

 [kjspring](#) / [dsc-phase-1-project-v2-4](#) Public

forked from [learn-co-curriculum/dsc-phase-1-project-v2-4](#)


Flatiron Data Science Crouse Phase 1 Project

☆ 0 stars  66 forks

☆ Star

👁 Watch ▼

<> Code

 Pull requests

🎬 Actions

📁 Projects

📖 Wiki

🛡 Security

📈 Insights

⚙ Settings

 master ▼

...

This branch is [51 commits ahead](#) of learn-co-curriculum:master.

 Contribute ▼

🔄 Fetch upstream ▼



kjspring updated results of README.md ...

14 seconds ago

🕒 69

[View code](#)

☰ README.md



Microsoft Movie Studio Production Recommendations

Overview

In 2021, digital, theatrical, and physical sales of movies were \$136.5 billion globally and \$36.8 billion in the United States. Digital sales account for 80% of revenue, theatrical sales are 12%, and physical media is the remaining 8%.

Microsoft is creating a new movie studio and wants to know what types of movies they should create to tap into this industry. This project will explore what types of movies are currently doing the best at the box office and recommend what types of movies Microsoft should create.

Objective - Business Questions

- What movie genres should Microsoft Movie Studios focus on?
- When should these movies be released?
- How much should be budgeted for the production of each movie?

Deliverables

- [Presentation: Genre, Release Time, and Budget for Microsoft Movie Studio](#)
- Jupyter Notebook
- GitHub Repository

Repository Navigation

- [/img](#) - contains image files
- [/img/fig](#) - stores plots and figures created with the analysis
- [/zippedData](#) - stores data used for this analysis
- [/license](#) - contains the license information
- [project_notebook.ipynb](#) - Jupyter Notebook

Technologies

- [Python](#)
- [pandas](#) library
- [matplotlib](#) library
- [sqlite3](#) library

Methods

1. Clone this repo (for help see this [tutorial](#)).
2. Gather and import [raw data](#)
3. Combine, join, and clean data
4. Exploratory Data Analysis to form recommendations
5. Plot visualizations

The Data

Data Description

In the folder `zippedData` are movie datasets from:

- `bom.movie_gross.csv.gz` : [Box Office Mojo](#)
 - stored in compressed CSV
 - contains domestic and foreign gross revenue
- `tn.movie_budget.csv.gz` : [The Numbers](#)
 - stored in compressed CSV
 - contains production budget

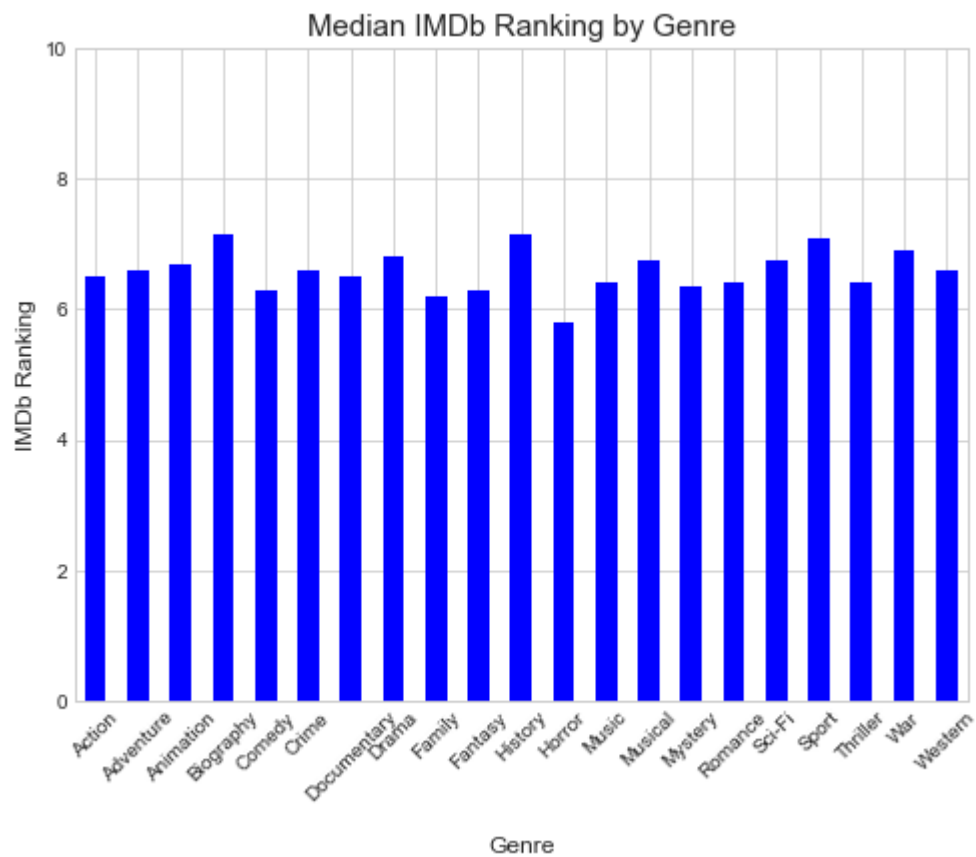
Box Office Mojo and The Numbers data are stored in compressed CSV (comma-separated values).

- `im.db.zip` : [IMDB](#)
 - stored in a SQLite database
 - contains user-generated rankings on a scale of 0 - 10 with 0 being the worse and 10 the best

Data Analysis and Recommendations

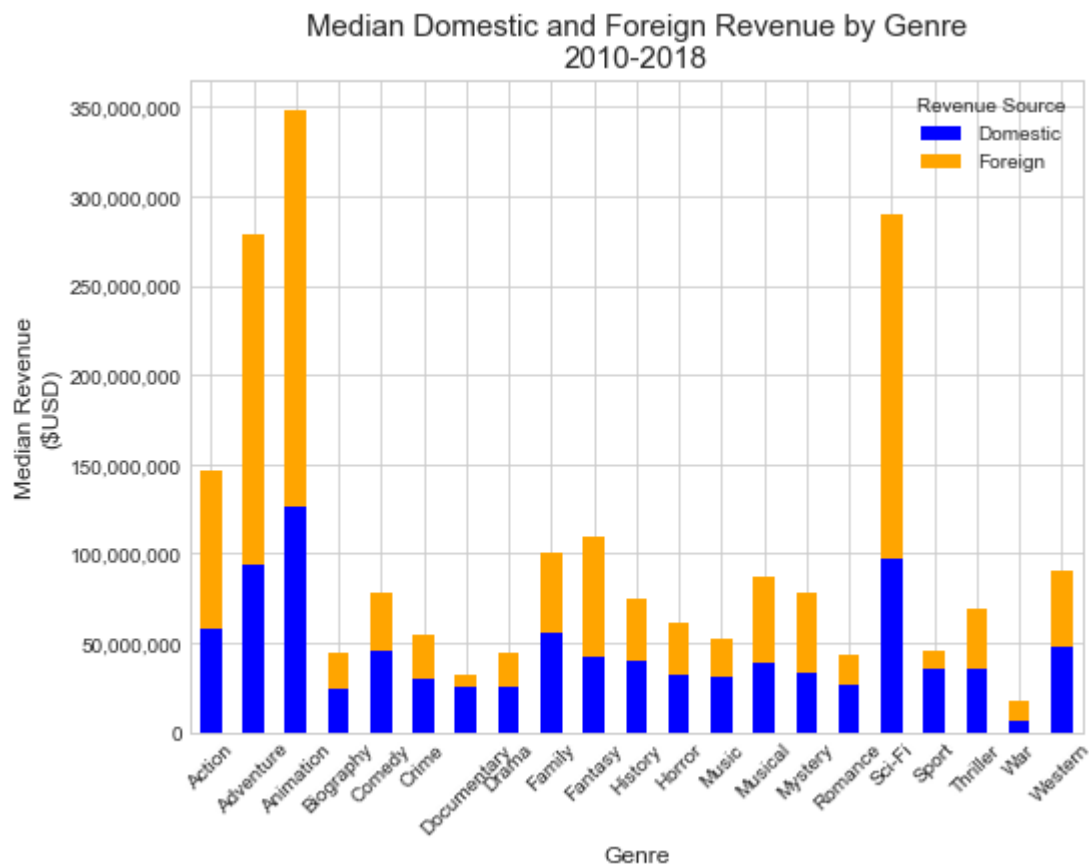
Movie genre does not have an effect on IMDb

- The median IMDb ranking for all analyzed movies was 6.5 and the rankings between the different genres is very different.



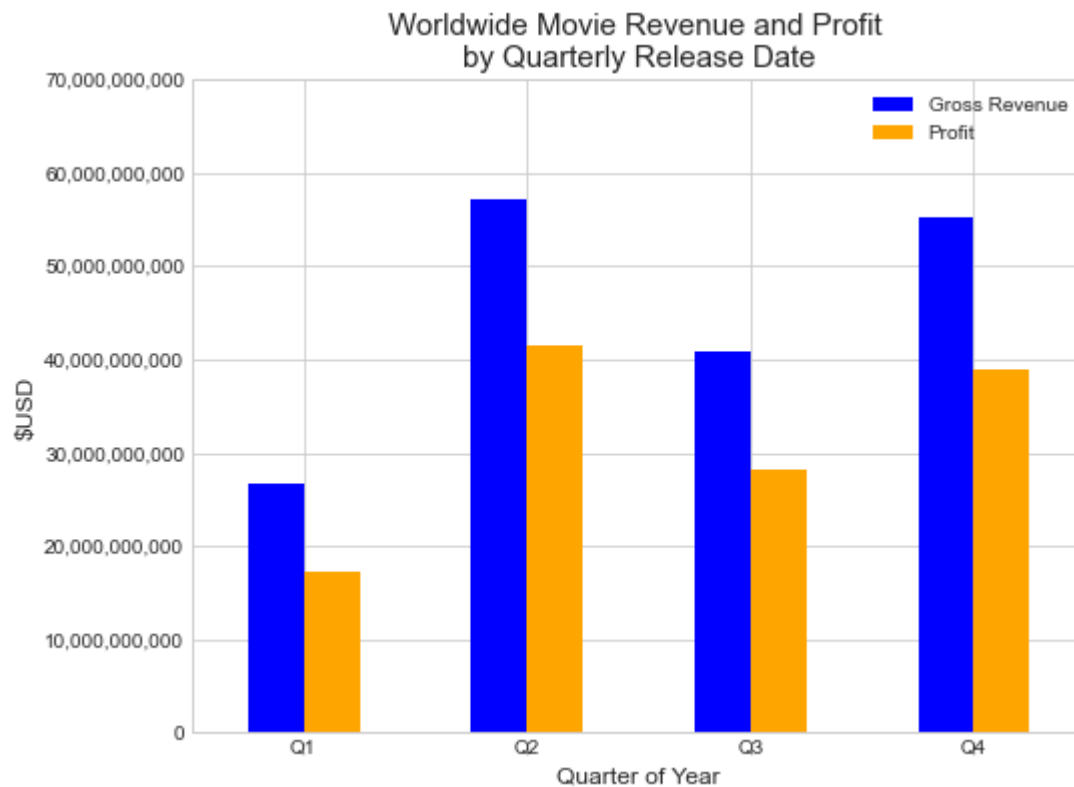
Movie genre should be a combination of Action, Adventure, Animation, or Sci-Fi

- These genres generated the most worldwide revenue.



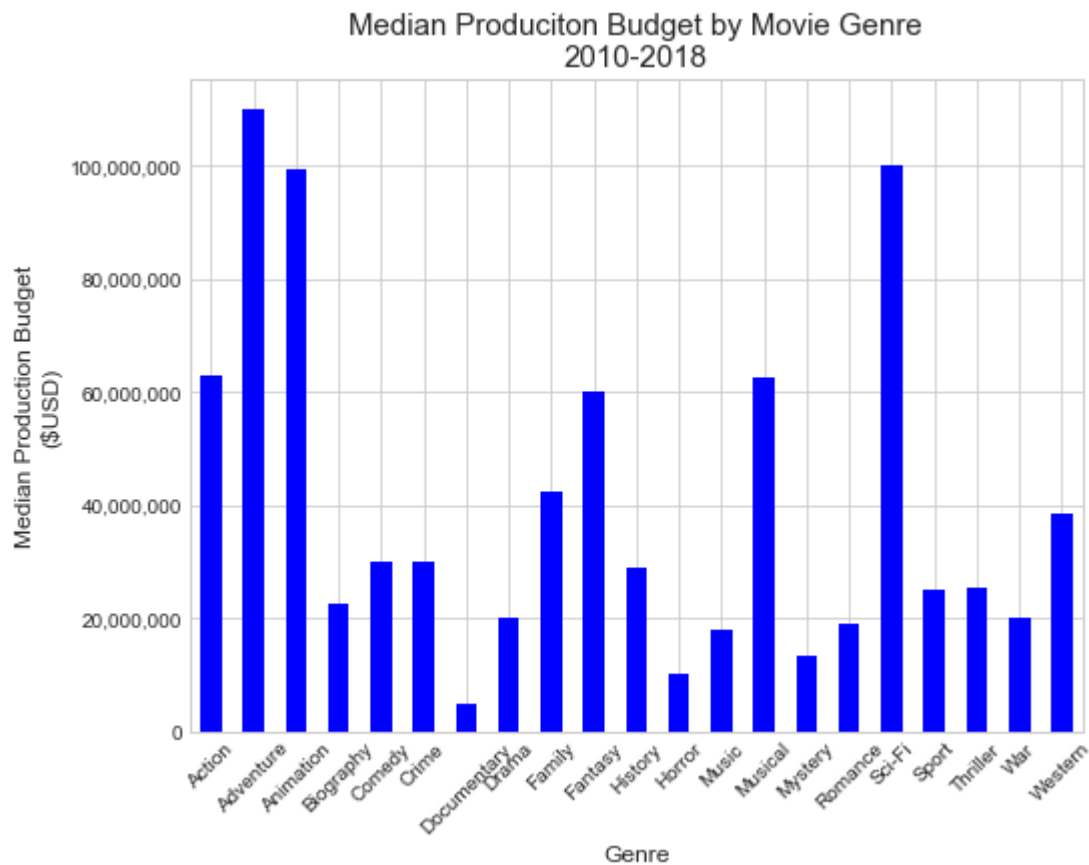
The movie should be released in Q2 or Q4

- Q2 and Q4 have the highest revenue because these are when people have more time as they are on holiday, working less, and not in school.



Expect to budget \$60 million to \$100 million per movie

- The highest revenue generators also have the highest production budgets.



Conclusions

Microsoft wants a part of the \$136 billion global movie industry by creating highly ranked and profitable movies. Microsoft is positioned well with its brand recognition and streaming services to capture most of the revenue generated after a movie's theatrical release.

- Microsoft should focus on action, adventure, animation, or science fiction genres
- The movies should be released in the summer during Q2 or during the fall or winter holidays in Q4
- The budget needed for these genres is between \$60 million to \$100 million.

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages

● Jupyter Notebook 100.0%