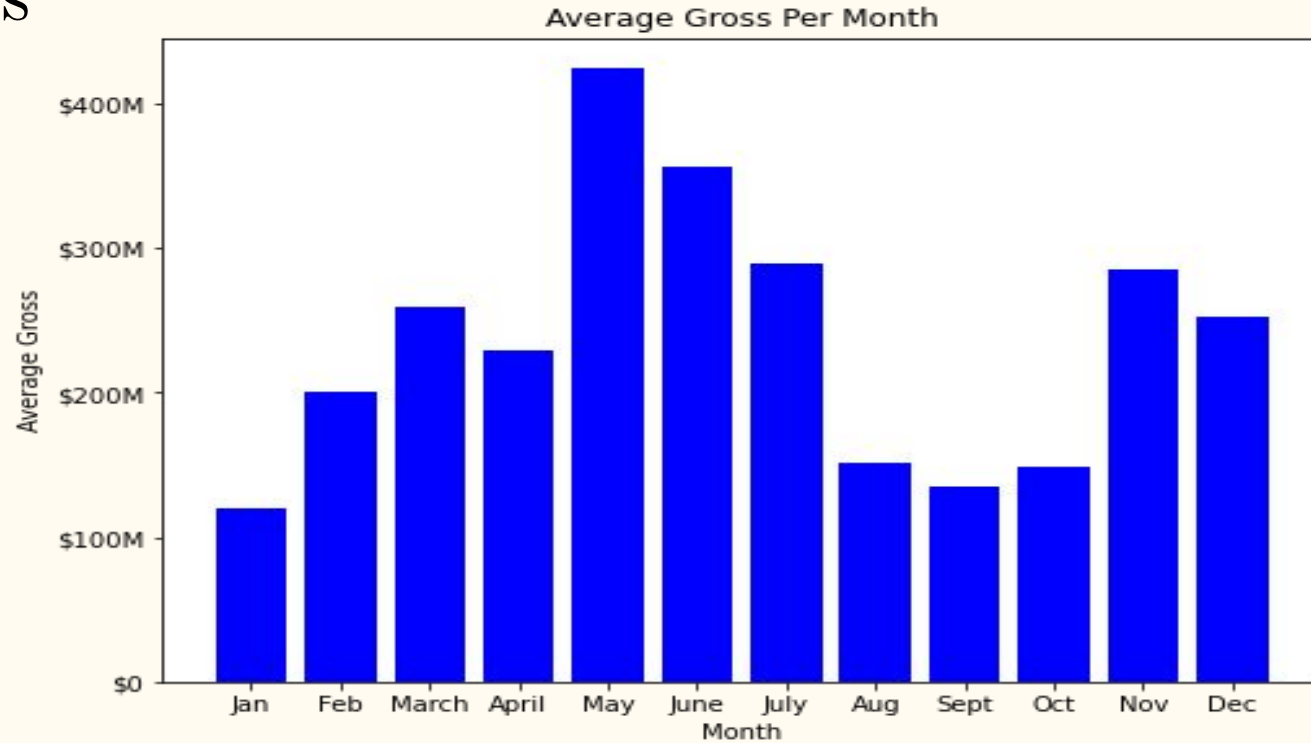
Methods

- Data cleaning: Identifying and handling missing values

- Data aggregation: Joining dataframes and using summary statistics

- Exploratory Data Analysis: Looking at the data in-depth, finding useful patterns and spotting anomalies while using visual aids to understand the data.
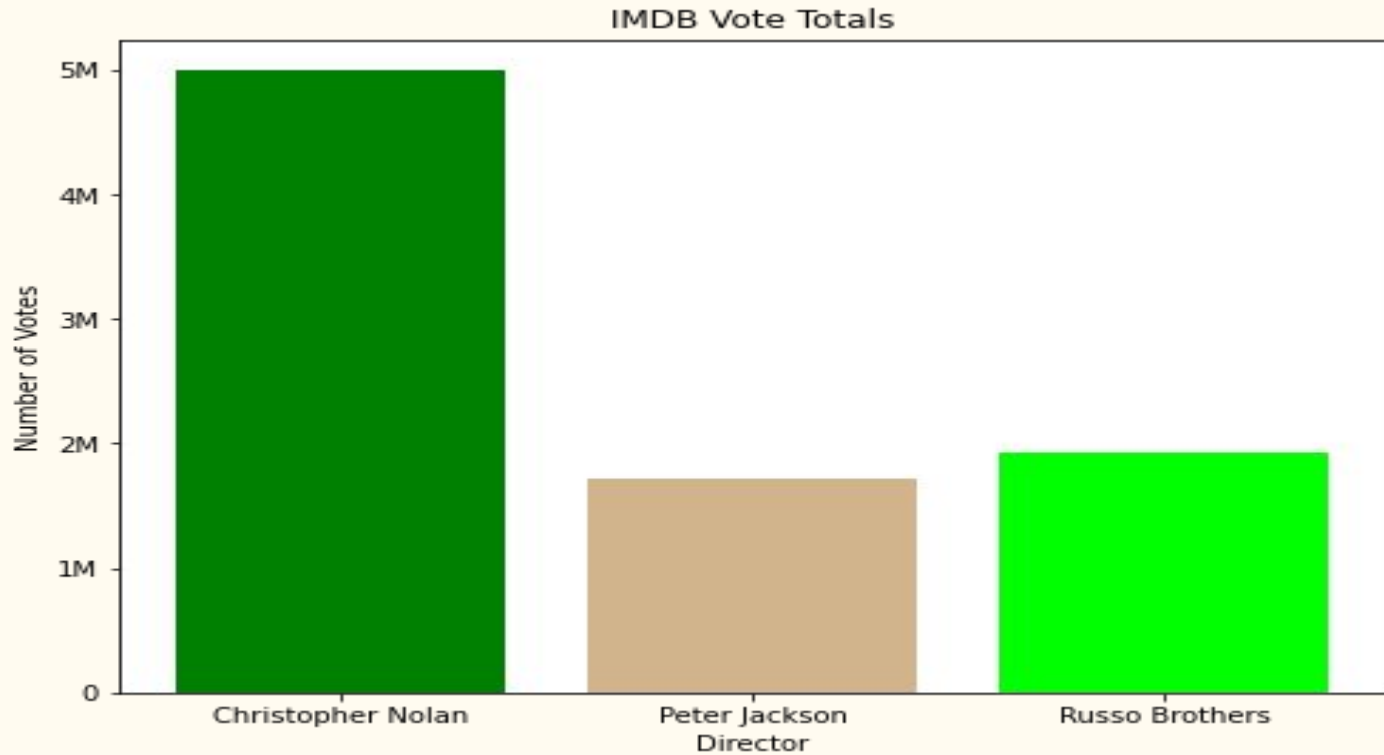
Results

Top Grossing Genres

Top grossing genres are Adventure, Action, and Comedy

Results

Average Gross Per Month

May is the highest grossing month, followed by June and July.

Results

IMDB Vote Totals

Christopher Nolan's films have nearly 5 Million votes on IMDB.

Conclusions

- Collaborate with Christopher Nolan

- Release a big budget action-adventure film

- Aim for a May release

Limitations

- Release window may be overcrowded
- Analysis does not predict far into the future
- Older films will have less online engagement on IMDB
- Additional data would be helpful:
    - movie streaming numbers
    - digital purchase and rental sales
    - VHS and DVD sales
    - data during the pandemic and the post-pandemic recovery period

Thank You!

Email: mario.mocombe@gmail.com
GitHub: @mmocombe
LinkedIn: linkedin.com/in/mario-mocombe-80117211/