



# PROJECT PRESENTATION

## ANALYSIS ON FILM PRODUCTION AND TACTICAL INSIGHTS

---

DATA SCIENTISTS INCLUDE  
SARAH JOSHUA, AUREL OCHIENG,DANIEL KARUE, LEVIS  
GICHUHI, MICHELLE ANYANGO  
MICHAEL NJOROGE.



# BUSINESS PROBLEM



Your company now 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. You are charged with exploring what types of films are currently doing the best at the box office



# THE AIM OF THE PROJECT

---

The project aims to translate findings into actionable insights that the head of our company's new movie studio can use, to help decide what type of films to create.

# GOAL

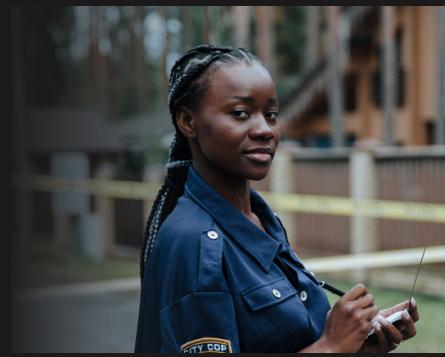
Our main goal as data scientists is to ensure the success of the company's new movie studio by offering the most appropriate business advice.



# KEY OBJECTIVES



- *To evaluate the impact of release timing.*
- *To identify the highly rated genre.*
- *To analyze the performance of films in different budget ranges.*
- *To identify the most produced Genre by other producers.*



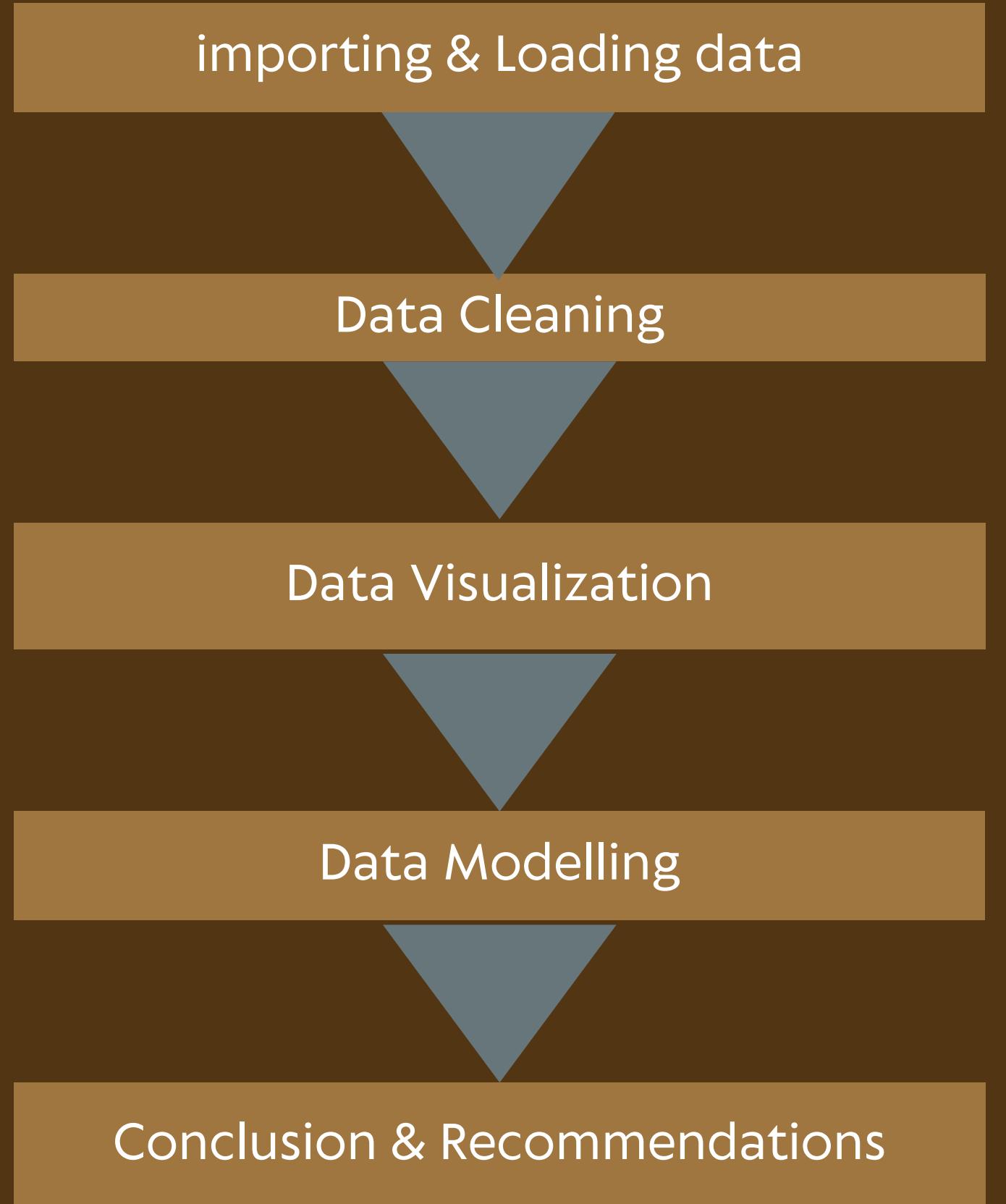
# DATA SOURCE

- Box Office Mojo: Provides comprehensive box office revenue data.
- IMDB: A rich source of movie ratings, genres, and other detailed movie information
- Rotten Tomatoes: Offers critical and audience reviews and ratings.
- TheMovieDB: Contains detailed movie information, including genres, ratings, and runtime.



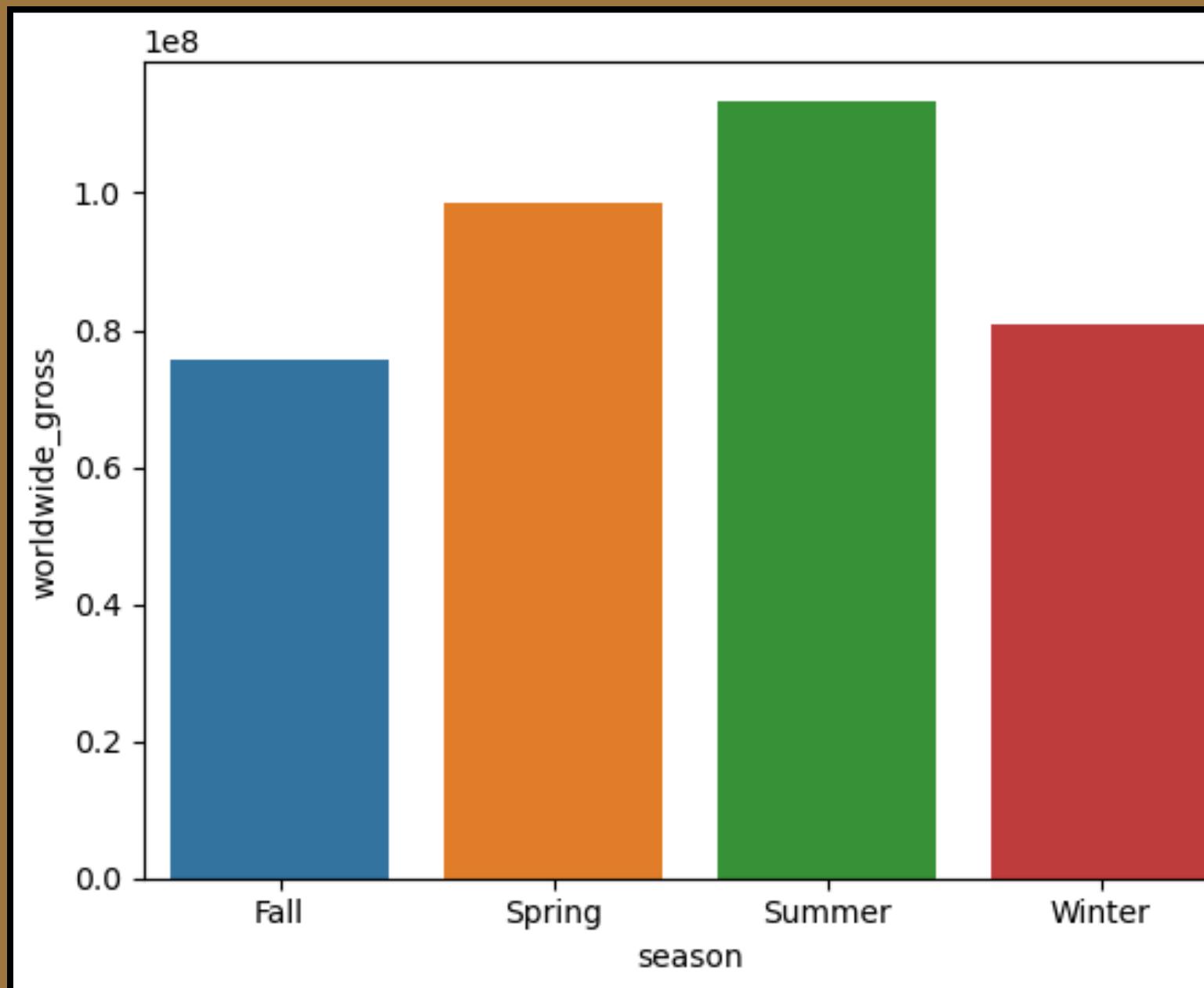
# GENERAL WORKFLOW

---

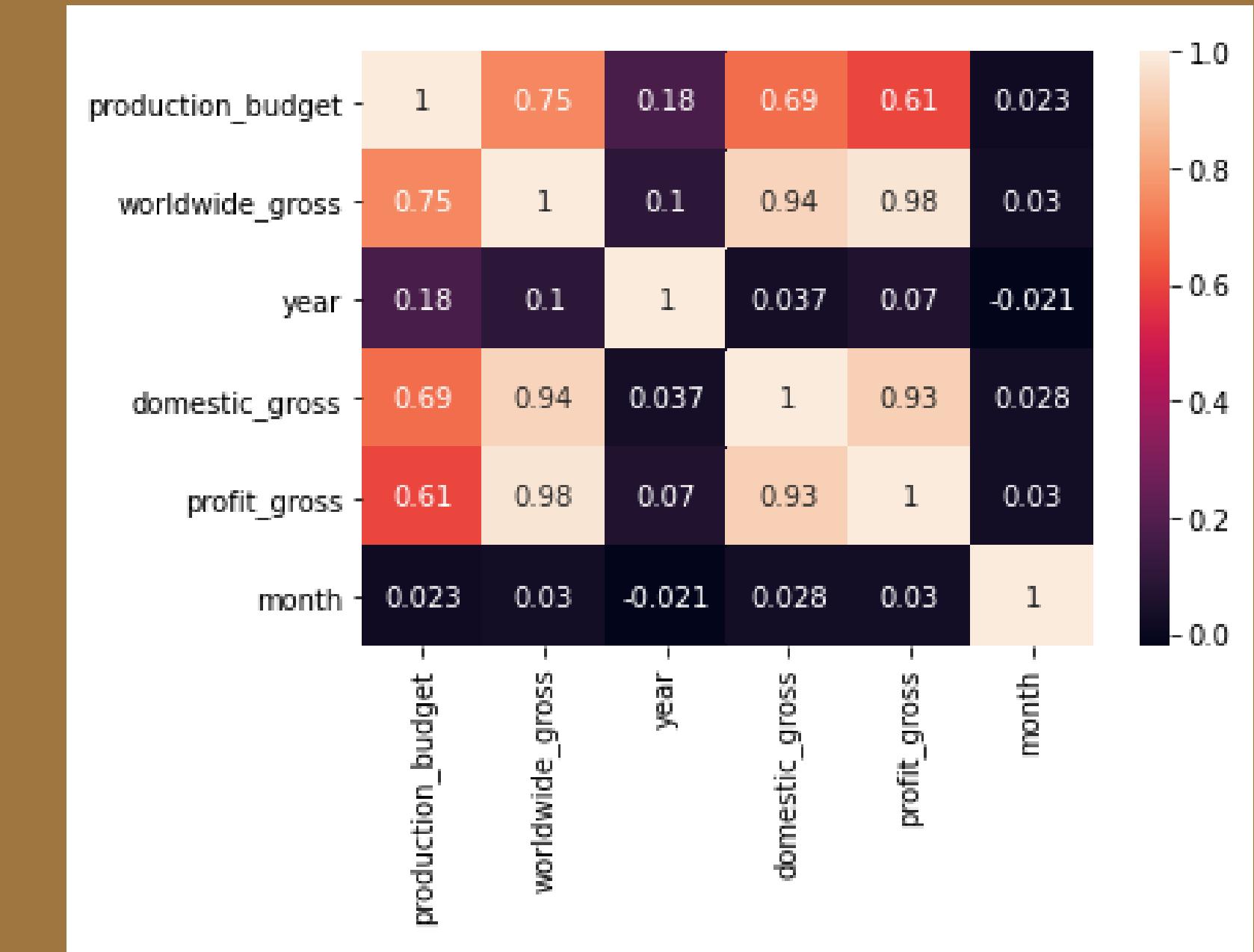


# Results

## 1. Data Visualization

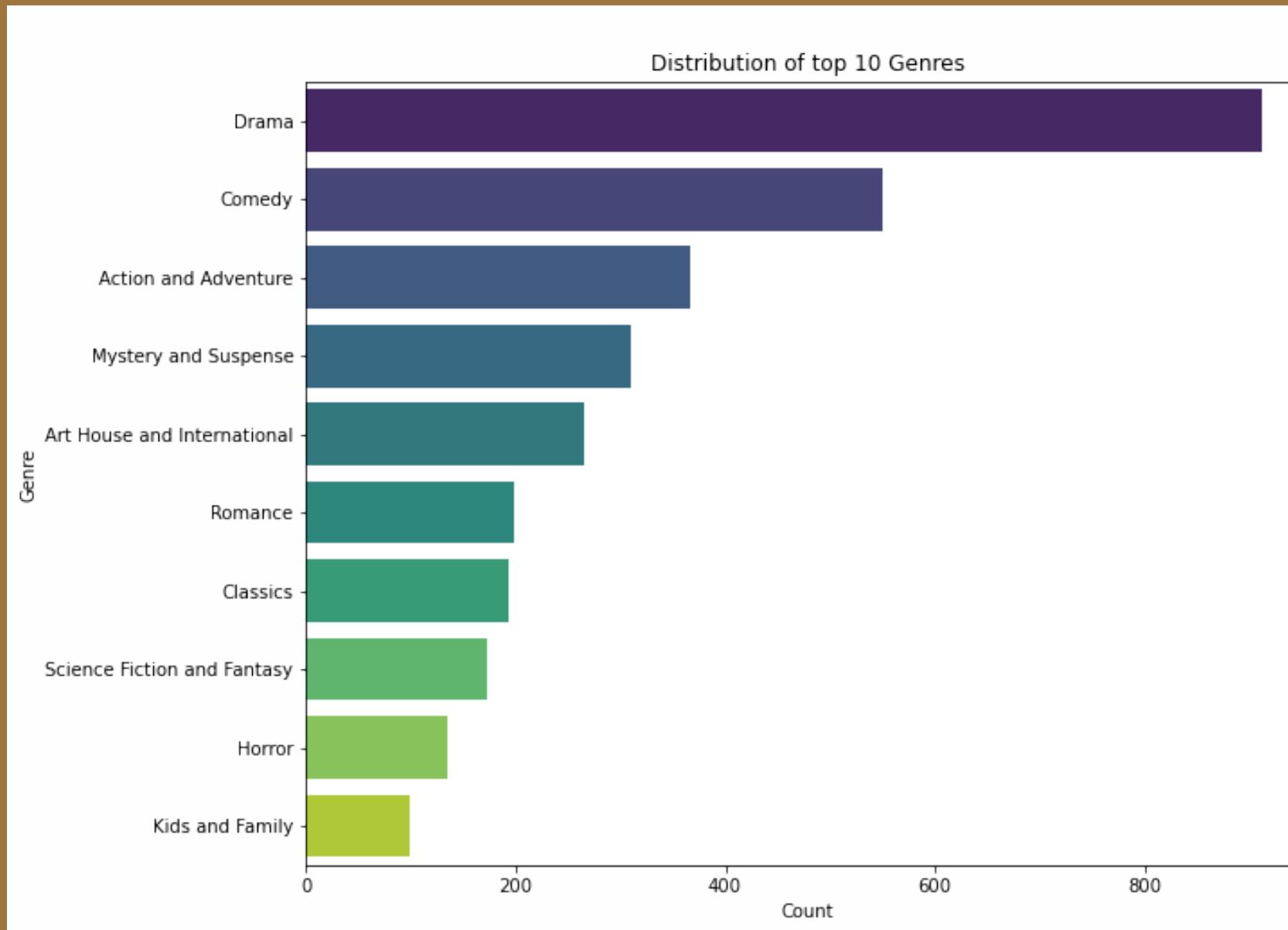


The bar graph above shows that most movies perform best in the summer and least in the fall

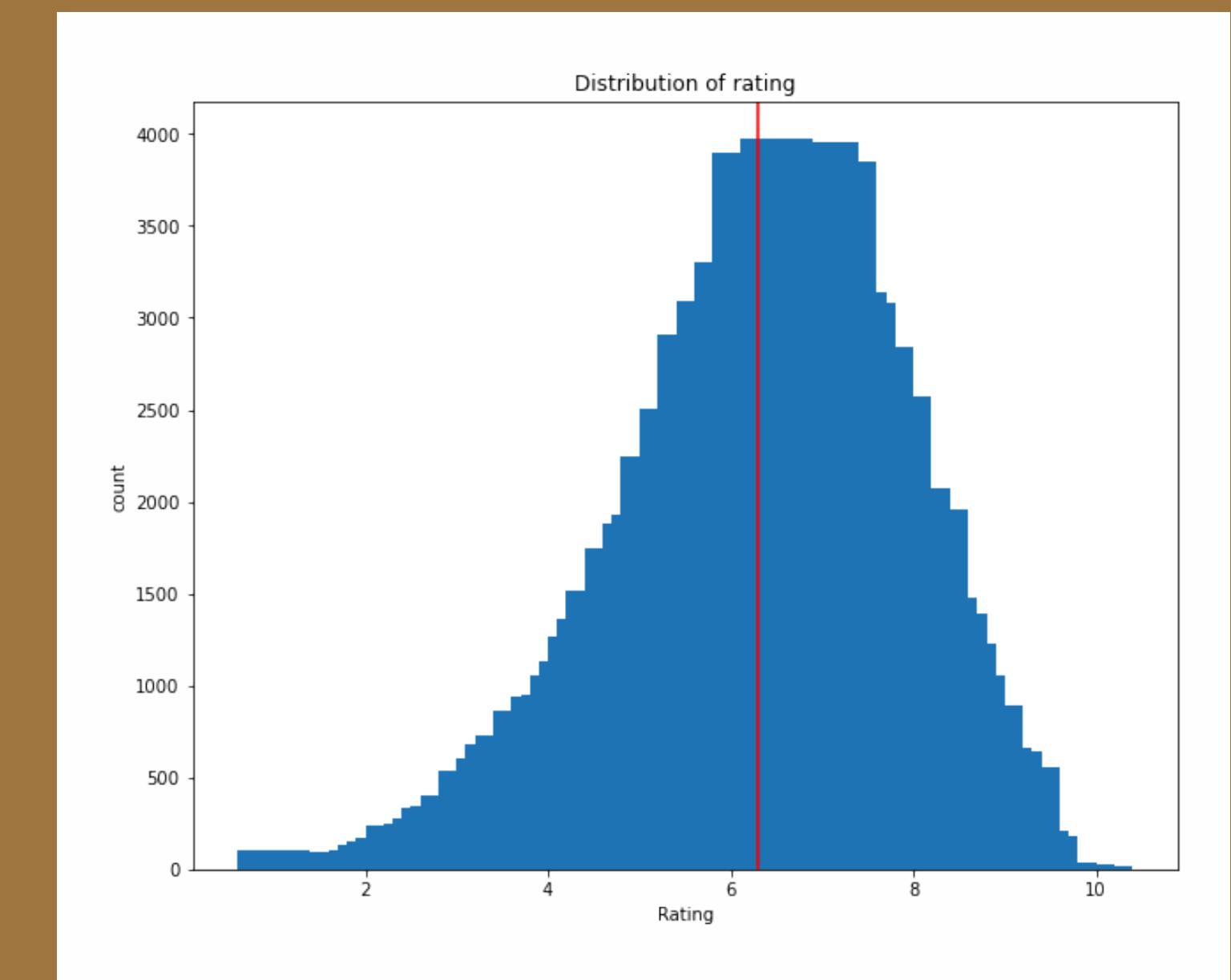


The heatmap above shows how different factors in film relate to each other, with the lighter colours having the highest relationship and the darker colors having the least relationship

# continuee..

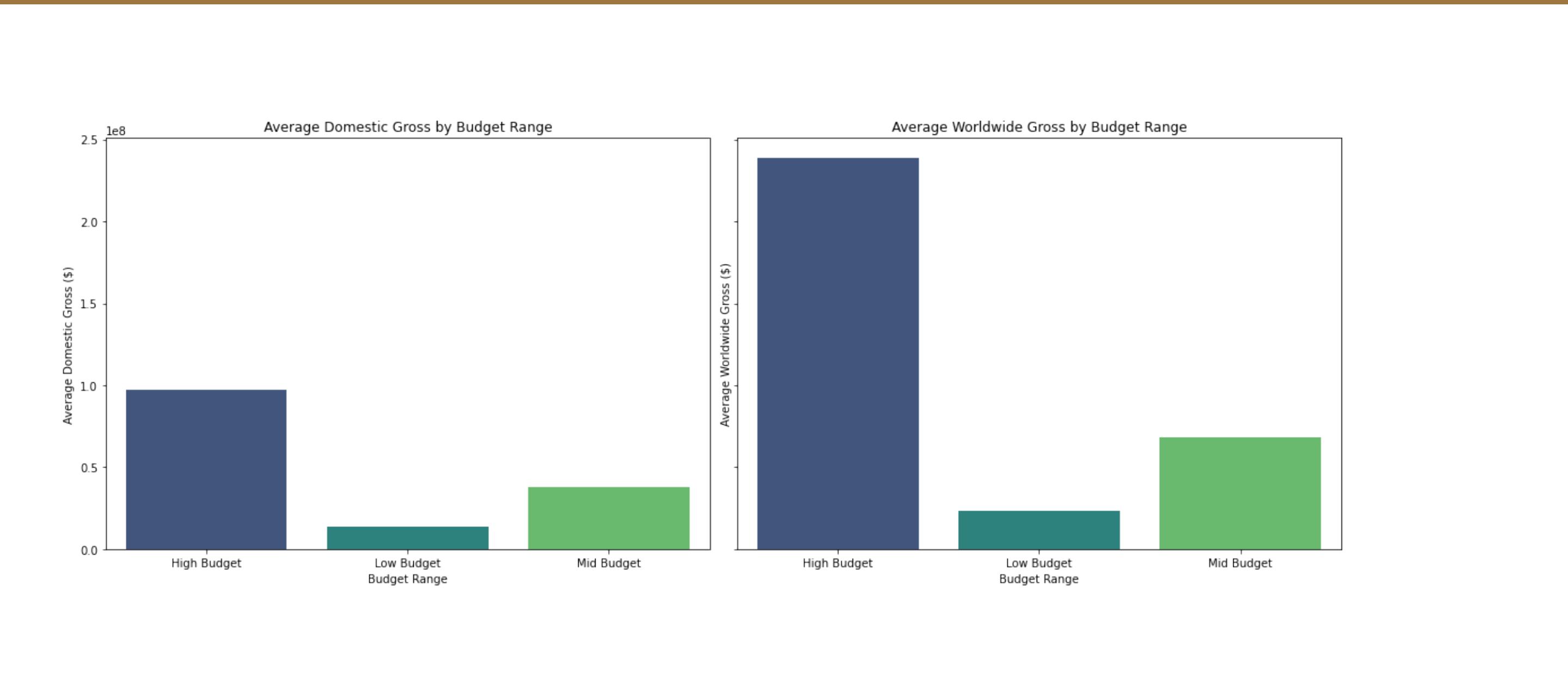


The above plot is a histogram showing the Distribution of the top 10 genres. It shows that the drama genre has the highest count, followed by comedy then action and adventure and so on.



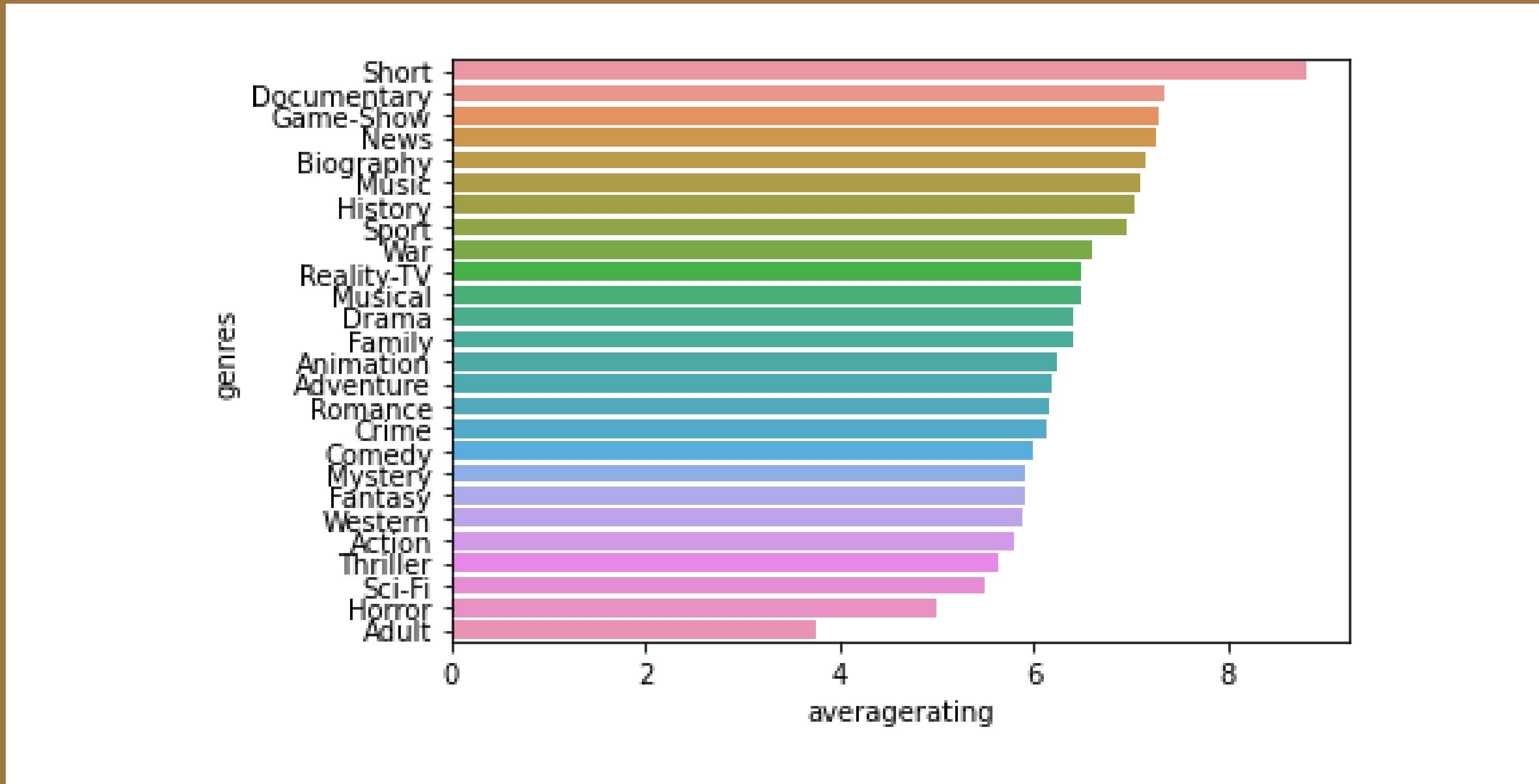
The above plot shows that the movie rating has a normal distribution. It can also serve as an independent variable in our simple linear regression model. It shows that most films have 6.2 ratings.

# continue



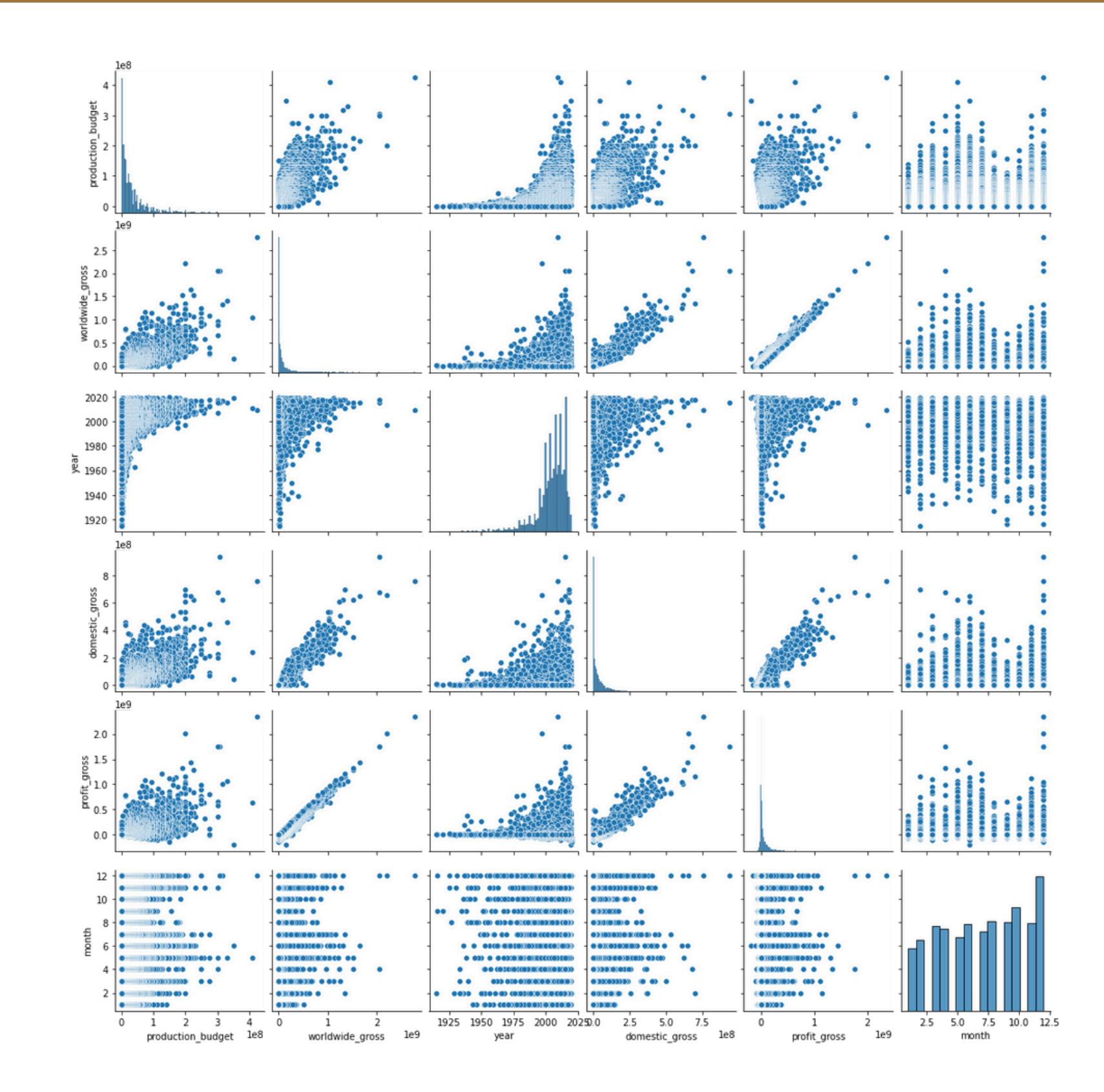
The visualization shows that the higher the production\_budget the higher the returns both domestic and worldwide gross.

# Continue..



The visualisations show that short films are highly rated while adult films are the least rated

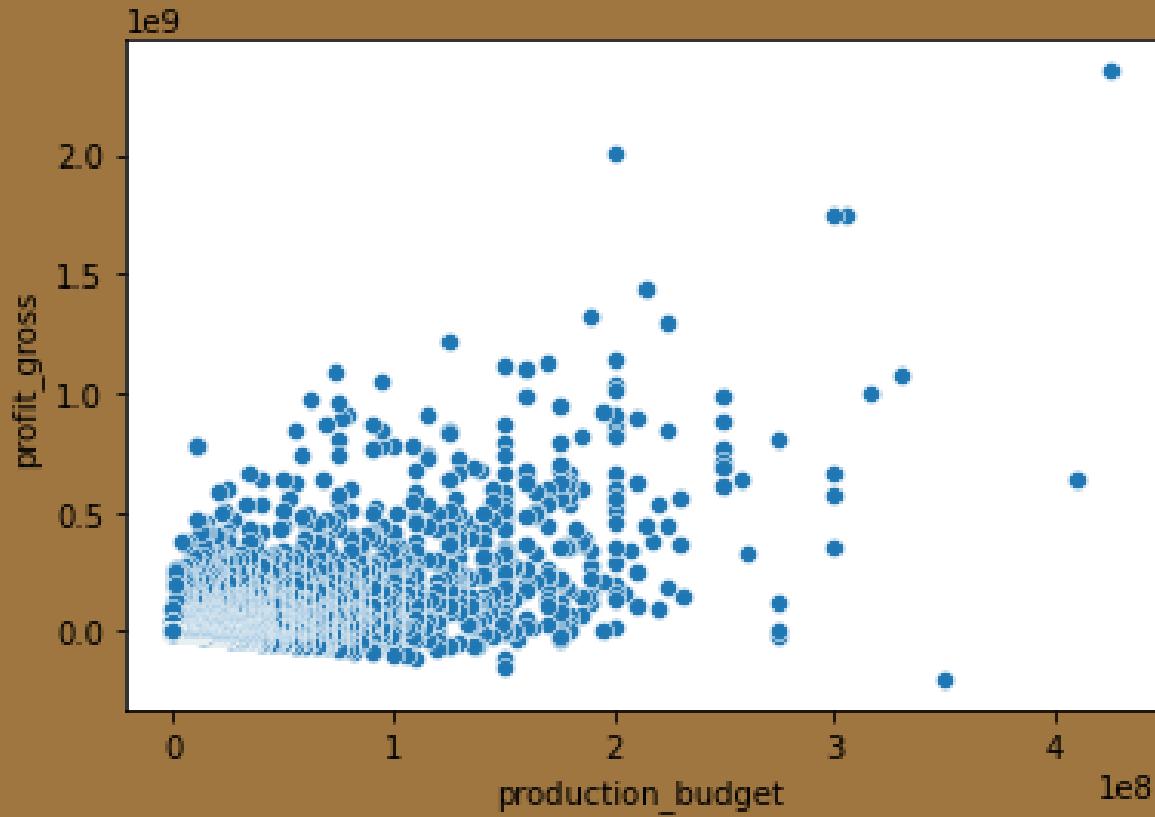
# 2. Data Modelling



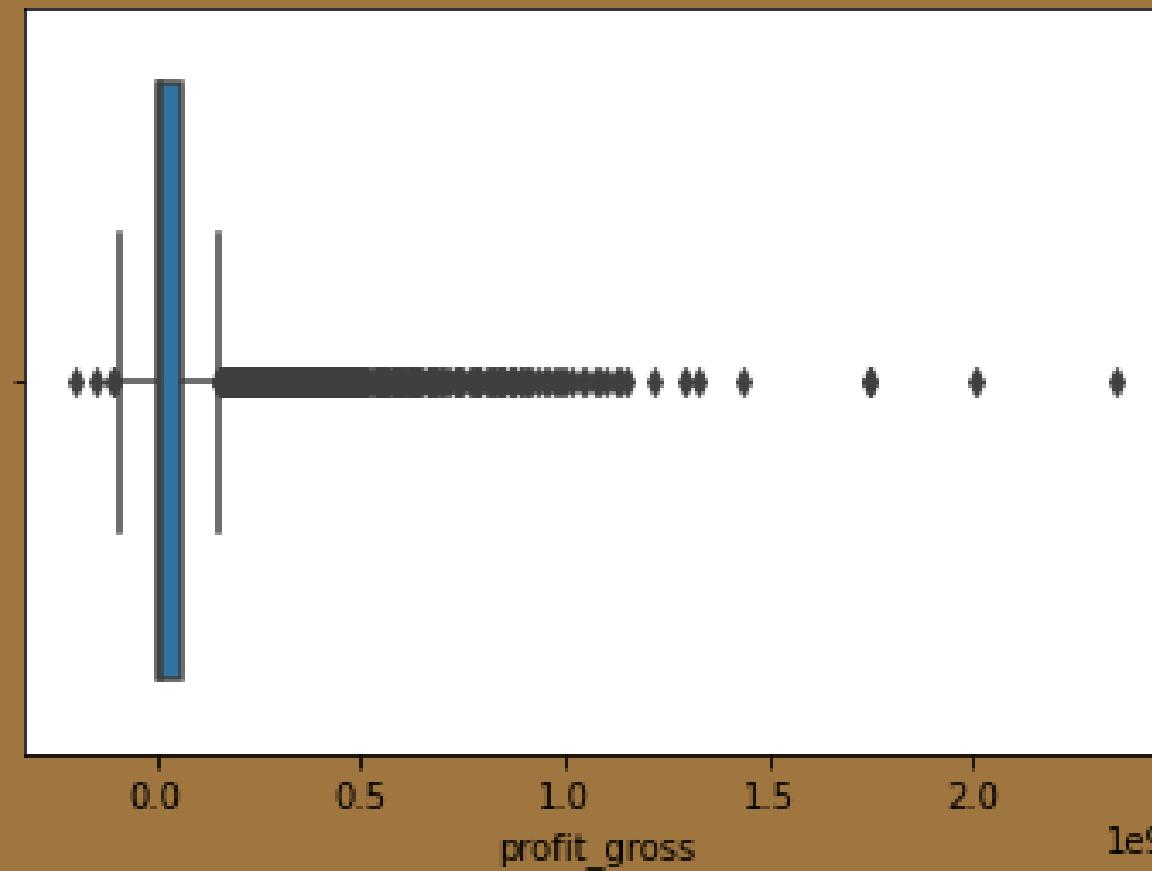
## Preparing modelling data

The visualization the figure provides a comprehensive overview of the relationships among numerical variables in the mb\_mg DataFrame, offering valuable insights for further analysis.

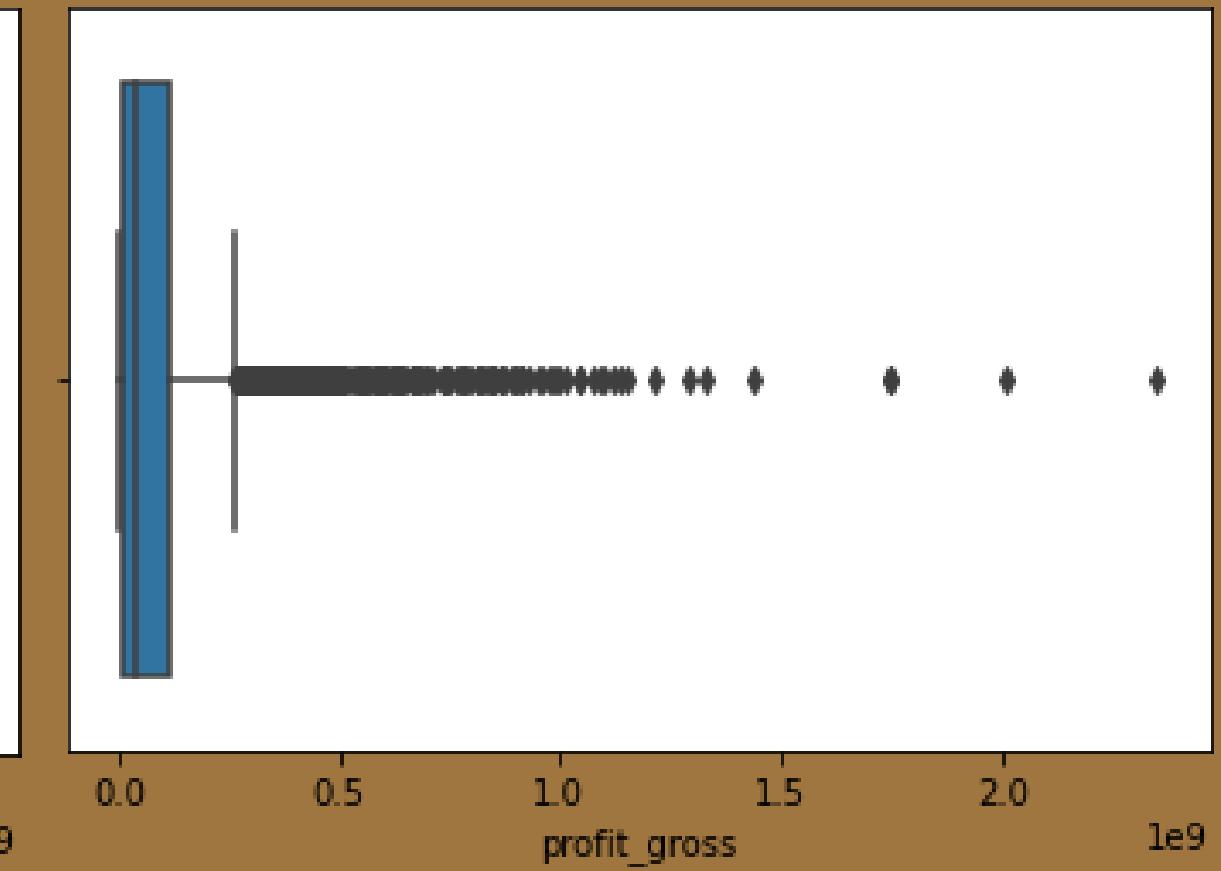
# continue..



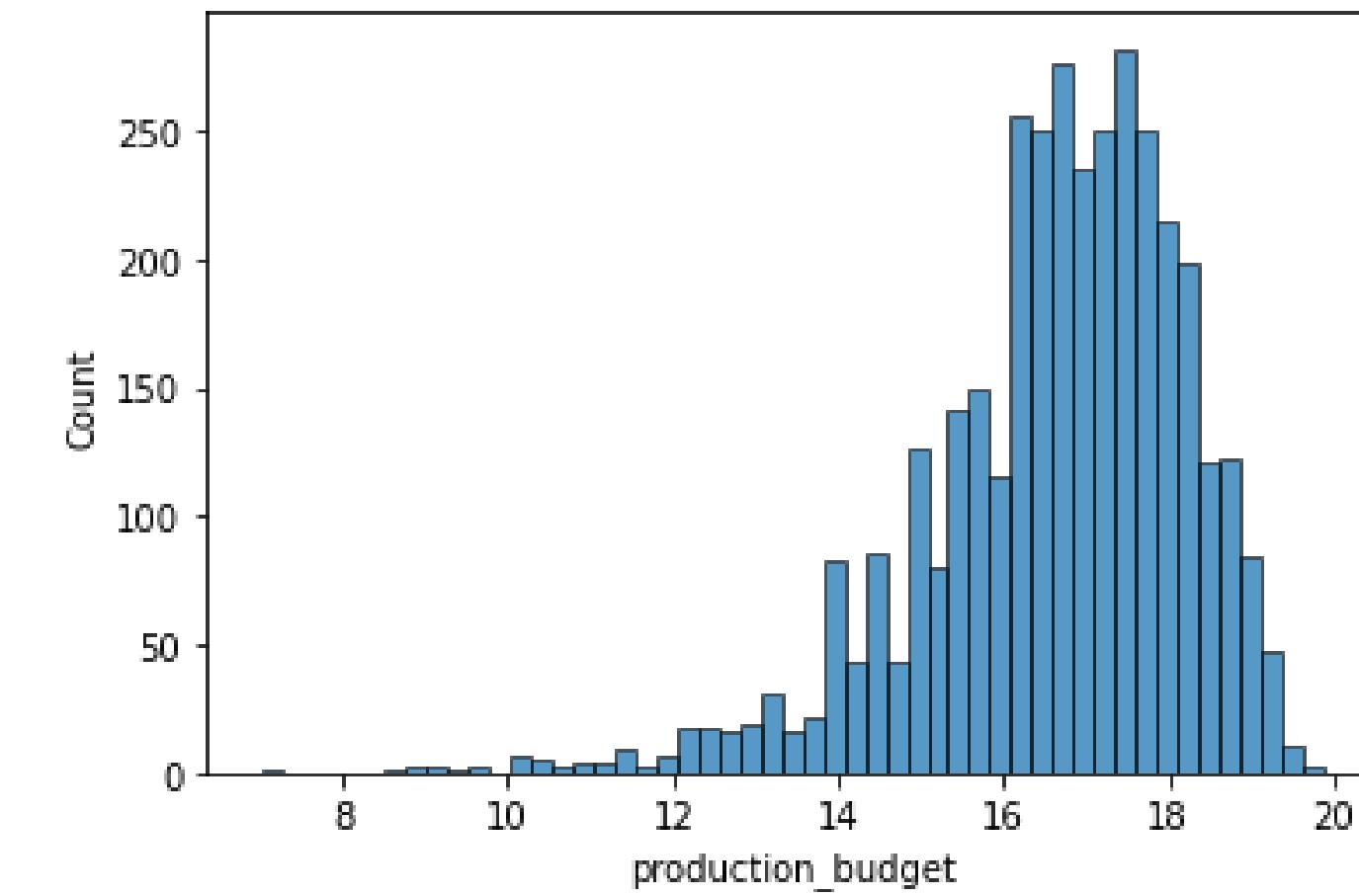
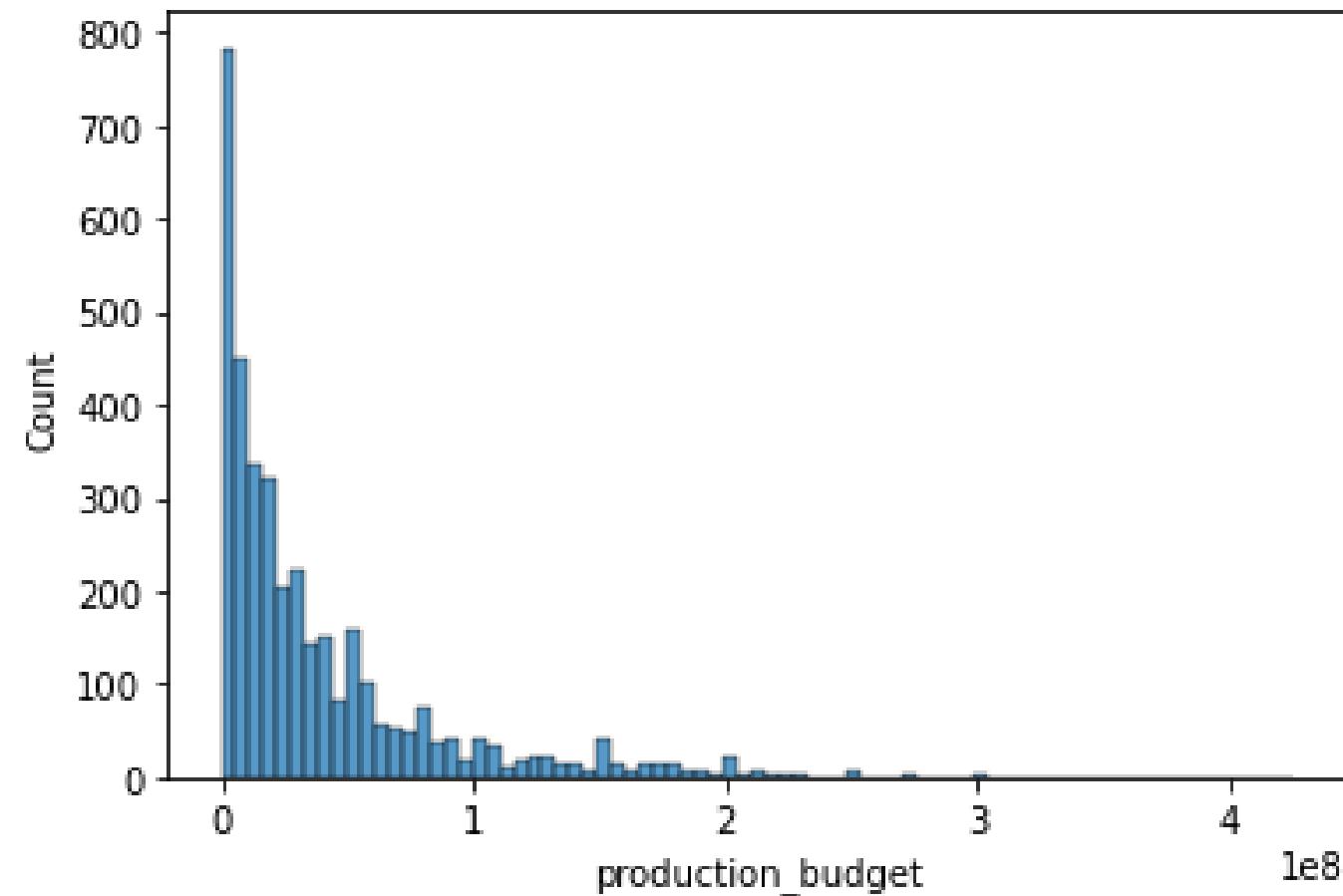
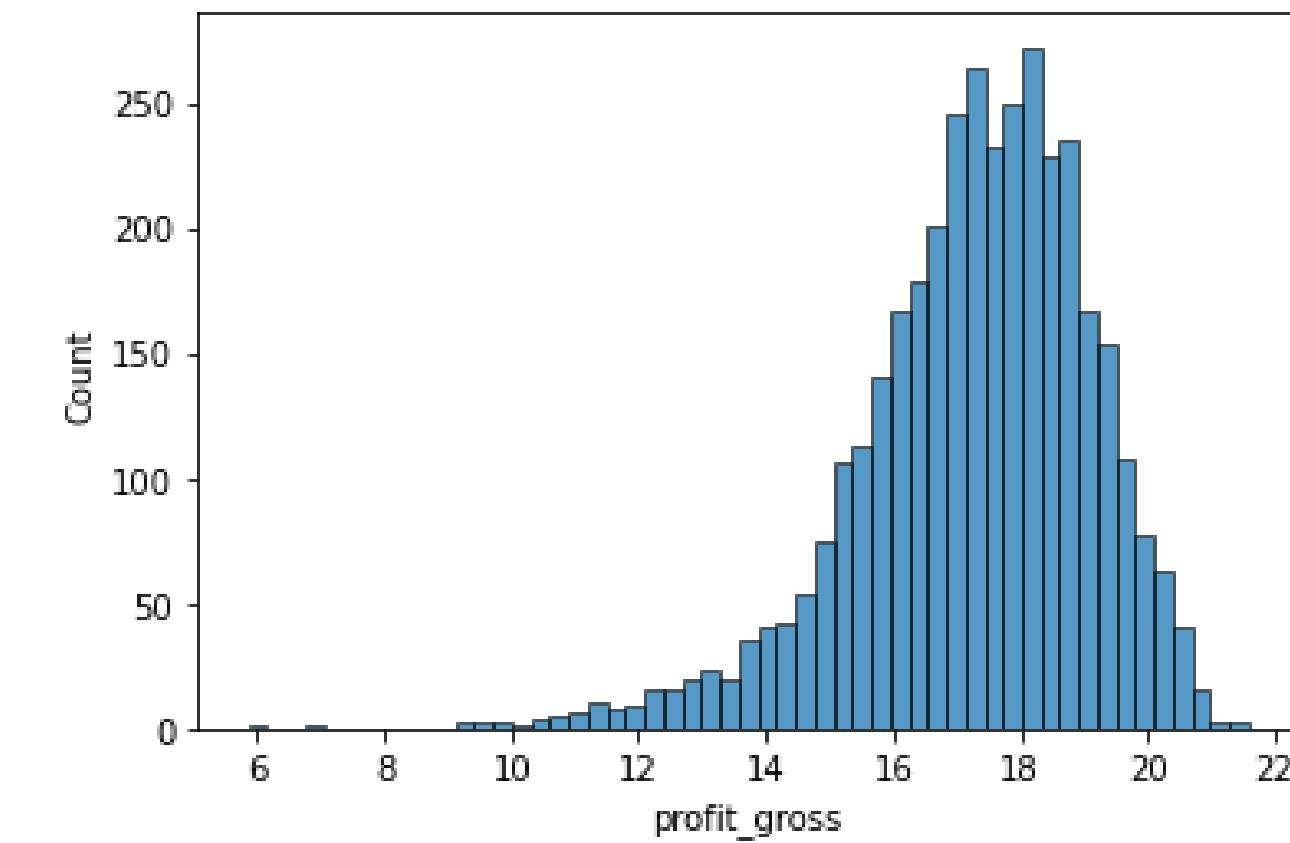
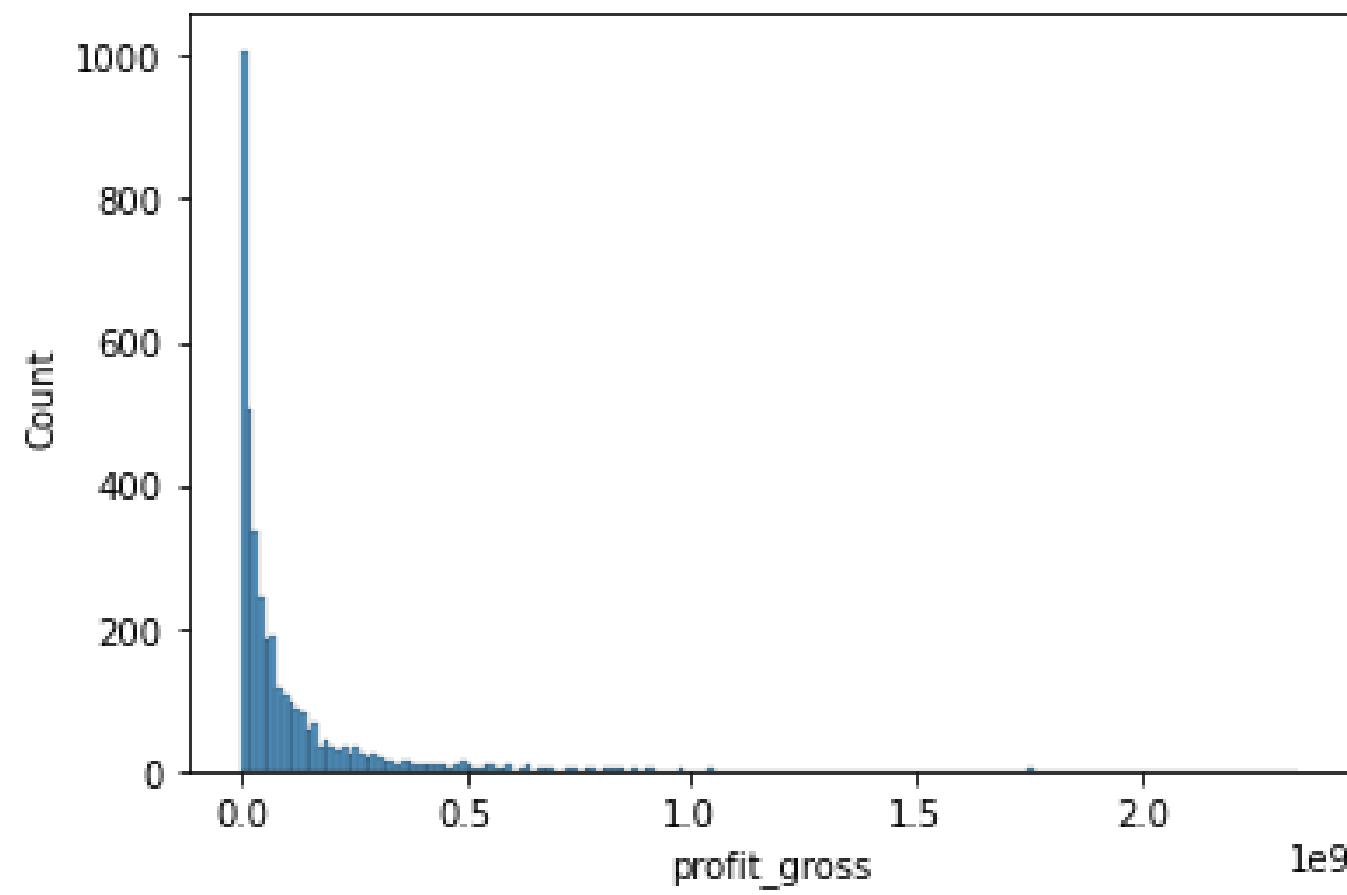
The resulting scatter plot shows the distribution of movies based on their production budget and profit gross. By analyzing the plot, you can identify potential trends, outliers, and correlations between the two variables.



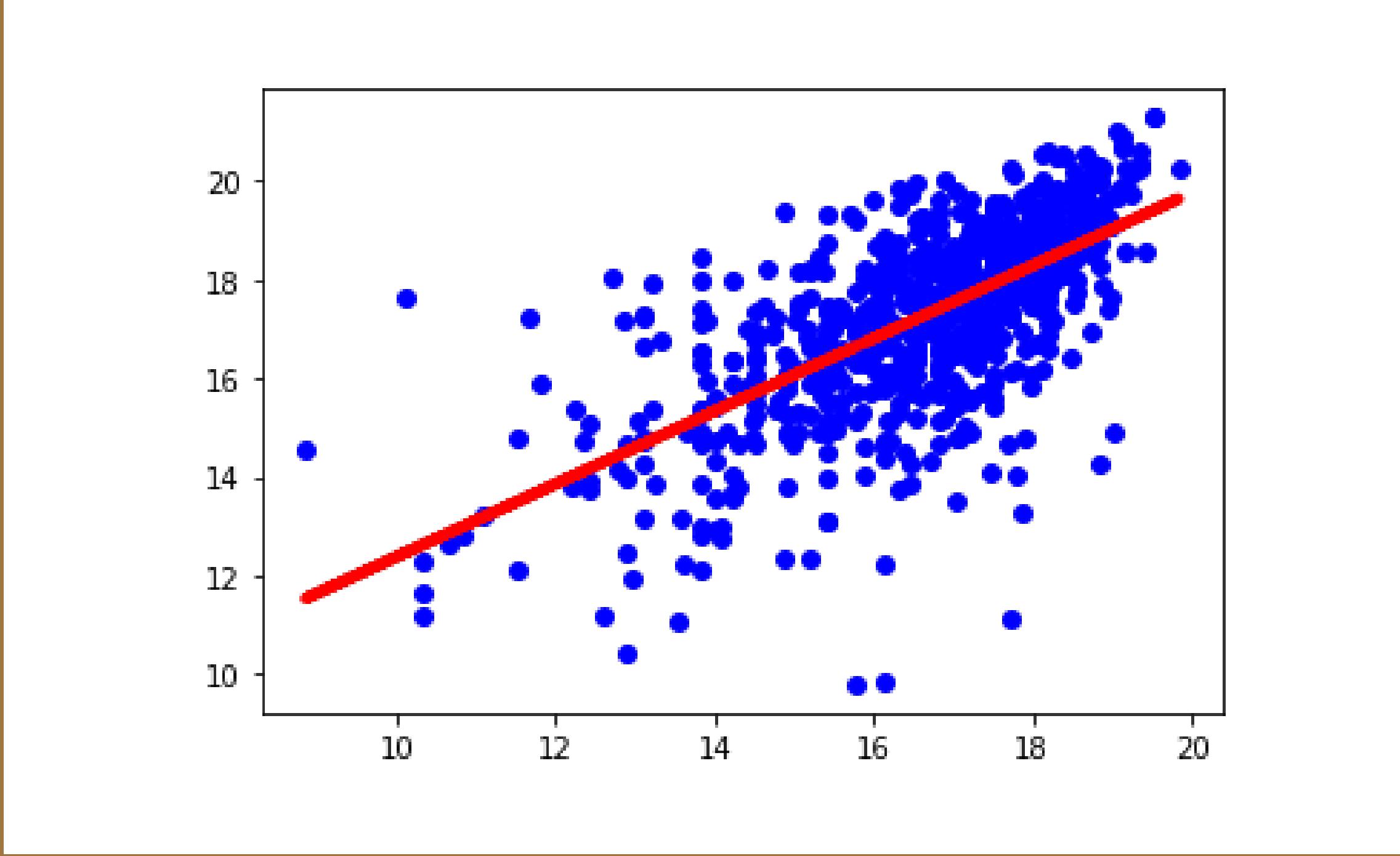
In the box plot above, it shows that there are a few negative values



The above box plot shows that there are no other negative values in the profit\_gross column.



Data distribution using logarithmic transformation



This visualization compares the model's predictions (red line)  
with the actual profit values (blue dots) on the test set

# RECOMMENDATION

---

1. Utilize holidays and festive periods to attract a larger audience, which happens mostly during the summer
2. Focus on producing films in genres that consistently receive high ratings from critics and audiences.
3. Invest in high-budget films with strong potential for high returns, while also exploring lower-budget films with lower risk.
4. Identify popular genres that are currently in demand and consider producing films in these genres.



# CONCLUSION

---

1. By understanding the relationship between budget and box office performance, studios can make informed decisions about budget allocation and risk management.
2. By focusing on high-rated genres, studios can attract a wider audience and improve their chances of critical and commercial success.
3. Studio owners can optimize their release calendars by carefully considering release timing to maximize box office performance and minimize competition.



CANVA STORIES F20

CANVA STORIES F20