NAME: VICTOR KABUI KARANJA

PROJECT TITLE: **ANALYSIS OF MOVIE DATABASE TO IDENTIFY A POSSIBLE ENTRY GENRE**

PRESENTATION OF FINDINGS

PHASE 1 PROJECT

DATA SCIENCE PART TIME (dsfpt01)

BUSINESS BACKGROUND:

• Microsoft is a company dealing with operating systems among many other deliverables. It intends to venture into the film industry.

• Having no prior knowledge on the film industry, consultation and adequate research had to be conducted to ensure they did a good investment.

PROBLEM STATEMENT:

• Lack of Knowledge:
Microsoft wills to venture into the movie creation sector however, having not been in the trade
before they risk failing.

• Resource Wastage:
Microsoft is not willing to go into losses or drain their resources
into a money pit

• Need to Keep with the Trend:
The other big companies 'competitors' having moved into the film industry, Microsoft has to ensure
it is not left behind in the old ways of running, else they will be out of the market with no point of
return

SOLUTION

In order, to mitigate through the problem statements, I was hired to perform Exploratory data analysis from the already existing movie datasets with records of 8 years and give insights on the way forward to be taken.
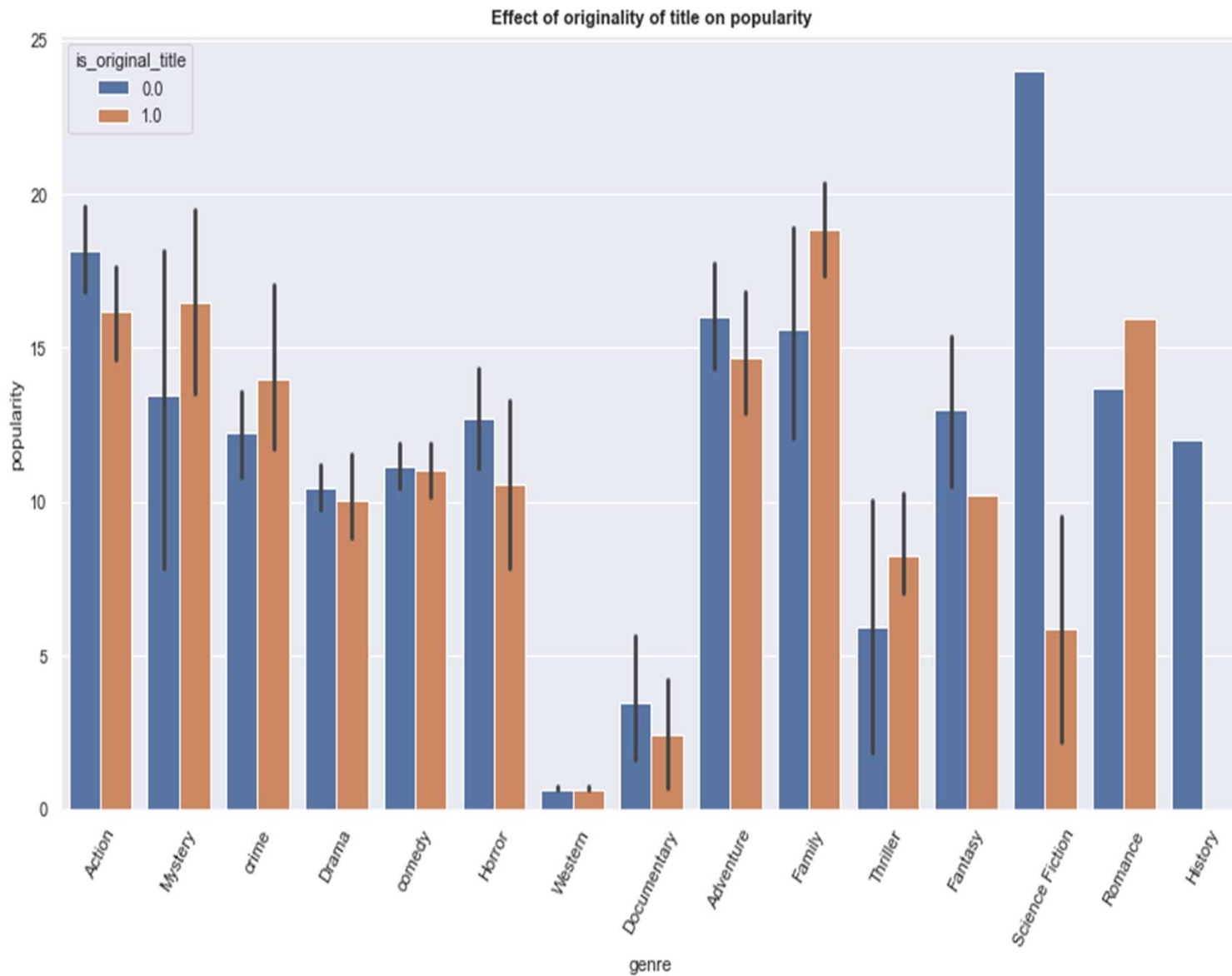
DEVELOPMENT PROCESS

1. Research and data analysis
2. Discussion of the findings
3. Installation of the Movie studio
4. Releasing of sample movies to a targeted group for feedback
5. Discuss any feedbacks obtained and make necessary corrections
6. Release to the larger population
7. Continuously receive feedback and improve
8. Research on what to improve so as to stay top of the market
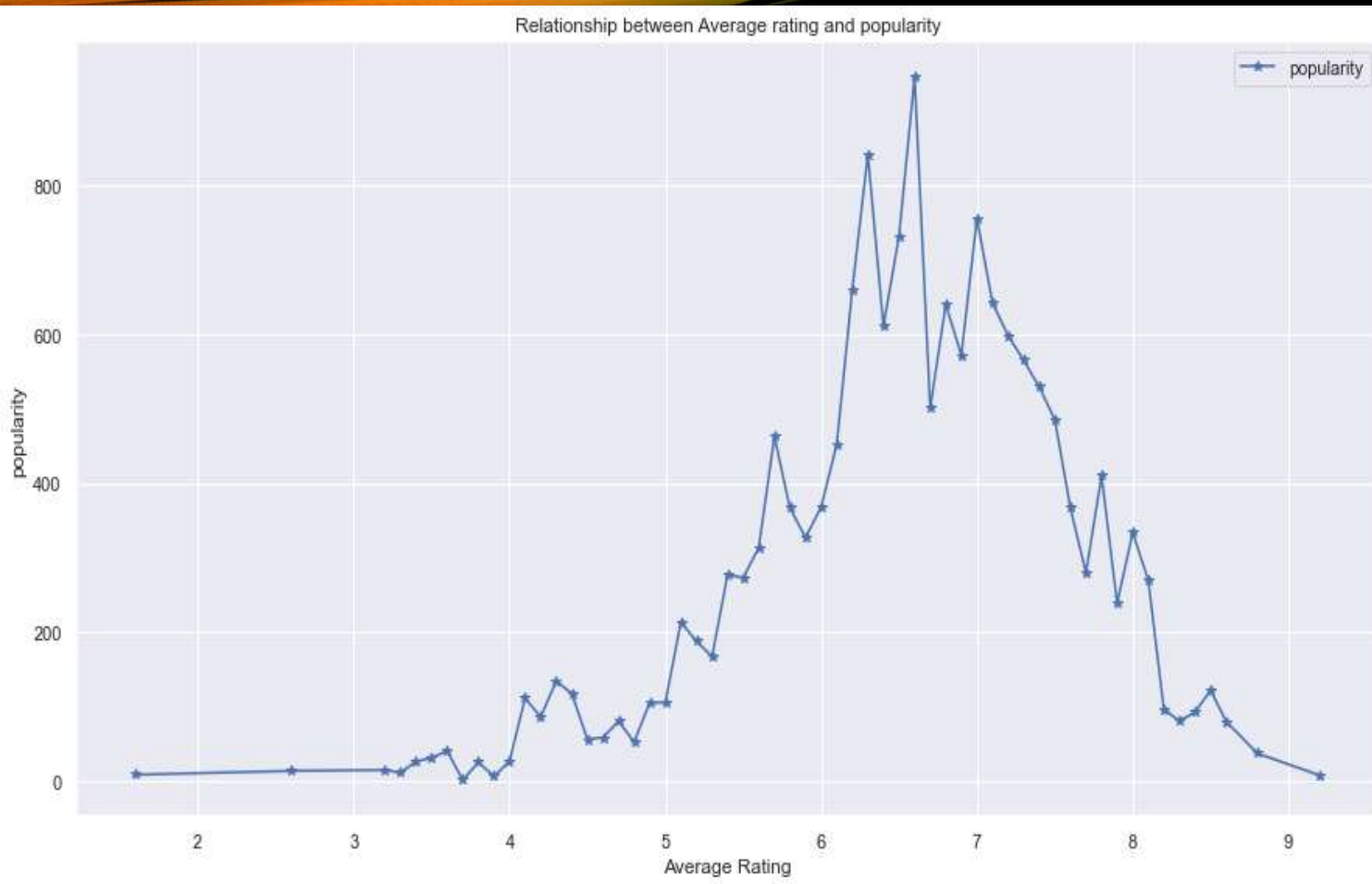
Technologies used:

The software used to perform the analysis included:
- Python programming libraries
- Matplotlip.pyplot
- Pandas
- Numpy
- Seaborn
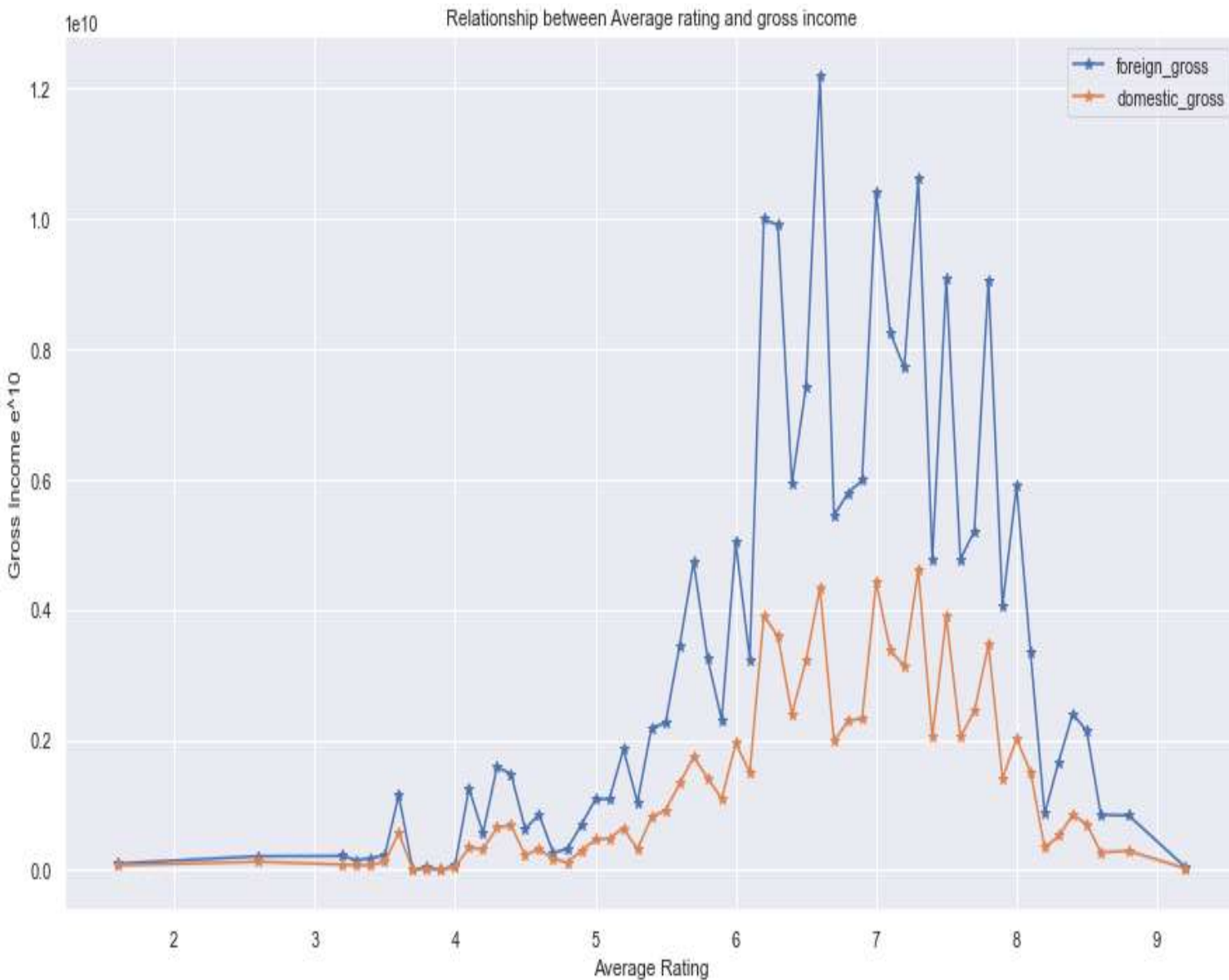- Jupyter notebook
- Github
- Git bash

Effect of originality of title on popularity

DATA VISUALIZATION AND INSIGHTS:

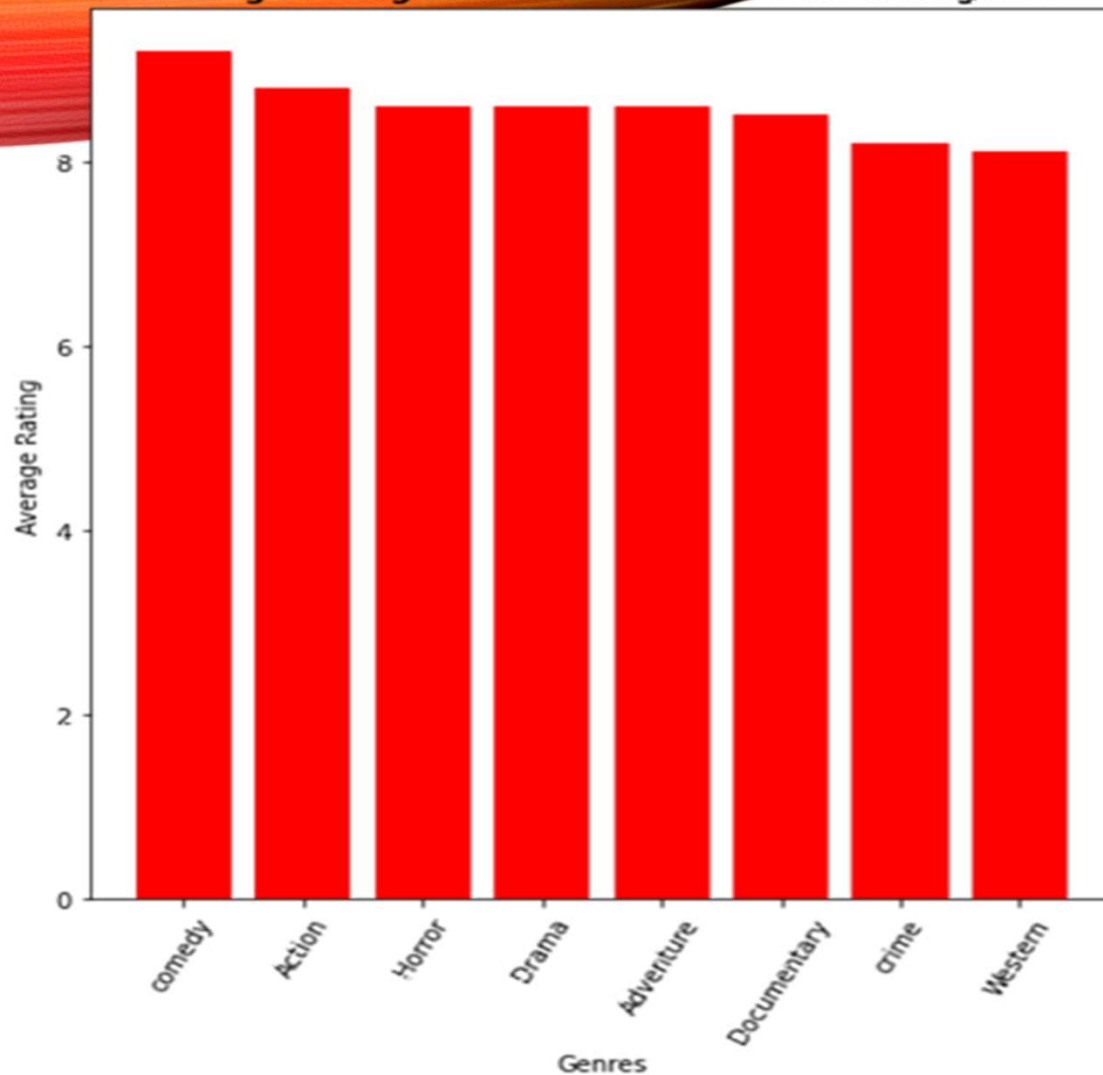In some genres the original titled movies had a higher popularity than the movies without.

Relationship between Average rating and popularity

DATA VISUALIZATION AND INSIGHTS:

From the representation it is clear that the highest popularity is between rating 6 and 7.5.

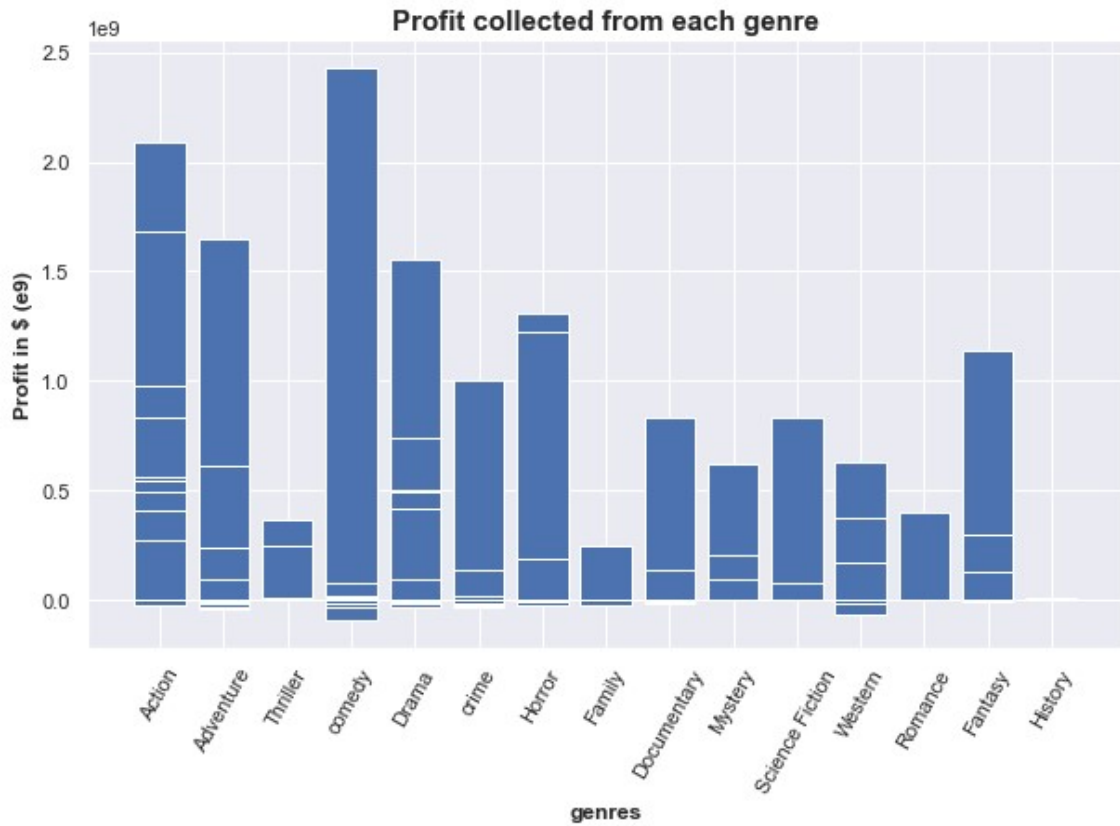Relationship between Average rating and gross income

DATA VISUALIZATION AND INSIGHTS:

From the representation it is clear that the highest income(total of foreign and gross) is between rating 6 and 7.5.

Average Rating above 8.0 obtained from each genre

- The average rating was highest for the comedy.

- The graph alongside shows the genres that had a rating of above 8.0

Profit collected from each genre

Where:
Profit = (foreign_gross + domestic_gross) – production_budget

Comedy had the highest profit in the analysis of the 8 years. With Family genre having the list profit.

CONCLUSION:

- The common genres in the top 4 list of ratings above 8 and popularity include action, drama, adventure, comedy.
- Popularity has a positive correlation with the number of people who gave a rating.
- An increase in the popularity resulted in an increase in the gross income raised from the movie and so did the profit.
- The originality of the title has a negligible correlation with the popularity.
- Popularity have a positive correlation coefficient with the production budget.
- Very high ratings have not been having highest popularity

RECOMMENDATION:

- Microsoft studio should venture into the comedy, drama, adventure and action genres for having the highest rating and being among the top 5 genres with a high popularity.

- Microsoft should concentrate on having very quality video movies since an increase in the production cost which would be as a result of the inputs done into the creation of the movie resulted to an increase in the domestic gross, foreign gross, popularity and average rating of the movie.

- Changing the title of a movie has a very low correlation to the popularity of the movie. This is because on comparing the genres that had titles changed, only 33% had a higher popularity than movies in the same genre with original titles.

- Microsoft should also venture in running ads and other advertisement medias to increase the popularity of the movies on their movie studio. Ensuring that they provide a good system of movie recommendation to their viewers would have them at the top of the studios