

MICROSOFT MOVIE ANALYSIS

AUTHOR: FRIDAH KIMATHI

BUSINESS BACKGROUND

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 don't know anything about creating movies.



PROBLEM STATEMENT

Microsoft has very little knowledge about the creation of movies and are unsure of what movies they should produce in there new movie studio.

DATA

Analysis was made using 3 datasets, two from *_imbd_* and one from *_bom_*

- BOM movie_gross dataset has important columns such as studio, domestic_gross and foreign_gross.
- IMBD title basics dataset has substantial columns such as start_year, runtime_minutes and genres.
- IMBD ratings dataset has important columns such as average rating and number of votes

The target variables were genres, start_year(movie release year), the total_gross (combination of domestic and foreign gross)



METHODS

- Data cleaning
 - ✓ Identifying duplicates
 - ✓ Identifying missing values
 - ✓ Changing data types
 - ✓ Flattened the genre column
 - ✓ Checking for placeholders or outliers
 - ✓ Merging of data sets

METHODS

- Data analysis using functions and methods such as
 - ✓ Group by and aggregate functions
 - ✓ .info(), .head(), .dtypes
 - ✓ .duplicated
 - ✓ Isnull()

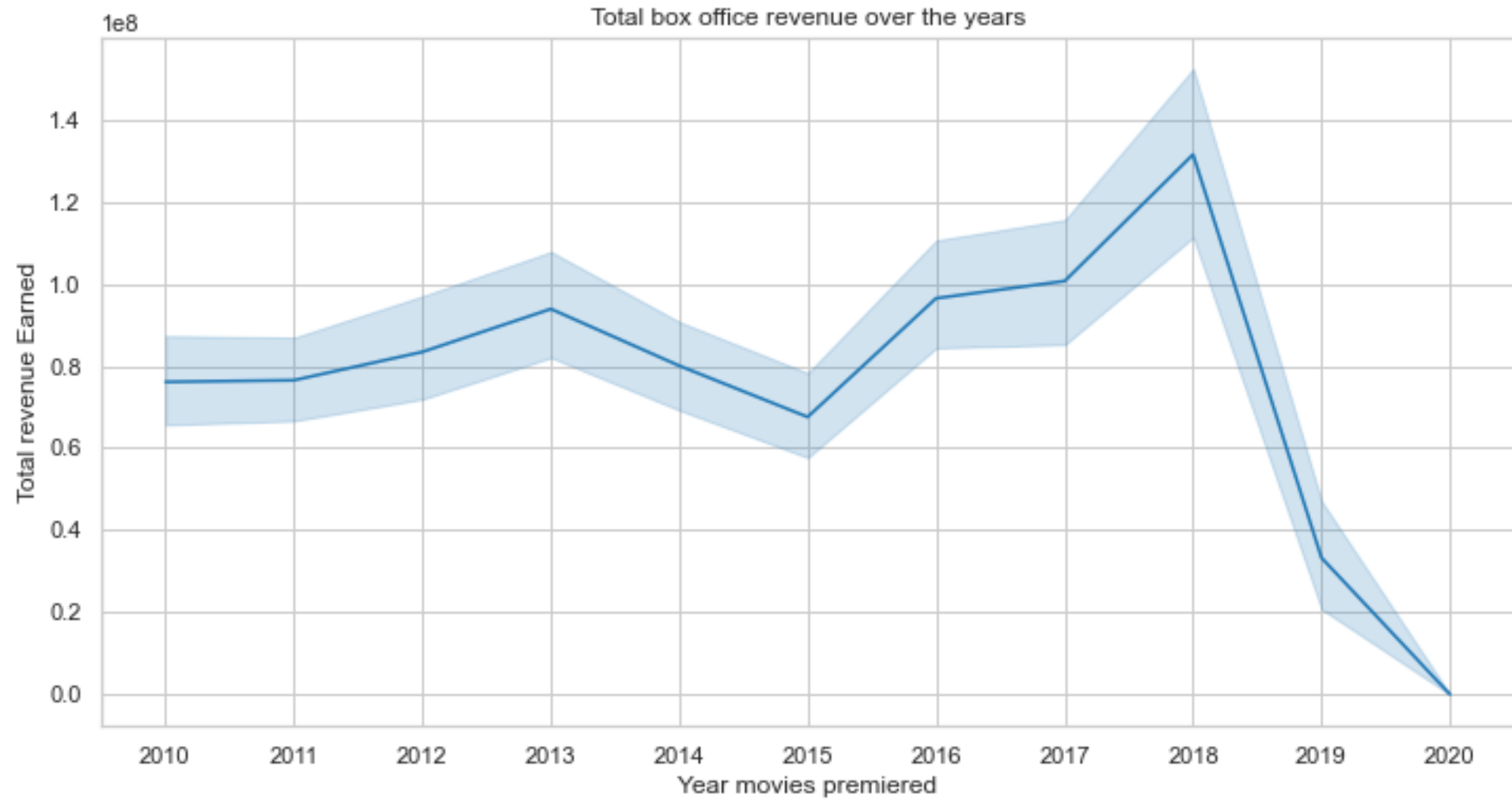
METHODS

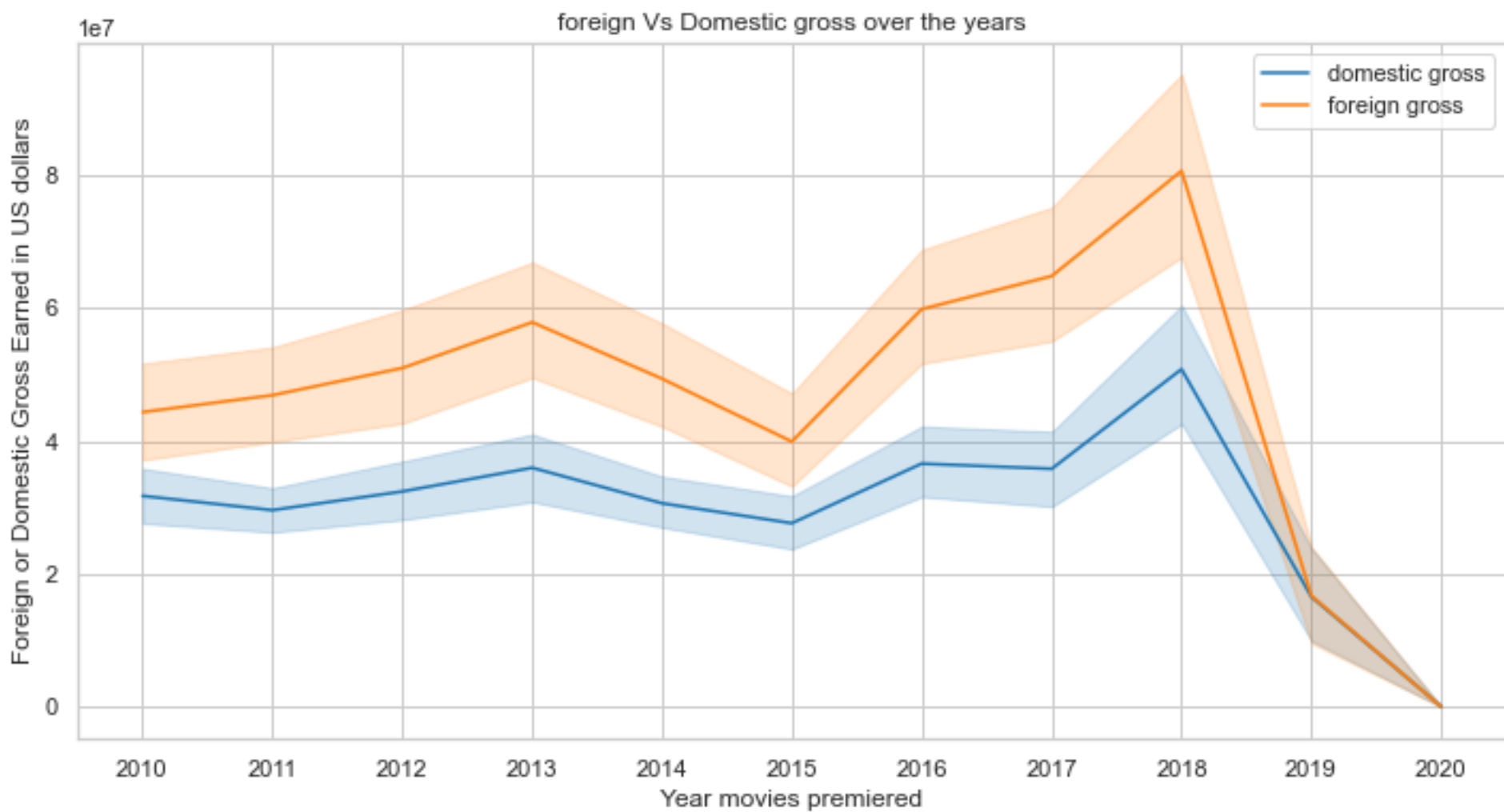
- Data visualization
 - ✓ Visualization of the total box office revenue over the years was done using a line plot
 - ✓ A comparison between the top_10 most preferred movie genres Vs the top_10 most profitable movie genres was done by plotting a bar plot in the same figure with different axes.
 - ✓ The top 5 highest rated genres were visualized using a bar plot.
 - ✓ The correlation between Average rating and total gross was done visualized using a scatter plot.
 - ✓ A visualization of the top 5 studios based on total revenue earned was done using a bar plot.

RESULTS



Total box office revenue over the years



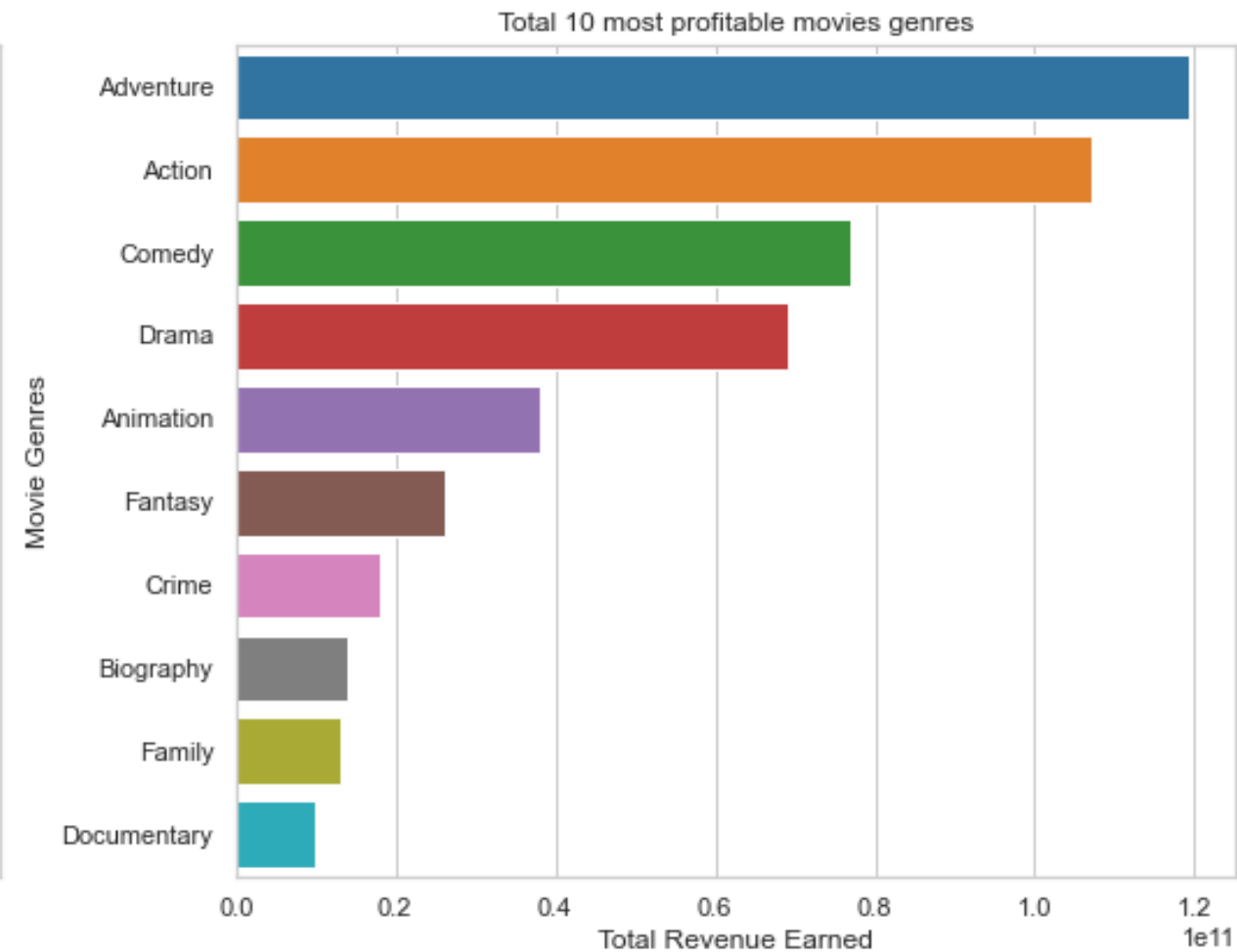
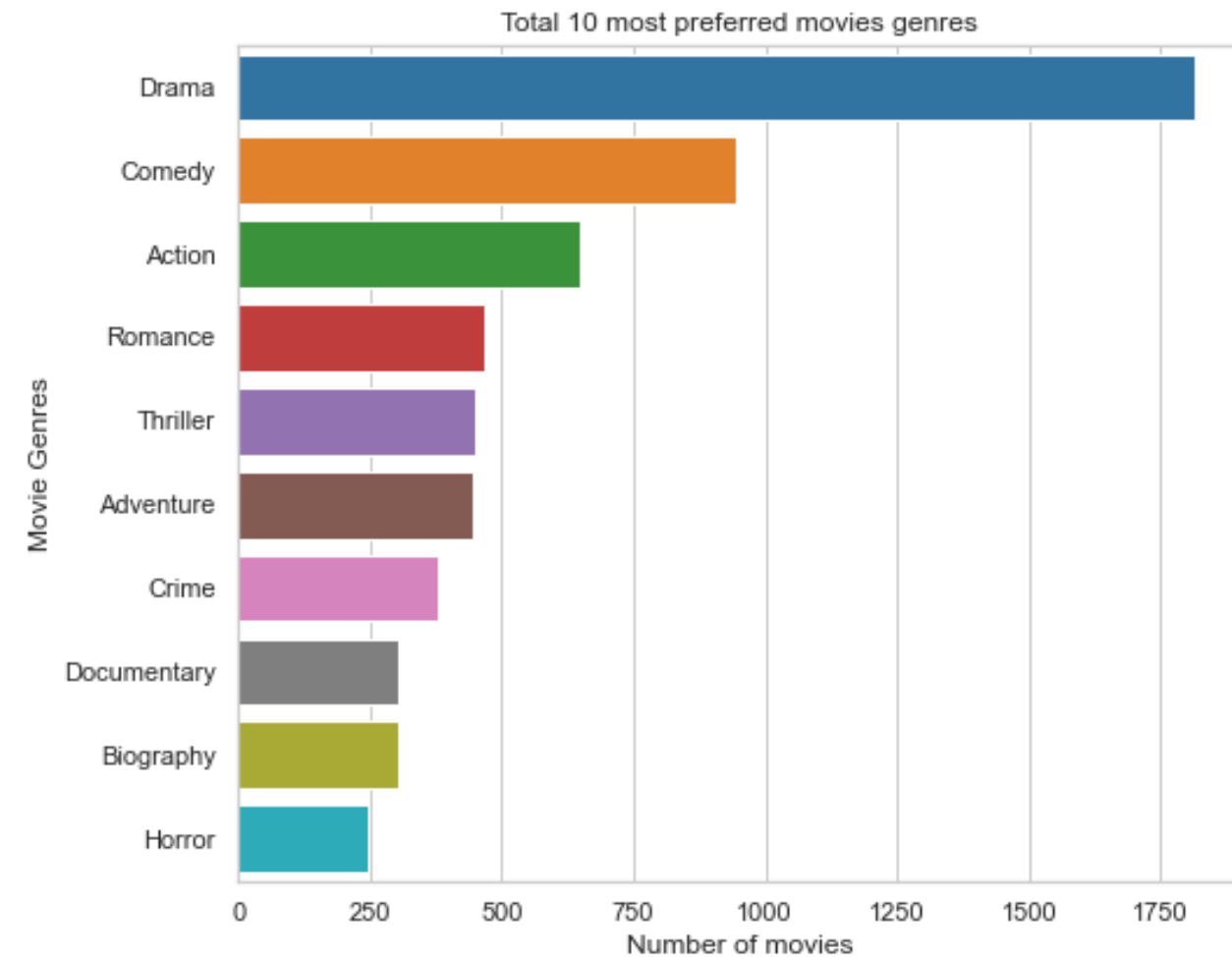


RESULTS

Observations for Total box office revenue over the years:

- movies made a lot of money in 2018.
- There has been a sharp decrease in total gross earned by movies since 2018, with the lowest year being 2020.
- This makes sense since streaming platforms started gaining a lot of hype in 2018, leading to lot of people preferring to stream movies on platforms such as Netflix hence the decline in box office revenues.
- Movies made zero box office revenue in 2020.
- This is because COVID started in 2020, which lead to the closure of movie theaters in a lot of countries hence there was no box office revenue for movies that year.

Preferred movie genres _Vs_ **The most profitable movie genres**

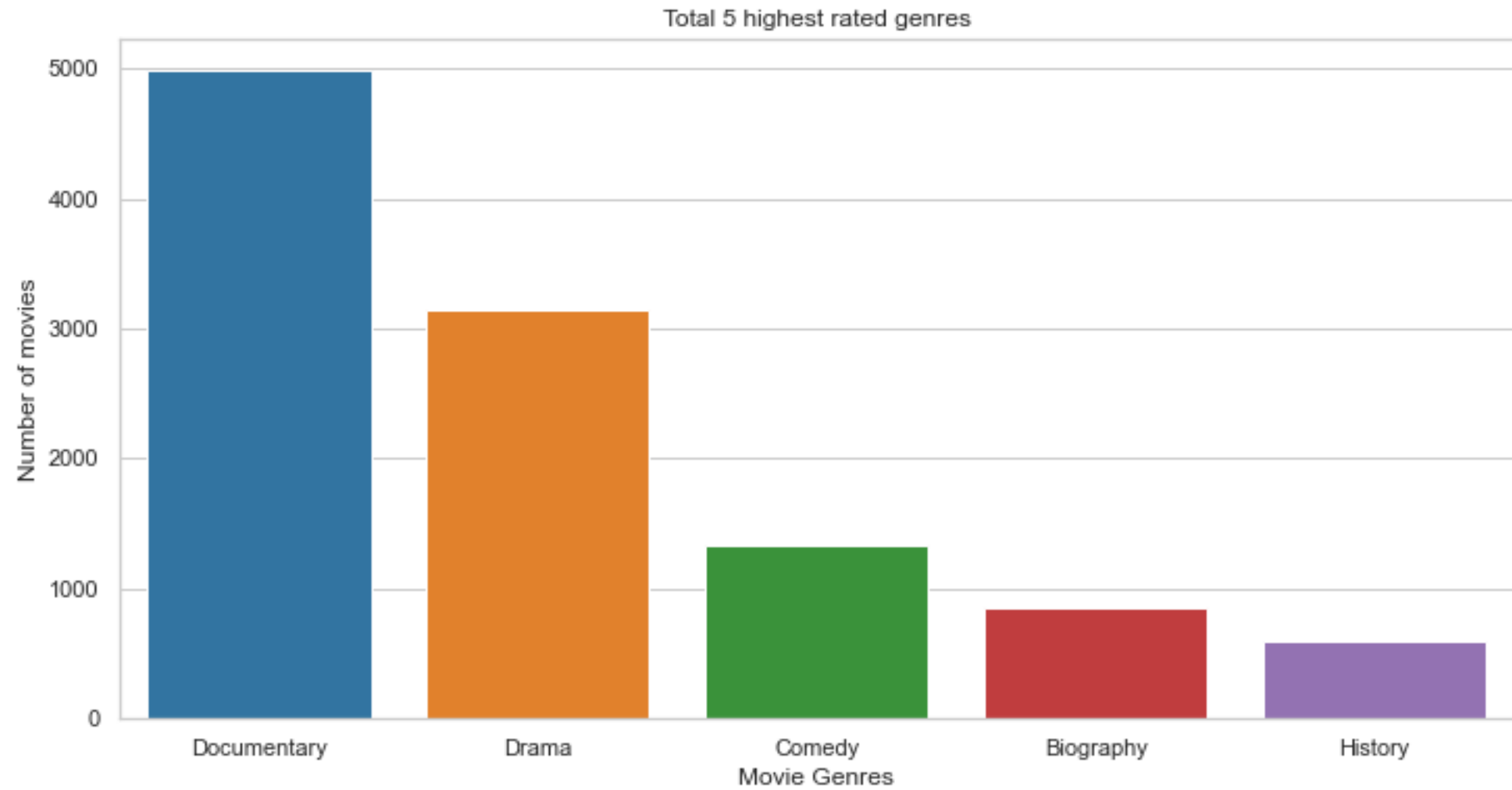


RESULTS

Observations for the preferred movie genres _Vs_ The most profitable movie genres

- Drama is the most preferred movie genre.
- Romance, Thriller, Horror which are in the top 10 most preferred movies is not in the top 10 most profitable movies
- Adventure is the most profitable movie genre.
- Animation, Fantasy and Family are in the top 10 most profitable movies but are not in the top 10 most preferred genres

Top 5 highest rated genres

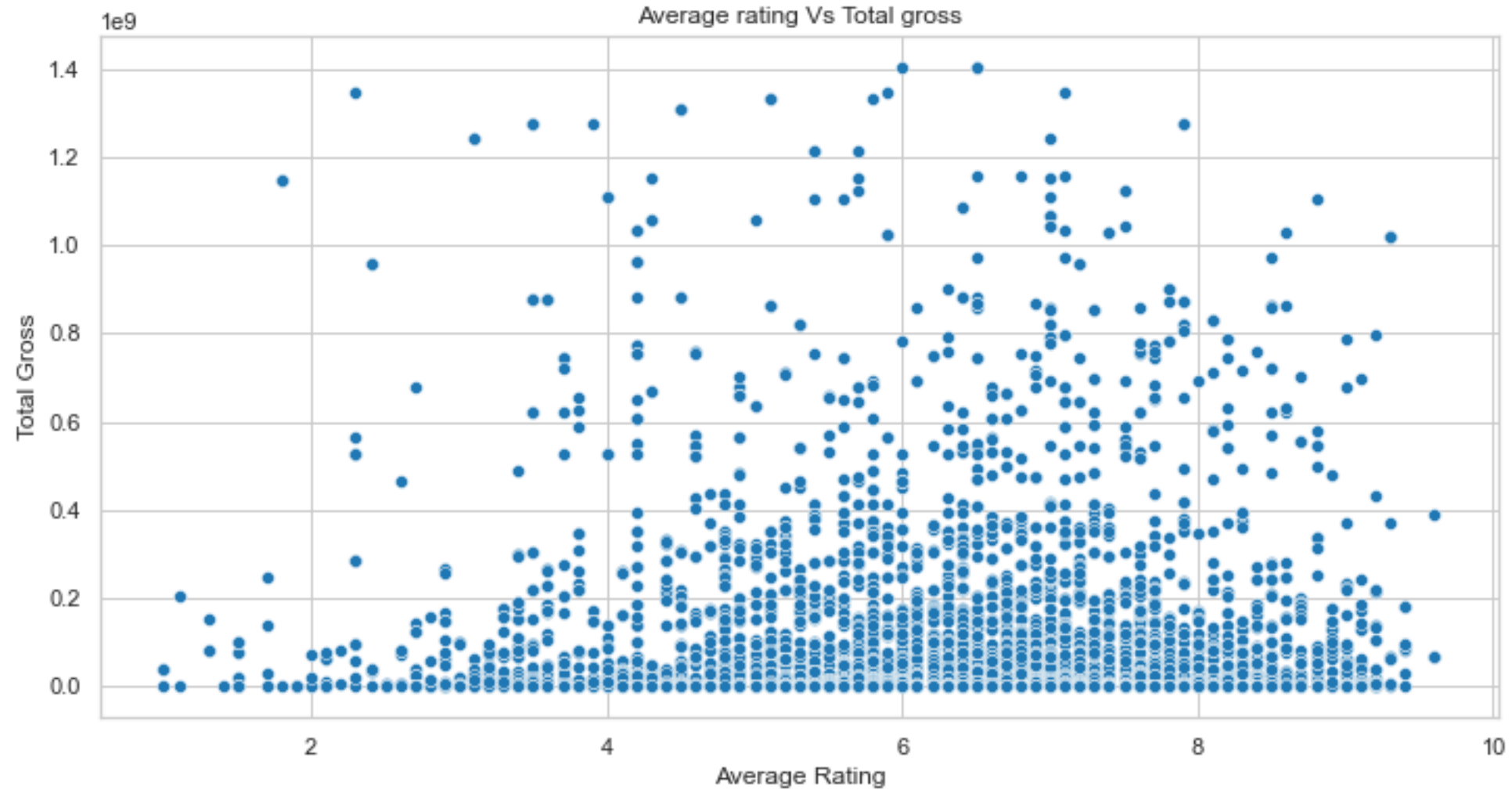


RESULTS

Observations for the top 5 highest rated genres

- Documentary is the highest rated movie genre.
- The first four genres, that is Documentary, drama , comedy and biography are also in the top ten most preferred/profitable movies
- History is not in the top ten most preferred/profitable movies

Correlation between Average rating and total gross

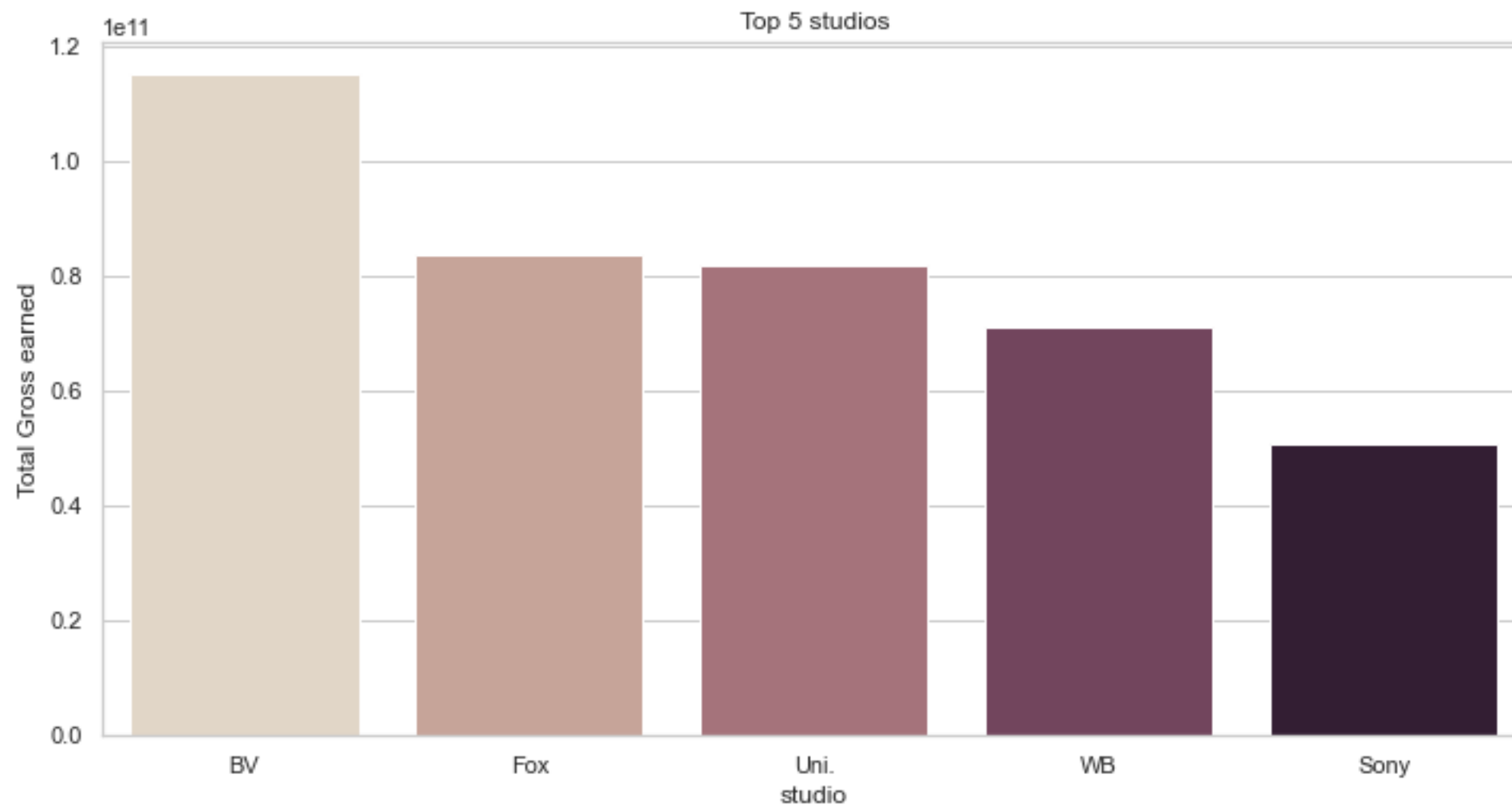


RESULTS

Observations for the correlation between Average rating and total gross

- There is no correlation between average rating and the total revenue earned.
- Majority of movies make a total gross of less than 400 million dollars.
- Majority of movies have an average rating higher than 4.

Top 5 studios based on total revenue earned



RESULTS

Observations for the Top 5 studios based on total revenue earned

- BV studio is the highest earning studio
- Fox and Uni.Studio have approximately the same gross earnings
- All 5 studios have a gross total greater than 400 million dollars

CONCLUSIONS

- There has been a sharp decrease in box office gross earnings since 2018. This can be attributed to the increase preference for movie streaming in platforms such as Netflix.
- Adventure, action, comedy and drama are genres that are highly preferred by audiences and also have high profitability.
- BV Studio has the highest total gross earnings.

RECOMMENDATIONS

From the results in my analysis, I recommended the following:

- Create a streaming platform for their movies
- Produce movies that have a combination of adventure, action, comedy and drama
- Benchmark with top earning movie studios such as BV studios

FUTURE PLANS

- Determine profit based on the movie's budget
- Compare streaming services gross earnings and box office gross earnings per movie
- Assess the relationship between genre and number of votes or what influences the number of votes per movie

Thank You!

Email: fridahnkiretekimathi@gmail.com

GitHub: @FridahKimathi

LinkedIn: [Fridah Kimathi](#)