

PHASE 1 PROJECT

DATA ANALYSIS AND VISUALIZATION

PROJECT OVERVIEW

This project aims to analyze box office data to provide actionable insights for Microsoft's new movie studio on what types of films to create and how to maximize revenue.

BUSINESS PROBLEM

Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they do not know anything about creating movies. This project seeks to explore the types of films that are currently doing the best at the box office and translate these findings into actionable insights that the head of Microsoft's new movie studio will use to help decide what type of films to create and how to maximize revenue.

THE DATA

The findings of this project were derived from the below datasets which were loaded as dataframes for effective analysis.

- 'bom.movie_gross.csv.gz'
- 'tn.movie_budgets.csv.gz'
- 'tmdb.movies.csv.gz'
- 'rt.reviews.tsv.gz'
- 'rt.movie_info.tsv.gz'
- 'im.db'

PROJECT OBJECTIVES

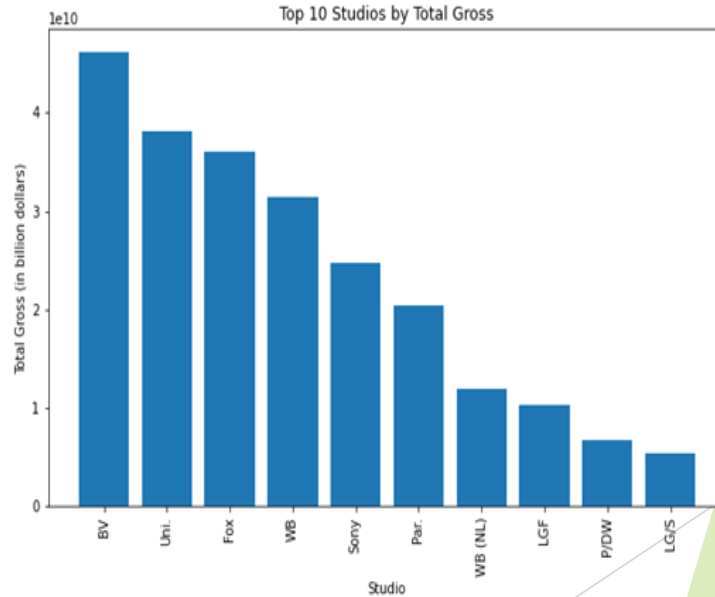
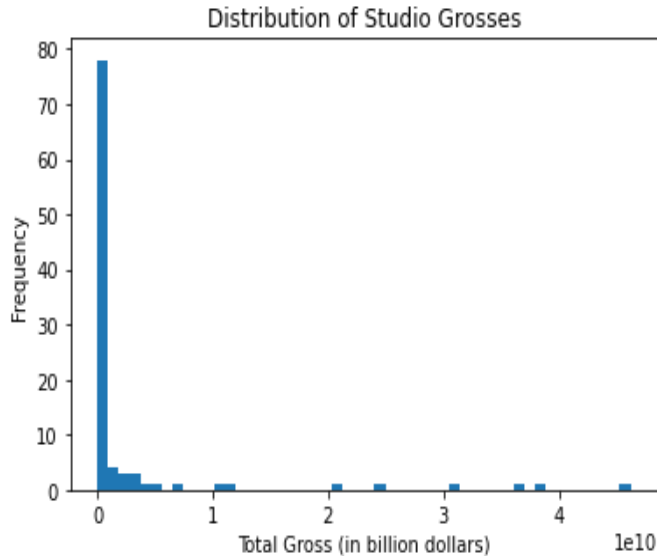
- ✓ Identifying the most successful studios based on domestic gross revenue.
- ✓ Identifying which genres tend to have higher gross revenues.
- ✓ Identifying the correlation between film production cost and its box office success.
- ✓ Investigating the relationship between ratings & reviews and box office success.

DATA ANALYSIS AND VISUALIZATION

- ▶ To analyze the datasets for this project, I used descriptive statistics, which involved summarizing and presenting data in a meaningful way to gain insights and understanding of the data.
- ▶ I then used various data visualization techniques to explore the data and identify patterns and relationships between variables.
- ▶ Let's look at my observations from my analysis:

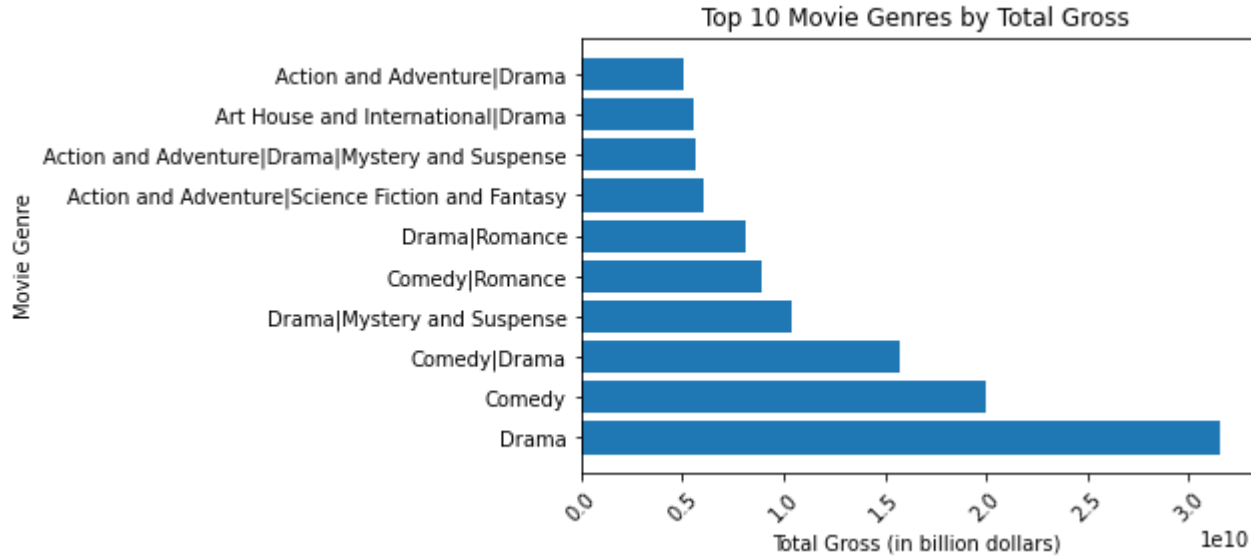
Visualization and observation of studio to gross analysis:

The movie industry is highly competitive with most studios earning modest gross revenue below 1 billion dollars and a few studios dominating the industry earning over 4 billion dollars.



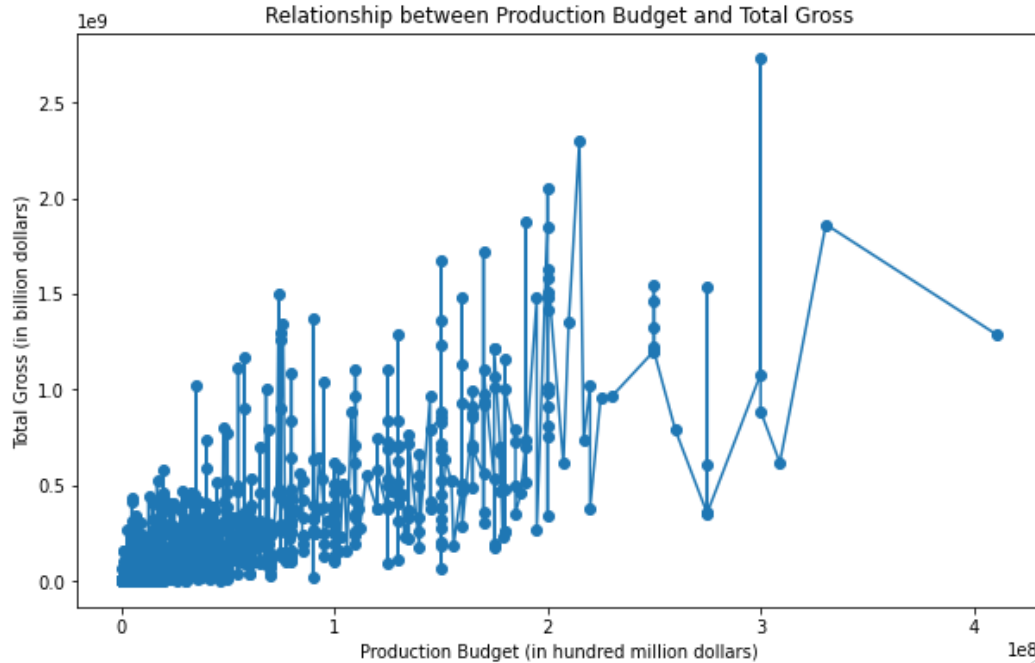
Visualization and observation of genre to gross analysis:

The highest grossing movies are drama and comedy genres grossing at over 1 billion dollars



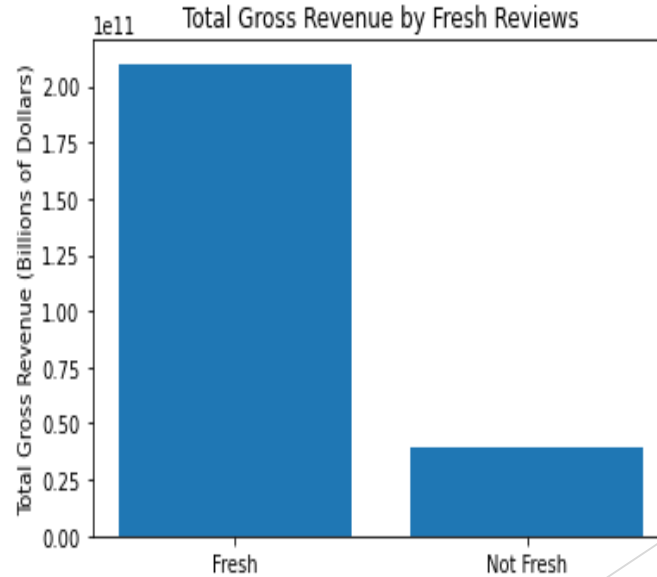
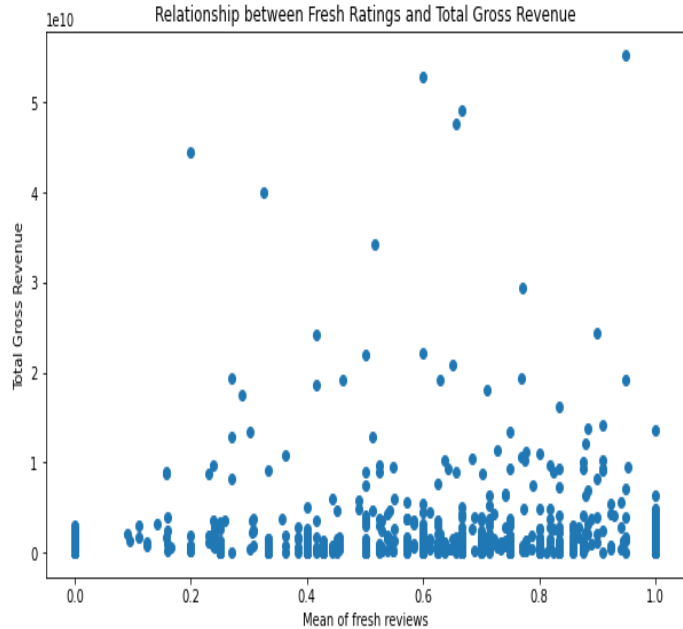
Visualization and observation of production cost and revenue correlation:

Movies with a higher production budget tend to have higher gross revenues with only a few exemptions.



Visualization and observation of correlation between ratings and gross revenue:

Although positive reviews and high ratings do not always reflect to box office success, movies with rave reviews and ratings tend to do better in the box office.



RECOMMENDATIONS

My analysis led to four recommendations useful for the inception of Microsoft's new movie studio:

- ✓ Seek to emulate the top 10 highest grossing studios through research and/or benchmarking or seek to partner with them.
- ✓ Put focus on producing drama and comedy genres which are the most successful genres in terms of gross revenue.
- ✓ Invest in the film's production cost because a larger budget will lead to better production which will lead to great reception and ultimately more box office success.
- ✓ Produce high-quality movies that receive rave reviews from critics and high ratings from the audience to increase the likelihood of the films' financial success.

NEXT STEPS

To stay up-to-date with the latest trends and insights, the further analysis below could be useful:

- ✓ Analyze trends over time to provide valuable insights into which genres are growing or declining in popularity.
- ✓ Look into the impact of streaming services in the movie industry.
- ✓ Investigate the impact of marketing on box office success.
- ✓ Analyze the success rate of sequels and franchises, originals and remakes.

THANK YOU!

For questions, concerns or feedback, you can reach me on LinkedIn.

NAME: JOYCE MUTHONI MUTHIKE

LinkedIn profile: [linkedin.com/in/muthoni-muthike-85b5831a6](https://www.linkedin.com/in/muthoni-muthike-85b5831a6)