

# Analyzing Box Office Trends for Microsoft's New Movie Studio

## Overview

### Background of the Business

Microsoft, a technology giant renowned for its software and hardware innovations, has recognized the growing success of big companies venturing into original video content production. With a desire to explore new horizons and tap into the thriving world of entertainment, Microsoft has made a strategic decision to establish a new movie studio.

### Domain of the Business

Microsoft's latest venture positions it squarely within the entertainment industry, with a specific focus on movie production. This strategic move underscores Microsoft's recognition of the immense potential within the film sector, which includes every facet of filmmaking, from scriptwriting to distribution. The objective is clear: to produce compelling and commercially successful films that resonate with diverse audiences, both creatively and financially.

This domain is characterized by its artistic and technical aspects, audience engagement, revenue generation through various channels, competition with established studios, and opportunities for collaboration with creative talents. Additionally, Microsoft's expertise in technology introduces the potential for innovative approaches to filmmaking, from visual effects to immersive experiences. In summary, Microsoft's foray into movie production blends artistry and technology in a dynamic and multifaceted space within the entertainment industry.

### Business Case

The business case for Microsoft's new movie studio is grounded in the evolving landscape of entertainment consumption. With the rise of streaming platforms and the increasing demand for original content, Microsoft aims to capitalize on this trend by creating its films. However, needing experience in the movie production business, Microsoft faces a significant challenge: determining what types of films are currently performing the best at the box office. To address this challenge, Microsoft has tasked our team with conducting a thorough analysis of the movie industry to uncover actionable insights. These insights will guide the head of Microsoft's new movie studio in making informed decisions about the types of films to create, thus ensuring a successful entry into the market.

# Business Understanding

## Details About the Business to Venture Into

Microsoft's new movie studio is venturing into the competitive movie production industry with the goal of creating films that connect with audiences and generate substantial box office revenue. This strategic expansion into entertainment recognizes the industry's potential in the digital age. Microsoft aims to compete with established studios, focusing on audience engagement and revenue generation. Challenges include intense competition and meeting diverse audience preferences. Microsoft seeks to produce high-quality, innovative content using its technological expertise. Success in this venture will depend on producing movies that resonate with viewers, both creatively and financially.

## Objectives

### **1. To Analyze Budget vs. Revenue Relationship:**

- Investigate how production budgets influence the box office success of movies.

### **2. To Explore Genre-Based Revenue Performance:**

- Examine how different movie genres correlate with box office revenue.

### **3. To Analyze the Influence of Release Time on Revenue:**

- Explore how the timing of a movie's release (e.g., seasonal trends, and holiday releases) impacts its box office revenue.

### **4. To Provide Actionable Insights for Movie Production:**

- Translate the budget vs. revenue analysis findings into actionable recommendations for Microsoft's movie production decisions.

# Data Understanding

## Understand the Columns of the Data

The data used for this project comprises information related to movies and their box office performance. The data source for this project could be a combination of publicly available datasets from movie databases which are The Movie Database(TMDB) and The Numbers(tn) budget data. These datasets may have been collected over time and can include information on a wide range of movies.

This data encompasses attributes such as production budgets, genres, release dates, and worldwide box office earnings. The primary target variable is "worldwide\_gross," which signifies box office revenue. Key properties of these variables include numerical budgets, categorical genres, date-based release dates, and continuous revenue values.

The columns of interest are as follows:

- Production Budget: The cost of producing a movie.
- Worldwide Gross Revenue: The total revenue generated by a movie globally.
- Genre: The category or genre of each movie.
- Release Date: The date on which a movie was released.

## Check the Relation of the Different Datasets

The datasets are related by common attributes such as movie titles, genres, and release dates, allowing for a comprehensive analysis of box office performance. The main relation is the individual movie names since the names are the same across the different datasets. This is what I used to merge the different datasets.

## Data Cleaning

Data cleaning procedures will involve handling missing values, outliers, and data inconsistencies to ensure the quality and accuracy of the analysis.

This data contained a lot of duplicate values in the movie names. The outliers in the case of revenue were not a main issue for this study since the values showed how varied the revenues of different films were.

## Data Analysis

### Exploratory Data Analysis (EDA)

#### Univariate EDA

In the univariate analysis, we will explore individual variables. This involves calculating summary statistics and creating visualizations, such as histograms and distribution plots, for key variables like production budgets and revenue figures.

Here I was able to create histograms to show the distribution of the worldwide gross which is in essence the revenue.

#### Bivariate EDA

Bivariate analysis will explore relationships between variables, including how production budgets relate to box office revenue. Bivariate analysis examines the relationships between pairs of variables. We will investigate how production budgets relate to box office revenue through scatterplots and correlation analysis. This helps us identify potential associations between these critical factors.

# Plots

The plots that are used in this study are like scatter plots to show a correlation between different factors, bar plots, histograms, and even box plots

Once the data was cleaned I investigated the relationship between production budget and revenue in this case profit, worldwide gross, domestic gross, and foreign gross. This showed that all of them had a strong positive correlation but of differing degrees.

## 1. Correlation between Production Budget and Domestic Gross (0.73):

A correlation coefficient of 0.73 indicates a strong positive linear relationship between the production budget and domestic gross.

This means that as the production budget of a movie increases, the domestic gross tends to increase as well. In other words, movies with higher budgets tend to earn more at the domestic box office.

## 2. Correlation between Production Budget and Domestic Gross (0.73):

A correlation coefficient of 0.80 indicates a strong positive linear relationship between the production budget and worldwide gross.

This means that as the production budget of a movie increases, the worldwide gross tends to increase as well. Movies with higher budgets tend to have higher worldwide box office earnings.

## 3. Correlation between Production Budget and Profit (0.69):

A correlation coefficient of 0.69 indicates a strong positive linear relationship between the production budget and profit.

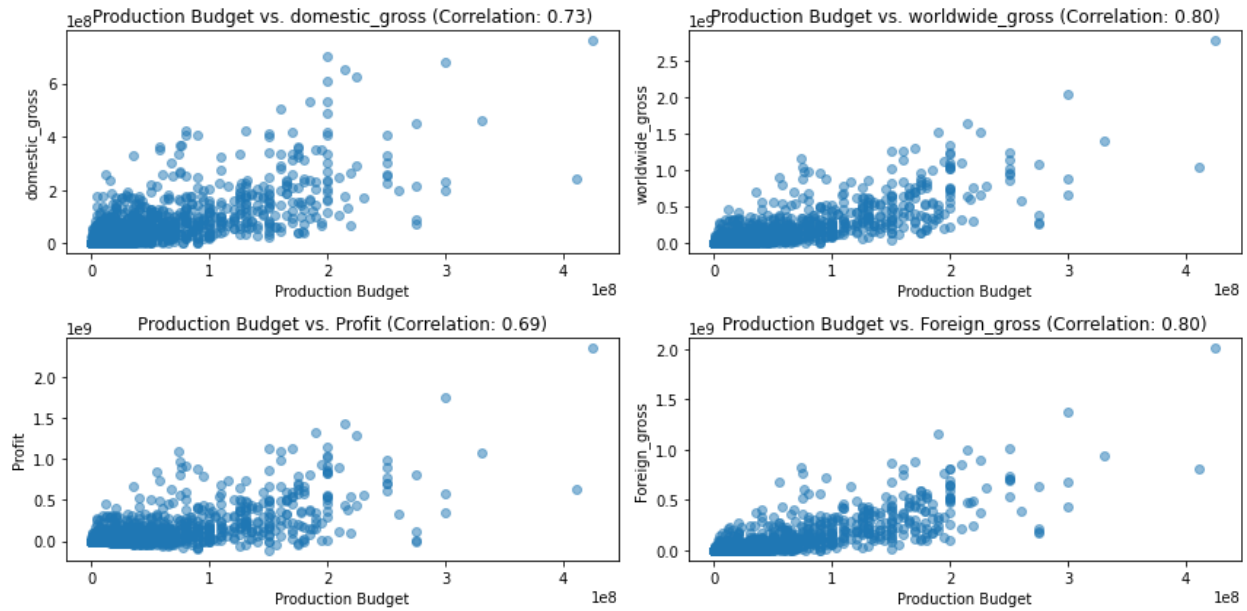
This means that as the production budget of a movie increases, the profit tends to increase as well. Movies with higher budgets tend to generate more profit.

## 4. Correlation between Production Budget and Foreign Gross (0.80):

A correlation coefficient of 0.80 indicates a strong positive linear relationship between the production budget and foreign gross.

This means that as the production budget of a movie increases, the foreign gross tends to increase as well. Movies with higher budgets tend to have higher worldwide box office earnings.

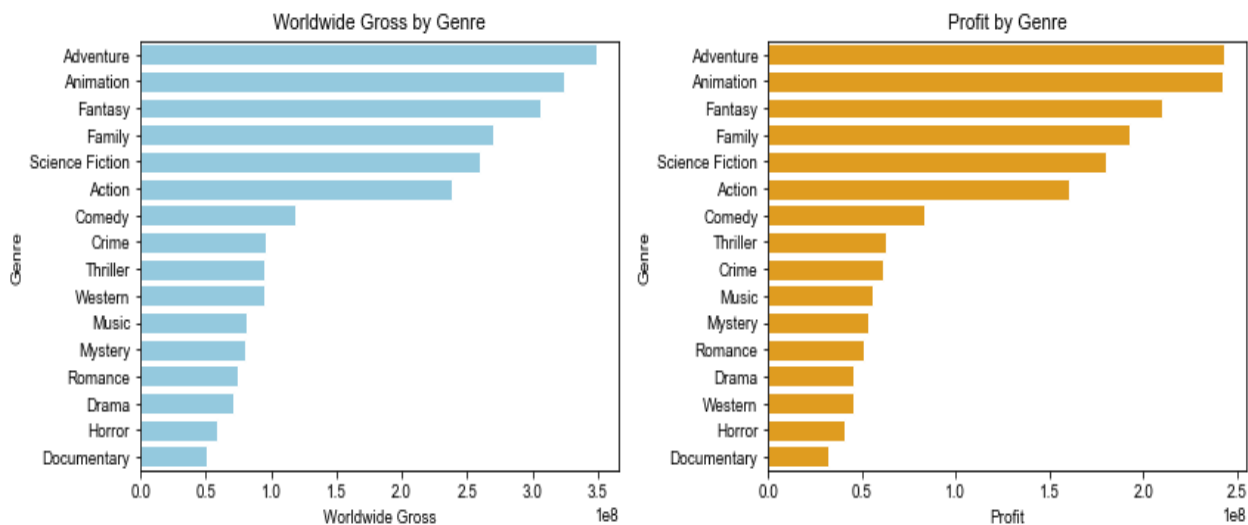
The plot explaining the above explanations is shown below



As for the genre-based revenue performance, I was able to create a horizontal bar graph to show how each genre was related to the total revenue(worldwide gross) and the profit gained by the different movies.

What is of note the genres which had the largest revenues are adventure, animation, fantasy, and family.

What is of note is that the genre with the most movies surprisingly was drama followed by comedy. This shows that many movies produced have this as a sub-genre in the movie genre list.



The interpretation of the boxplot below is as follows.

Best Months for Releases (December, May, April, November):

- These months have been identified as optimal for movie releases based on their higher median revenue. This means that, on average, movies released during these months tend to perform well at the box office.
- Despite the presence of outliers (movies with exceptionally high revenue), these months consistently demonstrate a higher central tendency in revenue, indicating a more reliable trend.
- December is noteworthy for its association with holiday-themed movies, which tend to attract a large audience during the festive season, leading to higher revenue.
- 

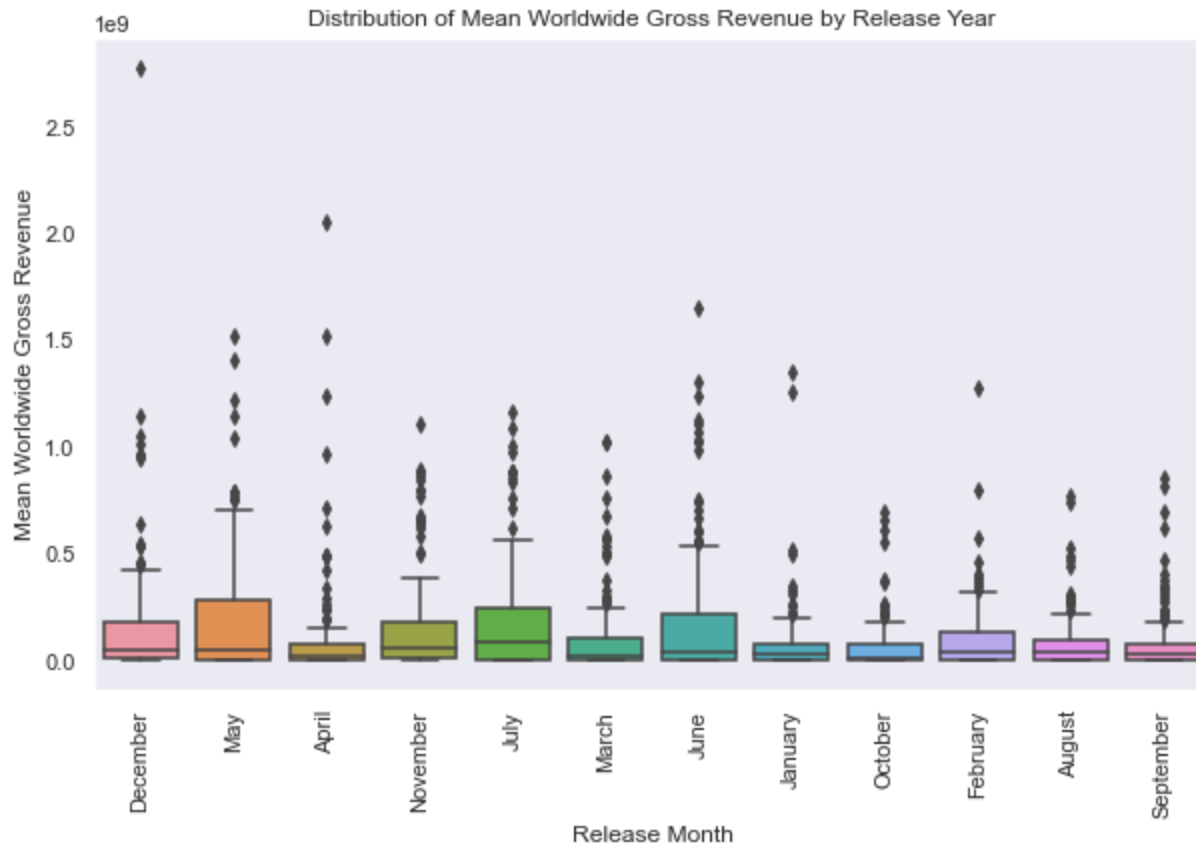
Greater Variability (May, July, June):

- Months like May, July, and June exhibit taller boxes in the box plots. Taller boxes represent a greater interquartile range (IQR), indicating higher revenue variability among movies released during these months.
- In these months, the box office performance of movies can be less predictable because the spread of revenue values (as indicated by the IQR) is wider. This suggests that some movies may perform exceptionally well, while others may not fare as well.
- Factors contributing to revenue variability during these months could include competition from other blockbuster releases, varying audience preferences, and the timing of major movie events or holidays.

Wider Revenue Range (May and Others):

- Months with longer whiskers, including May, have a wider range of revenue. Whiskers in a box plot represent the minimum and maximum values within a certain range from the median.
- A wider revenue range implies that box office performance varies more during these months. Movies released during these months can achieve both high and low revenue, resulting in a broader distribution of outcomes.
- Factors contributing to a wider revenue range may include the diversity of movie genres released during these months and the influence of external events or marketing campaigns.

As for seasonality from the plots it is shown that the summer is a good time to release movies but specific months are better than others. So more information on the genre and audience type is required so as to accurately use this information to maximize revenues.



## Recommendations

Based on the analysis findings, we offer actionable insights to guide Microsoft's new movie studio in its entry into the movie production industry:

**1. Strategic Investment:** The analysis reveals that higher production budgets can lead to better box office performance. However, it's crucial to allocate budgets strategically, considering genre and release timing. Microsoft should ensure efficient budget allocation to maximize returns.

**2. Genre-Centric Approach:** Genres play a significant role in revenue generation. Microsoft should focus on producing movies in genres that have consistently performed well at the box office. Stay adaptable to evolving audience preferences and emerging trends.

**3. Release Timing Strategy:** The timing of movie releases is pivotal. Microsoft should create a well-planned release calendar, capitalizing on months like December, May,

April, and November. Pay close attention to seasonal trends and avoid months with high revenue variability, such as May, June, and July.

**4. Outlier Investigation:** Keep a watchful eye on potential outliers, especially in December and other months. Investigate these exceptional performances, which may be influenced by genre-specific releases, and consider replicating successful strategies when appropriate.

**5. Continuous Market Research:** Stay committed to ongoing market research to stay informed about audience preferences, emerging genres, and evolving consumer behavior. This will enable informed decisions about genre selection and release timing.

**6. Collaboration and Expertise:** Collaborate with industry experts, consultants, and filmmakers to gain deeper insights into the film market. Leverage their expertise to identify opportunities and strengthen the production strategy.

**7. Advanced Analytics:** Consider implementing advanced statistical models or machine learning techniques to build predictive models for box office performance. This can enhance precision in budget allocation and genre selection.

By following these recommendations and maintaining a data-driven, adaptable approach, Microsoft can position itself for success in the dynamic and competitive movie production industry.

## Next Steps

In the future, we recommend:

- **Data Enhancement:** Focus on collecting additional data in areas where it is currently missing, such as comprehensive audience feedback and demographic data. This will provide a more comprehensive understanding of the audience.
- **Ongoing Market Monitoring:** Continuously monitor industry trends, audience preferences, and box office performance. Adapt strategies to changing dynamics and evolving tastes.