MICROSOFT’S CINEMATIC ENDEAVOUR: A DATA-BACKED APPROACH.

OVERVIEW

Microsoft Corporation Background.

Microsoft Corporation, frequently referred to as just "Microsoft," is a major player in the software sector and a leader in the field of worldwide technology. Microsoft, established in 1975 by Bill Gates and Paul Allen, has continuously changed how we utilize computers and digital technologies. Microsoft's current mission statement is, "Empower every person and every organization on the planet to achieve more." The business is still at the vanguard of technical advancement, influencing how computing, cloud services, artificial intelligence, and digital experiences are developed globally.

Microsoft Corporation Domain.

Microsoft's products, including the Windows operating system, Office program, and Azure cloud computing platform, have become indispensable to people and companies all over the world. Microsoft has a strong presence in hardware, gaming through Xbox, and AI research in addition to software. Microsoft is dedicated to innovation and is on a mission to enable everyone and every business to achieve more.

Business Case.

Despite having no past expertise in the field, Microsoft hopes to build a successful studio with its foray into filmmaking. The main objective is to have a thorough grasp of the movie industry, in particular, the kinds of movies that routinely perform well. Microsoft wants to find insightful information by thoroughly analyzing popular films. The corporation will be able to make informed decisions about the genres, content, budgetary constraints, and marketing plans for its own film projects thanks to this information. Microsoft has a lot of technology know-how that may be used to streamline the film production process. The company's extensive data analytics capabilities also allow for data-driven decision-making in the choice of movie genres, budget allocation, and marketing techniques, ultimately improving the likelihood of creating successful films.

BUSINESS UNDERSTANDING

Entering the world of studio movie production demands a comprehensive understanding of its intricacies. This journey begins with script development, collaborating with seasoned screenwriters and directors to select engaging stories. Pre-production requires a delicate balance between artistic freedom and cost management, while filmmaking and post-production demand meticulous management of schedules and talent. Successful marketing and distribution strategies, guided by data-driven insights, are pivotal. Microsoft must stay attuned to industry shifts and emerging technologies, exercise prudent financial management, and foster creative partnerships. This venture combines technological expertise with storytelling, aligning with Microsoft's mission to empower creativity in the digital era.

OBJECTIVES

1. **Problem Statement 1: To assess the success of different languages based on their worldwide gross earnings.**

   Description: We want to determine which languages now control the global by comparing the global gross revenues of films released in a variety of languages. Microsoft will be able to make educated decisions about the preferred languages for its film productions.

2. **Problem Statement 2: To profile the competitive landscape within the movie industry, including the performance of key players.**

   Description: This inquiry aims to analyze the competitive dynamics of the film business. To comprehend their respective performances, it entails a detailed research of the studios. Microsoft can make strategic decisions about prospective alliances, collaborations, or areas of opportunity.

3. **Problem Statement 3: To assess the financial implications of production budgets on the performance of movies.**

   Description: This explores the financial aspects of the film industry, focusing on the connection between production costs and total film quality. by looking at how spending on a film's budget affects its commercial success. These data will enable Microsoft's film division to allocate resources intelligently, assuring the best possible financial results while creating high-caliber cinematic experiences.

DATA UNDERSTANDING

We tapped into insights from many databases that offer a comprehensive view of the business in our investigation of the environment of film production. These datasets include:

1. Box Office Mojo: This dataset contains essential details about movie performance, including the Movie Title, Studio responsible for production, Domestic gross earnings, Foreign gross earnings, and the Release year.
2. The Movie Database: The second dataset from TMDB offers insights into movie attributes such as Genre ID, Original Language, Original Title, Popularity, Release Date, Title, Vote Average and Vote Counts.
3. The Numbers: The third dataset, sourced from The Numbers, focuses on the financial aspects of movies, encompassing Release Dates, Movie Titles, Production Budget, Domestic gross earnings, and Worldwide gross earnings.

Together, these datasets offer a complete picture of the movie business, including measurements for movie performance, financial information and movie qualities. These datasets are directly addressed by the data analysis questions, which seek to glean valuable insights for Microsoft's filmmaking endeavor.These data links enabled us to create a substantial and well-organized dataset that forms the basis of our analysis. The dataset's important columns are listed below:

1. Title: The title of the film serves as an identification number.
2. Studio: The entity responsible for production of the movie.
3. Genre ID: The genre ID representing the genre of the movie.
4. Original Language: The language in which the film was first made.
5. Popularity: A measure of a film's popularity.
6. Vote Average: The overall score that viewers gave the film.
7. Vote Count: The total number of votes and reviews received by the movie.
8. Production Budget: The sum of money set aside for the movie’s production.
9. Domestic Gross: Total revenue made from the movie's domestic release in the nation where it was produced.
10. Worldwide Gross: The sum of the movie's worldwide box office receipts.

DATA CLEANING

Data preparation is a crucial stage in this project for a number of reasons, including:

- Data Quality Assurance: Ensuring data quality is crucial. Errors, missing values, and inconsistent data are all possible. We can find and fix these problems by carefully gathering the data, ensuring the dependability of our analysis.
- Data Integration: We are utilizing a number of datasets from various sources. By addressing differences in variable names, units, and formats, data preparation enables us to harmonize these datasets, allowing for meaningful cross-referencing and analysis.
- Feature engineering: New features might be developed or current ones modified in order to improve analysis.
- Handling Missing Data: Analysis results can be greatly impacted by missing data. In order to achieve a robust analysis, handling missing values must be decided, whether through imputation, deletion, or other suitable approaches.
- Outlier detection: For statistical validity, outliers must be found and dealt with. We can use methods like visual inspection or statistical testing to find outliers and handle them correctly with the help of data preparation.

The dataset is now clean and ready for analysis. To improve the data, several crucial actions were conducted. At first, extraneous columns were removed, yielding a more concentrated dataset of shape (1395, 10). A single missing value in the'studio' column was quickly deleted when we meticulously used the `.isna()` method to check for missing values. As a consequence, we were left with a dataset that had the shape (1394, 10). The next crucial step was to use the `.duplicated()` method to find and handle duplicates. This revealed 124 duplicated values that were immediately removed, giving us a DataFrame with the shape (1270, 10). We data casted the `production_budget`, `domestic_gross_y`, and `worldwide_gross` columns to float data type to guarantee numerical consistency. Additionally, we used the `.str.replace()` method to remove dollar signs and commas from these columns and the `.astype()` method to convert the data types. These thorough data cleaning procedures have now prepared the dataset for intelligent analysis.

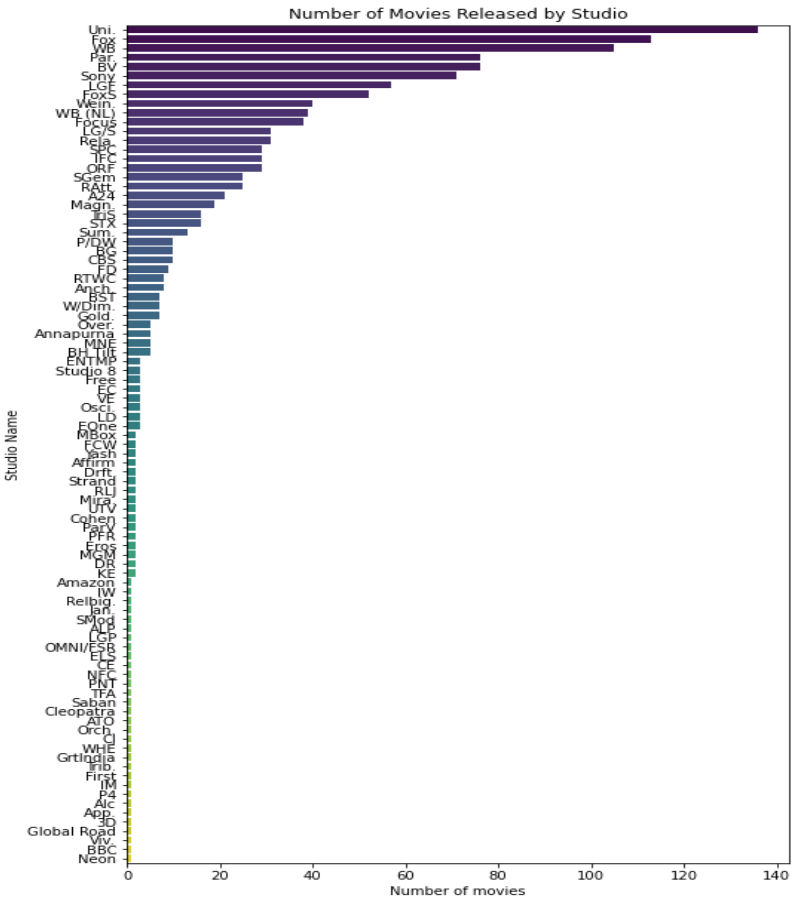
DATA ANALYSIS

A crucial turning point in our investigation of the film industry occurs during the data analysis stage, which enables us to extract significant patterns and insights from the compiled dataset. We used univariate, bivariate, and multivariate exploratory data analysis (EDA) methodologies in a comprehensive manner to properly analyze the data. Let's dive into our analysis.

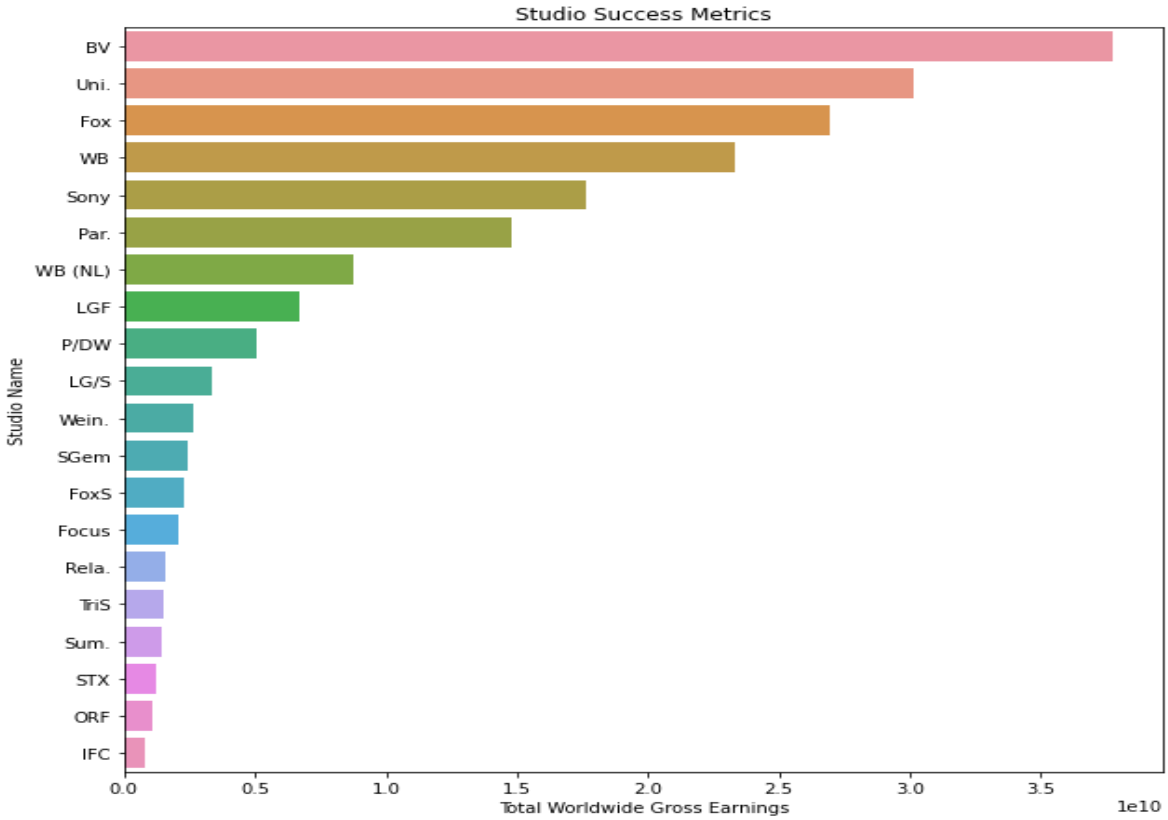
Problem Statement 1: To profile the competitive landscape within the movie industry, including the performance of key players

This inquiry aims to analyze the competitive dynamics of the film business. To comprehend their respective performances, it entails a detailed research of the studios. Microsoft can make strategic decisions about prospective alliances, collaborations, or areas of opportunity to build a solid footing in this competitive sector by acquiring knowledge into the competitive forces at play.

Using bivariate analysis we plot a graph to check the dominance of the studios based on the number of movies they have released.



Next we examine the top studios based on worldwide gross and identify the dominant players in the industry. This helps in understanding which studios have consistently delivered successful movies. Analyzing the distribution of worldwide gross among the top 20 studios can reveal how market share is distributed. You can identify if a few studios hold a significant portion of the market or if it's relatively evenly distributed.



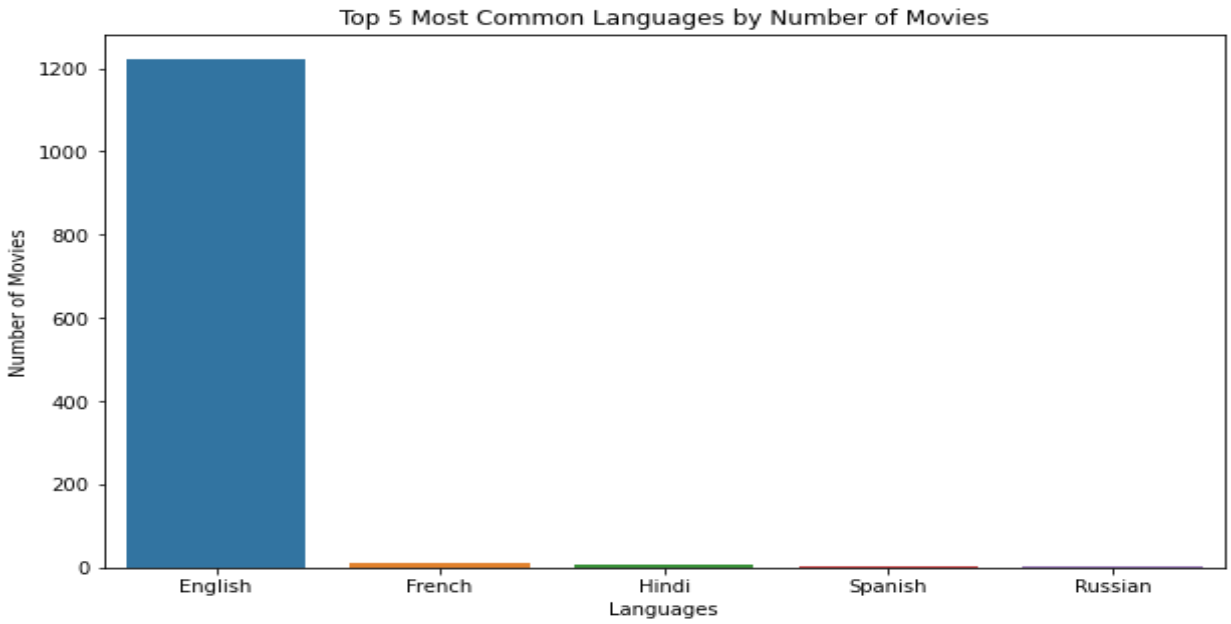
The above analysis arises the following insight:

- Competitive Benchmarking: A corporation can evaluate its position in the market by comparing its performance to that of its rivals through the strategic process known as competitive benchmarking. This analysis entails the identification of immediate competitors, the gathering and analysis of information, including financial metrics, and market share. Companies use this assessment to uncover best practices used by top competitors, identify strengths and weaknesses, create performance goals, execute changes, and continuously track progress. In the end, this iterative approach helps firms thrive in a competitive environment by fostering performance improvement, informed decision-making, market positioning and efficient strategy planning.To Microsoft, as a new company entering the industry, this analysis provides a benchmark to compare against. It helps in setting realistic goals for market share and revenue targets.

Problem Statement 2: To assess the success of different languages based on their worldwide gross earnings

This analysis tries to reveal the importance of languages in the current box office scene. We want to determine which languages now control the global by comparing the global gross revenues of films released in a variety of languages. By taking into account audience preferences and enhancing global appeal, Microsoft will be able to make educated decisions about the preferred languages for its film productions.

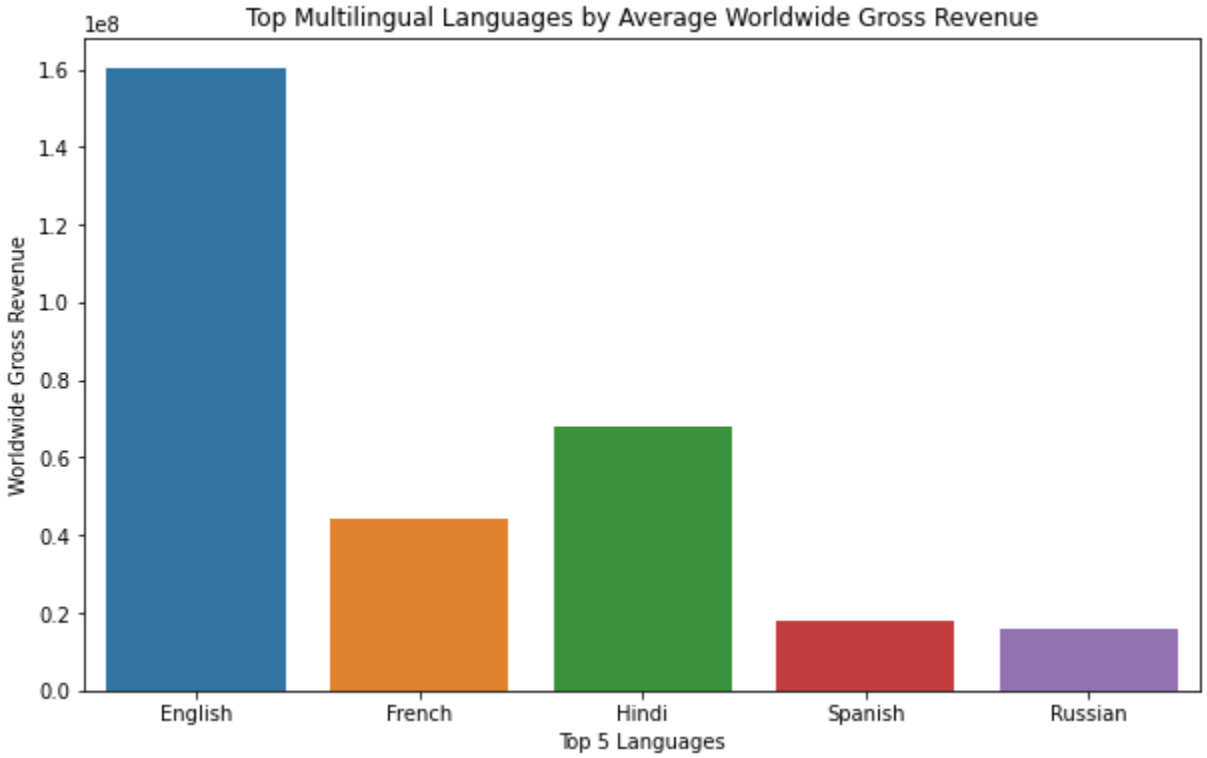
We plot a bar graph to check the top 5 most common languages used based on the number of movies made in that specific language.



The top 5 most common languages in order of most used are:

- English
- French
- Hindi
- Spanish
- Russian

Next we evaluate the worldwide_gross associated by each of the top 5 most common languages and analayze which one produces high gross values.



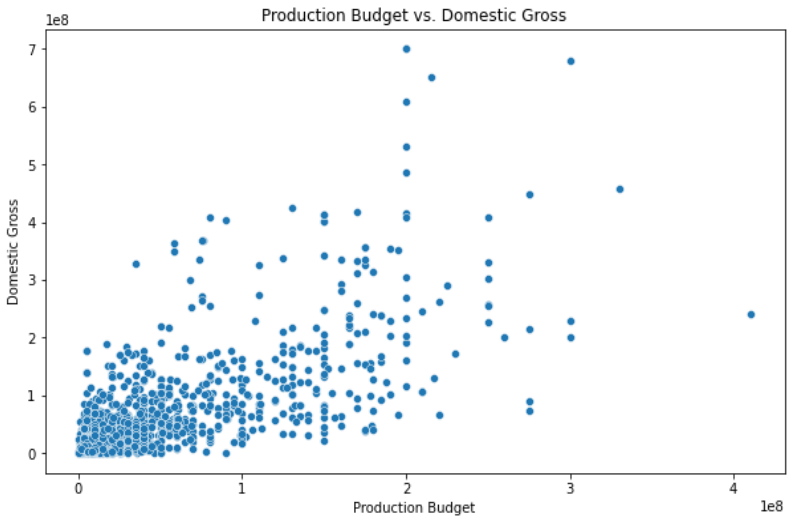
When compared to films in other languages, the study of the movie dataset shows that English-language films regularly have the greatest worldwide box office take. Several things can be used to support this conclusion. First off, as English is one of the most extensively used languages in the world, a larger audience for English-language films is possible. In addition, many English-language films are produced by renowned studios, which can increase their marketability and appeal to a broad audience. Therefore, Microsoft should consider producing a significant portion of its movie portfolio in English. This aligns with the global appeal of English-language cinema and its potential to reach a broader audience.

Problem Statement 3: To assess the financial implications of production budgets on the performance of movies

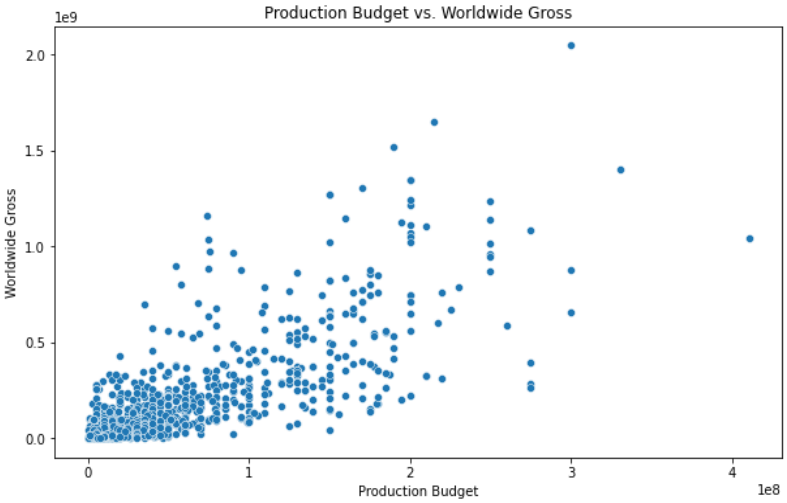
This analysis explores the financial aspects of the film industry, focusing on the connection between production costs and total film quality. by looking at how spending on a film's budget affects its commercial success. These data will enable Microsoft's film division to allocate resources intelligently, assuring the best possible financial results while creating high-caliber cinematic experiences.

The aim is to plot scatter plots which are crucial for data-driven decision-making in the movie industry. They provide insights into how financial investments in production budgets correlate with revenue generation and how domestic and global markets interact. This information informs budgeting, marketing, and distribution strategies, ultimately contributing to the financial success of movie projects.

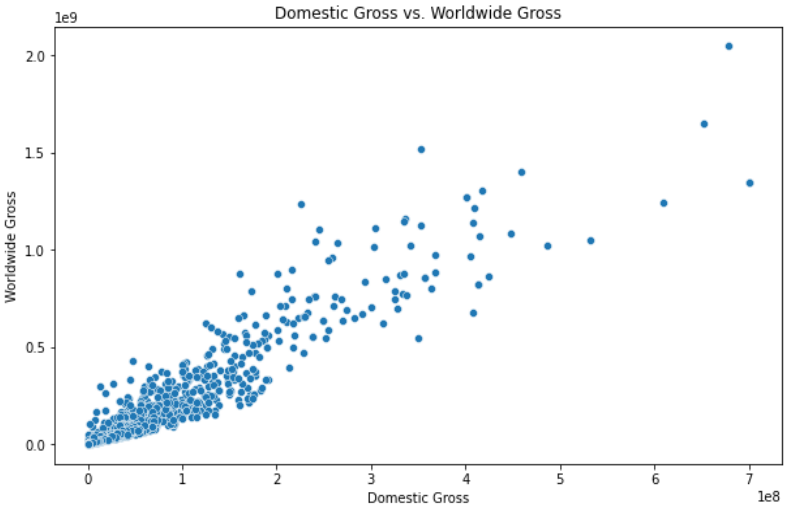
- Domestic Gross and Production Budget: There exists a moderately strong positive relationship between the production_budget invested in a movie and its domestic_gross_y earnings. This finding suggests that allocating higher budgets to movie production is associated with a greater likelihood of achieving higher earnings in the domestic market – 0.6985841005507146



- Worldwide Gross and Production Budget: Similarly, the positive correlation between worldwide_gross and production_budget indicates that higher production budgets are linked to increased worldwide gross earnings. This emphasizes the financial impact of strategic budget allocation in movie production – 0.7791527956492643



- Domestic Gross and Worldwide Gross: The notably high correlation between `domestic_gross_y` and `worldwide_gross` highlights that movies performing well domestically tend to perform exceptionally well on the global stage. This underlines the interconnected nature of these markets and the potential to leverage domestic success for international appeal – 0.9404679571013228



These strong positive connections highlight the financial significance of production budget allocation in the film business, to sum up. In light of the fact that wise production investments can have a major impact on both domestic and international profitability, Microsoft's new film studio should take these findings into account when deciding how much money to allocate to filmmaking. Recognizing the interdependence of the domestic and international markets further highlights the importance of flexible distribution and marketing tactics for maximizing profit potential.

CONCLUSION

The research of the movie dataset has revealed some significant trends that can direct Microsoft's new movie studio in its foray into the film industry:

- Language Choice: English-language movies tend to achieve the highest worldwide gross, highlighting the global appeal of English cinema. Microsoft should consider producing a significant portion of its movie portfolio in English to tap into this broad audience.
- Budget Allocation: The production budget and both domestic and global gross have a strong positive link. This suggests that strategic budget allocation can have a big impact on generating revenue. To maximize revenue, Microsoft should carefully manage and spend its budgets.he remarkably strong correlation between domestic and international box office revenue shows that films that do well domestically also frequently do well internationally. Microsoft may use its local success to broaden its appeal internationally.
- Collaboration and Competition: It is crucial to comprehend the competitive environment. Microsoft ought to be aware of the market share and effective strategies of other significant studios. To increase its influence and reach, chances for partnership with other studios or distribution channels can also be investigated.

It's important to recognize that this study has significant limitations, though. Because the film industry is subject to changing trends and audience preferences, the dataset used only represents historical data. In addition, elements including marketing tactics, timing of the release, and critical acclaim play important roles in a movie's success and weren't completely examined in this analysis.

Next Steps

Microsoft's film division might take the following into account to enhance this project in the future:

- Conducting market research: To identify current consumer preferences and trends can help writers decide on scripts and genres.
- Comprehensive Data: Adding more recent data and a larger range of factors to the dataset, such as marketing spending, genre-specific insights, and important reception measures.
- Competitive investigation: A thorough investigation of the competition to compare studios with success and find best practices. By using predictive modeling and machine learning, movie performance may be predicted based on a variety of variables, allowing for data-driven decision-making.