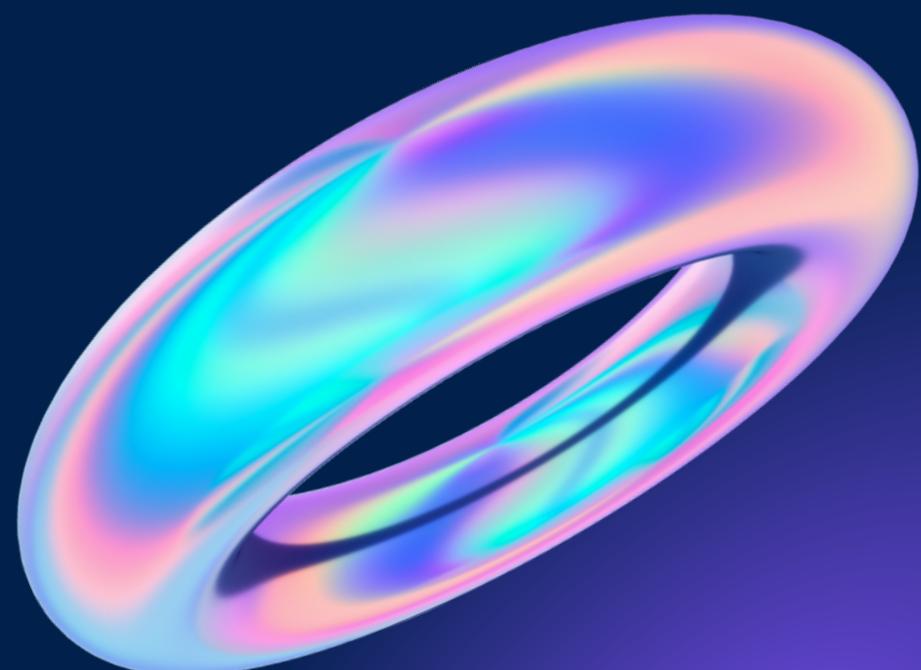




# Microsoft's Movie Studio



Presentation



# Introduction

Microsoft is looking to create a new movie studio, this presentation is exploring what types of films are currently doing the best at the box office and translates findings into actionable insights.



# Objectives

1. To get the most popular genres
2. To check the most popular studios
3. To use the domestic gross to estimate the budget used and how it affects the ratings.

# Data Sets Used



**Box Office**

A worldwide organization dedicated to innovating the box office through rich technology & data for the entertainment industry



**Rotten Tomatoes**

The most trusted measurement of quality for Movies & TV



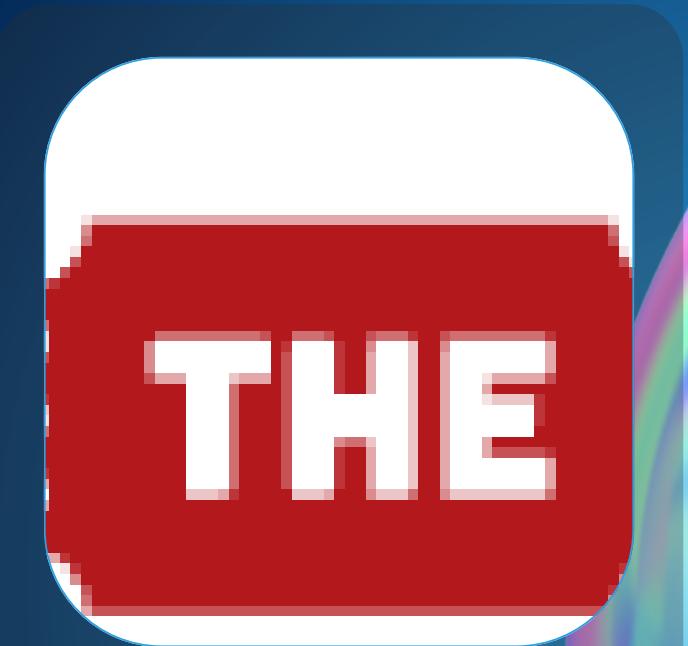
**The Movie DB**

The Movie Database (TMDB) is a popular, user editable database for movies and TV



**IMDb**

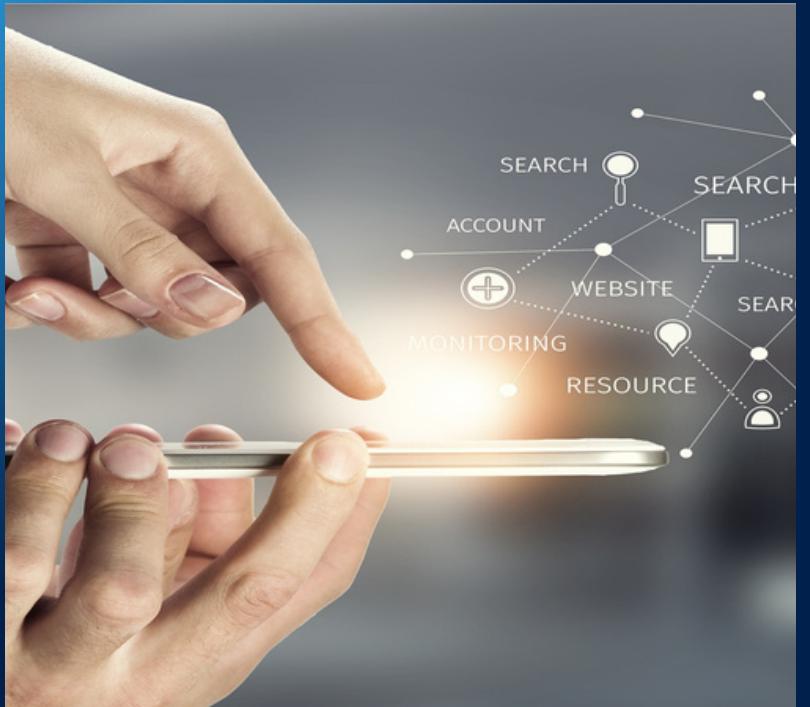
The world's most popular and authoritative source for movie, TV, and celebrity content. Find ratings and reviews for the newest movies and TV shows.



**The Numbers**

Where data and the movie business meets.

# Methods Used



## Data Processing

1. Loading the data
2. Inspecting the data

	0	0	0	(1,000)	0	0	(186)
Net cash provided by financing activities	208	660	544	837			
Common stock issued	(1,042)	(5,052)	(2,976)	(9,451)			
Common stock repurchased	(1,683)	(1,363)	(3,024)	(2,481)			
Net cash used in financing activities	(2,513)	(6,751)	(5,382)	(7,390)			
Net investing activities	(498)	(491)	(934)	(1,055)			
Additions to property and equipment	(8,627)	(69)	(9,502)	(69)			
Acquisition of companies, net of cash acquired, and purchases of intangible and other assets	(10,047)	(5,896)	(21,346)	(13,313)			
Decreases of investments	6,061	1,836	8,886	2,706			
Dividends paid	7,835	2,603	15,371	4,030			
Net investing activities	(292)	447	(358)	1,174			
Net investing activities, equivalents, end of period	(5,568)	(1,570)	(7,883)	(6,527)			
	\$ 10,610	\$ 4,023	\$ 10,610	\$ 4,023			

## Data Cleaning

1. Separate datasets
2. Duplicated values
3. Missing values
4. Conversion of data type



## Data Exploration

1. Highest number of production genres
2. Common studios used in production
3. Domestic gross effect on ratings
4. Correlations of genres and other attributes.

# RESULTS

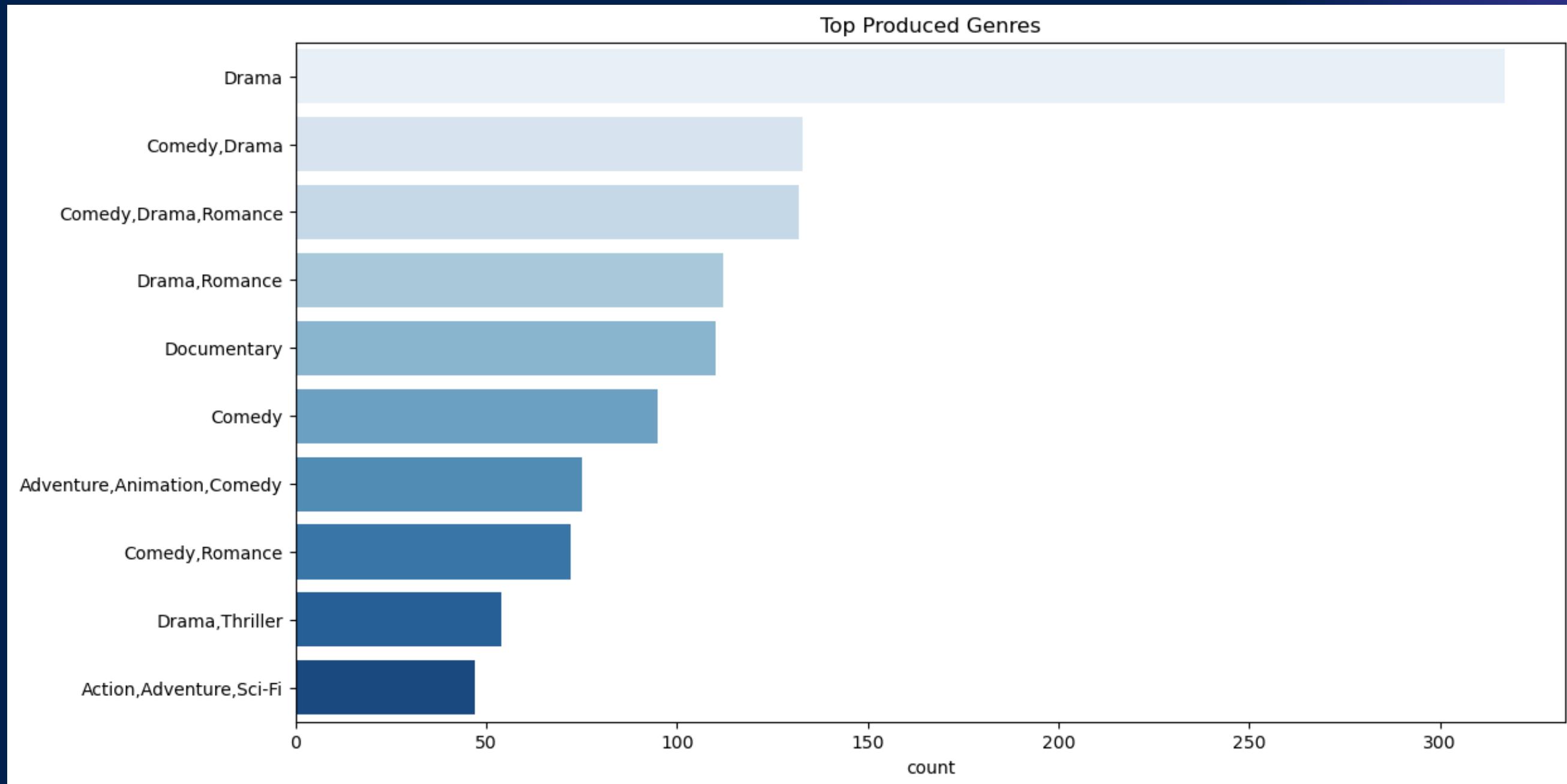
I used the following analysis methods to visualize the data to help us identify the patterns, trends, and relationships

1. Univariate Analysis
2. Bivariate Analysis
3. Multivariate Analysis

The presentations are in the form of

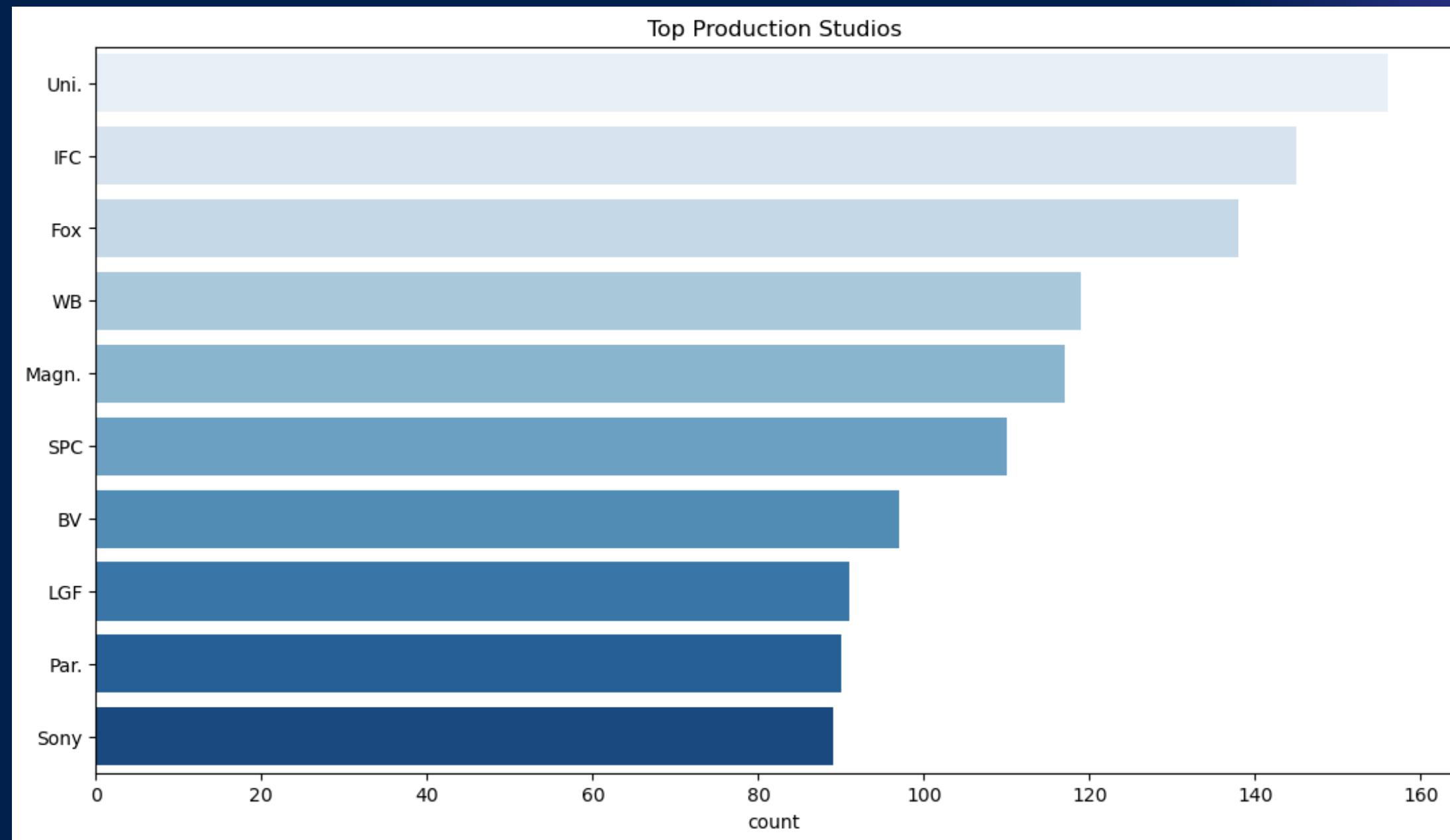
1. Bar plots
2. Scatter plots

# Top Produced Genres



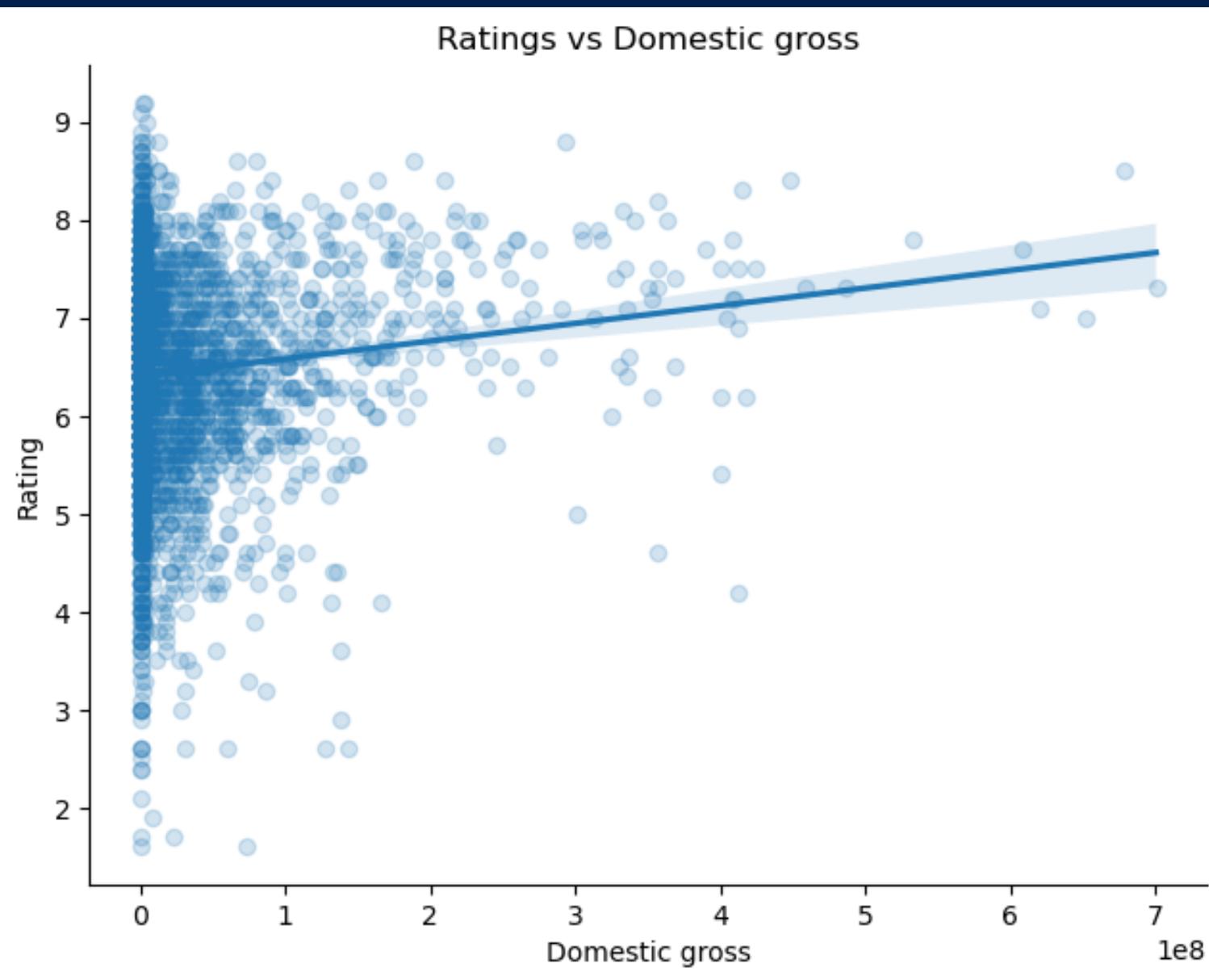
Drama is the leading genre

# Top production studio



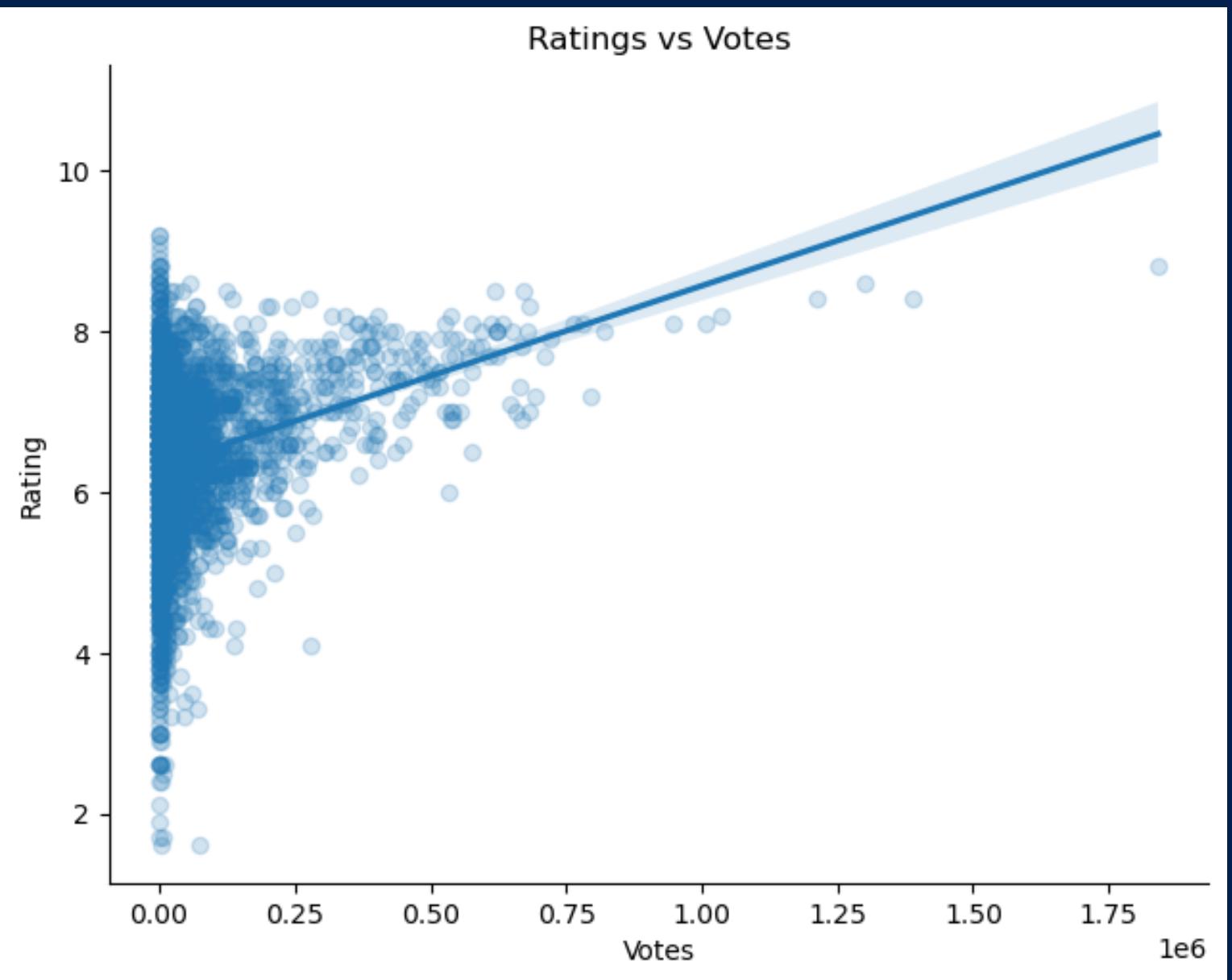
Universal Studios is the leading studio

# Ratings vs Domestic Gross



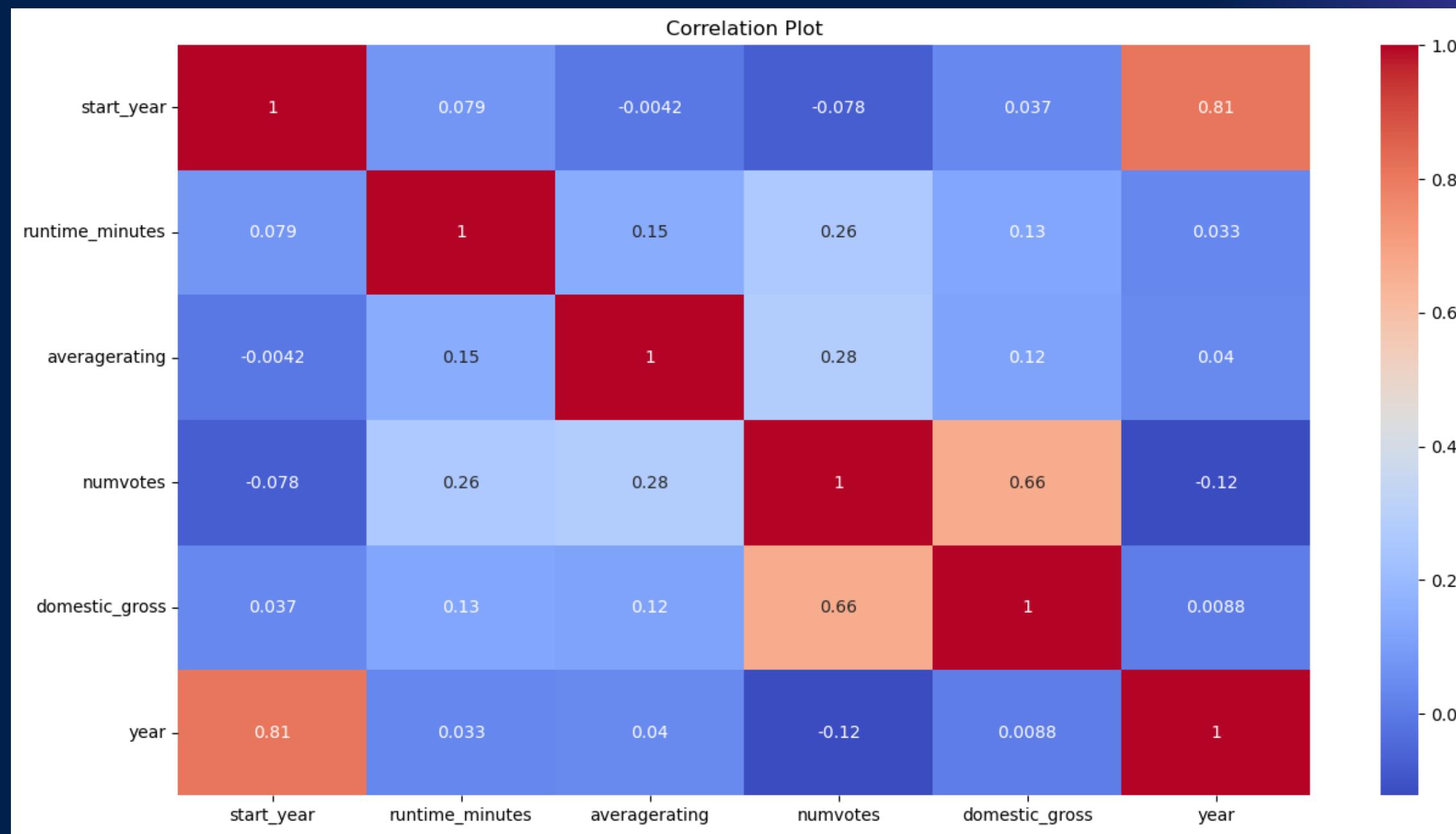
Higher domestic gross positively affects the movie ratings

# Rating vs Votes



**Taking votes into consideration before movie production increases chances of higher ratings**

# Rating vs Votes correlation matrix



The light shades show a high correlation between the elements while dark shades show a low correlation between the elements. For, example, the averagerating has a weak positive correlation with total\_gross meaning that a high rating indicates a descent gross return.

# CONCLUSIONS

We have drawn many interesting inferences from the dataset, here's a summary of the few of them:

The most genre type is drama with the least as war,

The amount of Domestic Gross/Budget does affect the rating of the movie,

Some Genre Combinations are more relatable to each other unlike others.

The number of votes also affect the rating of the movie.

# RECOMMENDATIONS

**The studios that should be invested in as partners or for collaborations should be either Universal, IFC or Fox.**

**The company should consider starting with the production of drama genre movies.**

**The company should conduct a survey to see which movies are popular.**