UNLOCKING SUCCESS IN THE WORLD OF ENTERTAINMENT: Microsoft's Movie Studio

Overview

Microsoft is a multinational technology company known for developing software, hardware and cloud-based services. Microsoft has been a trendsetter in technology for nearly half a century. We are renowned for our innovation, commitment to quality and global impact on the digital landscape. From operating systems to cloud computing, our expertise has touched every corner of the tech universe.

In an era where technology and entertainment intersect to define the boundaries of creativity and engagement, Microsoft is embarking on a remarkable journey. Our focus is to establish a great presence in the world of original video content creation. We leverage the power of data to derive insights that will provide a path forward for Microsoft's new movie studio.

According to a research study published by Custom Market Insights, the demand for the global entertainment industry market size and share revenue is valued at approximately $29.86 billion and is expected to reach $49.56 billion by 2032. The market is characterized by vibrant and evolving growth dynamics.

Business Understanding

Original video content creation has emerged as a cultural and economic force. It enables storytellers to explore diverse themes, genres and narratives. High-quality content can create a deep connection with the audience, leading to dedicated fan communities and long-term viewership. Successful content can generate revenue through box office earnings, streaming deals, merchandise distribution and licensing. With the potential to reach a global audience, original content can transcend cultural boundaries and gain international recognition. Exceptional content can win awards, thus enhancing the studio's reputation.

However, venturing into a new territory can have its challenges. The art of creating successful movies requires a diverse set of skills and insights. The entertainment industry is known for its high-risk nature. Many movies suffer financial losses when they do not regain their production costs. Developing and producing movies can be a long process which can tie up resources for extended periods. The field is also highly competitive.

We are conducting exploratory data analysis to uncover valuable trends and patterns in the film industry. Our goal is to inform strategic decisions and enhance our chances of box office success. The objectives of this analysis are:

1. To investigate how the production budget impacts a movie's performance at the global box office. This will cover a wide range of budgets allowing us to draw insights across budget ranges.
2. To investigate the relationship between seasonal trends and box office performance. We seek to understand whether certain release months or seasons influence a movie's worldwide revenue.
3. To determine whether the duration of a movie plays a role in its box office success. Examining successful movies with varying runtimes will help us establish whether there exists an ideal duration that maximizes worldwide revenue.

Data Understanding

The datasets were obtained from IMDB, Box Office Mojo, Rotten Tomatoes, TheMovieDB and The Numbers. Since they were collected from various locations, the different files had different formats. Some were compressed CSV (comma-separated values) or TSV (tab-separated values) files that could be opened using spreadsheet software or Pandas, while the data from IMDB was located in an SQLite database.

Data Preparation

The data used in our analysis was IMDB data which was in the form of an SQL database. The primary tables used for this investigation were 'movie_ratings' and 'movie_basics'. We used a series of SQL queries to extract the relevant information from the IMDB database. The movie_basics tables provided essential information about the movie ID, primary title, start year, runtime in minutes and genres. The movie_ratings table provided the movie ID, average audience rating and the number of votes. To create a dataset suitable for analysis, we merged the information from the two tables into one unified dataframe. The dataframe was then converted into a more accessible format- a CSV file.

In addition to the IMDB data, we incorporated information from the movie_budgets dataframe. This additional data was stored in a CSV file. It contained information on important attributes of the movies such as the release date, name, production budget, domestic gross earnings and worldwide gross earnings. We merged this dataframe with our existing dataframe to create the merged_movies dataframe which was our final dataframe.

To ensure data accuracy and consistency, we performed essential data cleaning and preprocessing tasks. These included handling the missing values, dropping the unnecessary rows and addressing the data types. This was vital in preparing our dataset for analysis. The transformation of these datasets into a well-structured dataframe equipped us with a solid foundation. With the dataset in hand, we can delve into the factors that affect box office performance and revenue generation for films.

Data Analysis and Visualization

Analysis of Production Budget vs. Worldwide Gross

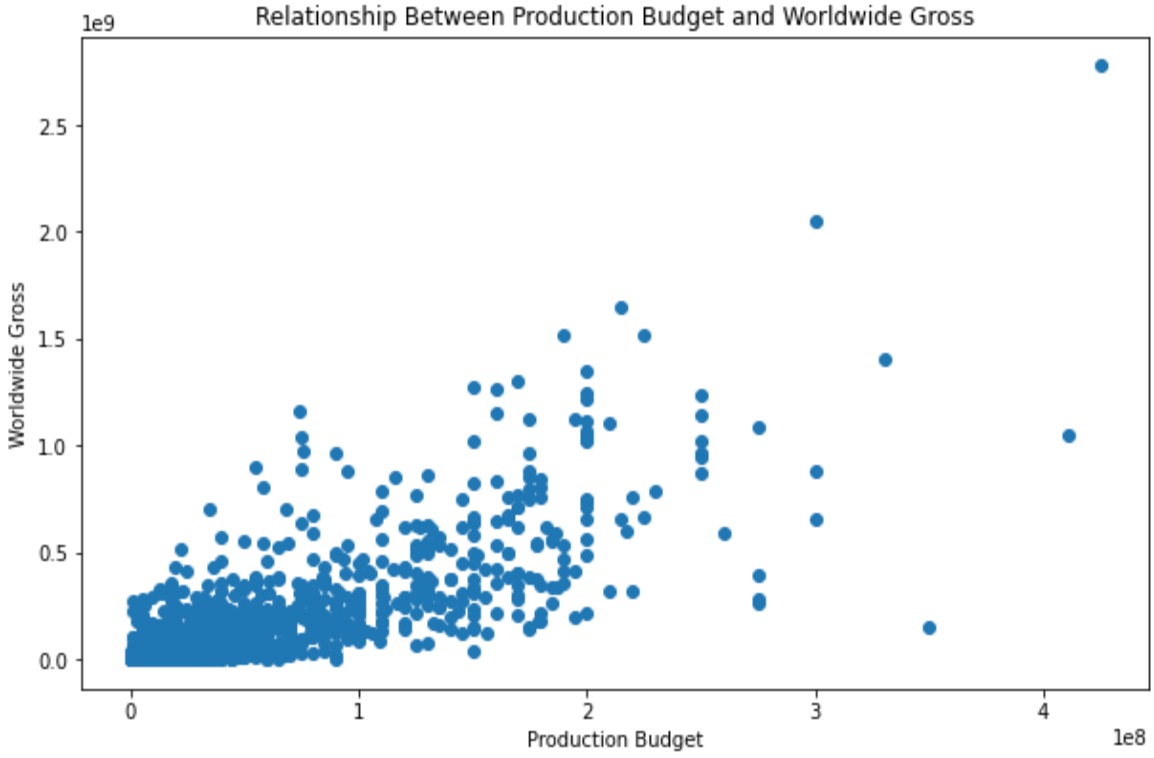
Worldwide gross is the total revenue generated by a movie from all its sources, including ticket sales and other distribution channels, on a global scale.
In order to understand the influence of the production budget on box office performance, we used a statistical measure known as the correlation coefficient and visualized our findings through a scatter plot.

- Correlation Coefficient

  The correlation coefficient quantifies the strength and direction of the linear relationship between two variables. It ranges from -1 to 1 where 1 indicates a perfect positive correlation, 0 suggests no correlation and -1 represents a perfect negative correlation.
  In our analysis, we obtained a correlation coefficient of approximately 0.79. This value suggests a strong positive correlation between the production budget and the worldwide gross. As the budget invested in a movie increases, its worldwide revenue also tends to increase.

- Scatter Plot Visualization

  To visualize this relationship, we used a scatter plot as shown below:



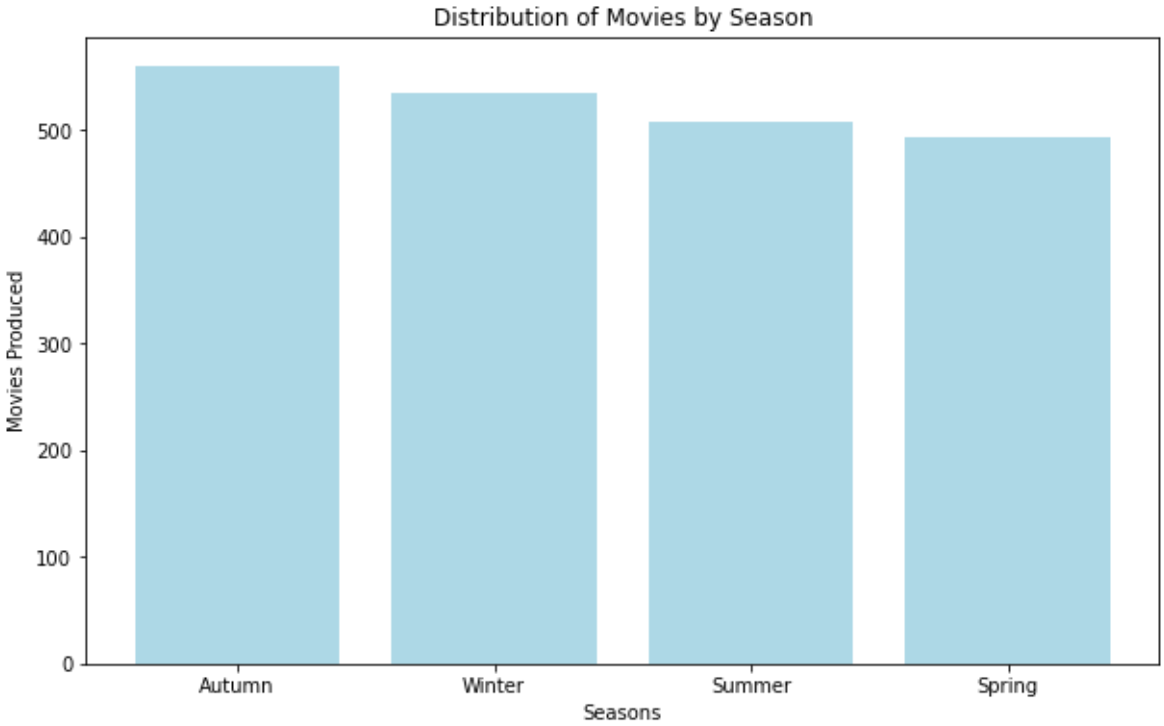
  From the scatter plot, we observe that there are a lot of movies with relatively low budgets and corresponding low worldwide gross as seen by the data points clustered around the origin. However, as we move along the x-axis, the data points gradually spread out towards the right, indicating a positive trend. We can conclude that movies with higher budgets tend to have higher worldwide earnings, as indicated by the scatter plot gradually moving towards the right and the high correlation value.

Analysis of Seasonal Trends in Box Office Performance

We explored the impact of seasonal trends on box office success using two steps. First, we extracted the release months and years from the release dates. We then categorized the months into the four seasons: Spring with the months March, April and May, Summer with the months June, July and August, Autumn with the months September, October and November, and Winter with the months December, January and February. We examined both the distribution of releases and their average worldwide gross across these seasons.

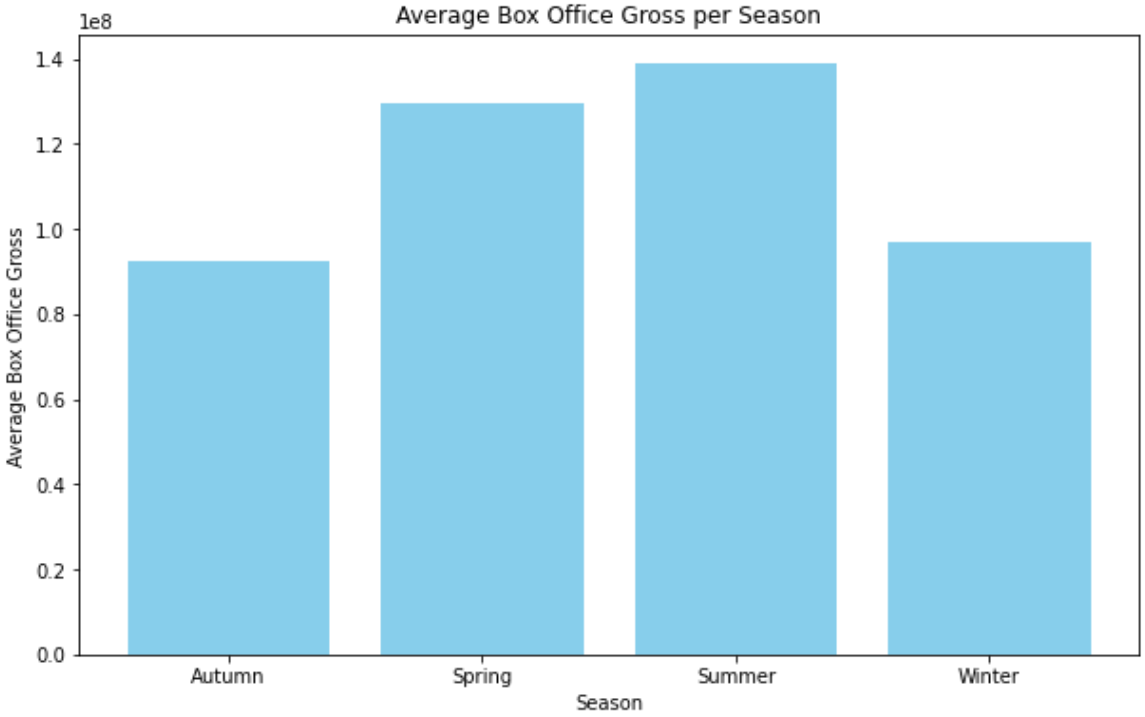
- Distribution of Movie Release by Season

  We visualized the distribution of movie releases across the four seasons using a bar plot. Our graph revealed that the highest number of movies were released during the Autumn season. This sheds light on the industry's preference for Autumn as a period to launch movies. The graph is given below:



- Average Box Office Gross by Season

  We calculated the average box office worldwide gross for movies within each season and displayed our results using a bar plot. Our analysis revealed that movies released during the summer season had the highest average worldwide gross. Despite a high number of releases during Autumn, movies in this season had the lowest average box office gross of all four seasons.



We can conclude that while Autumn may witness and increase in the number of movie releases, the summer season is best in terms of average box office

performance. It is important to strategically time the movie releases to align with periods of high viewer engagement and spending habits.

Analysis of Movie Duration and Box Office Earnings

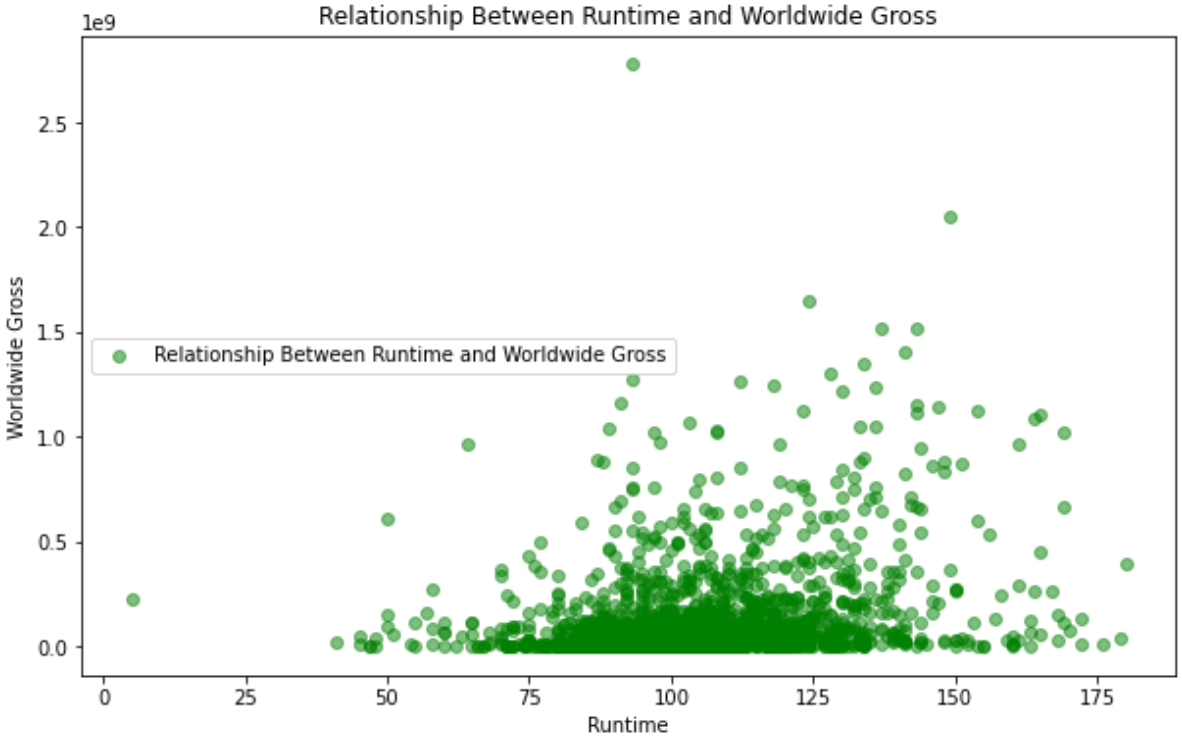
We employed various analytical methods to explore the connection between the movie duration and box office earnings. These were calculating the correlation coefficient, visualizing scatter plots and categorizing the runtimes into small, medium and long films based on the runtime of the lowest and the highest movie duration.

- Correlation Coefficient

  We first started by calculating the correlation coefficient between the runtime of the movie and the worldwide earnings. We obtained a correlation coefficient of approximately 0.28, suggesting that there is a weak positive relation between the two variables. While there exists a positive trend, it is not strong, which indicates that other factors may have a greater impact on box office success.

- Scatter Plot Visualization

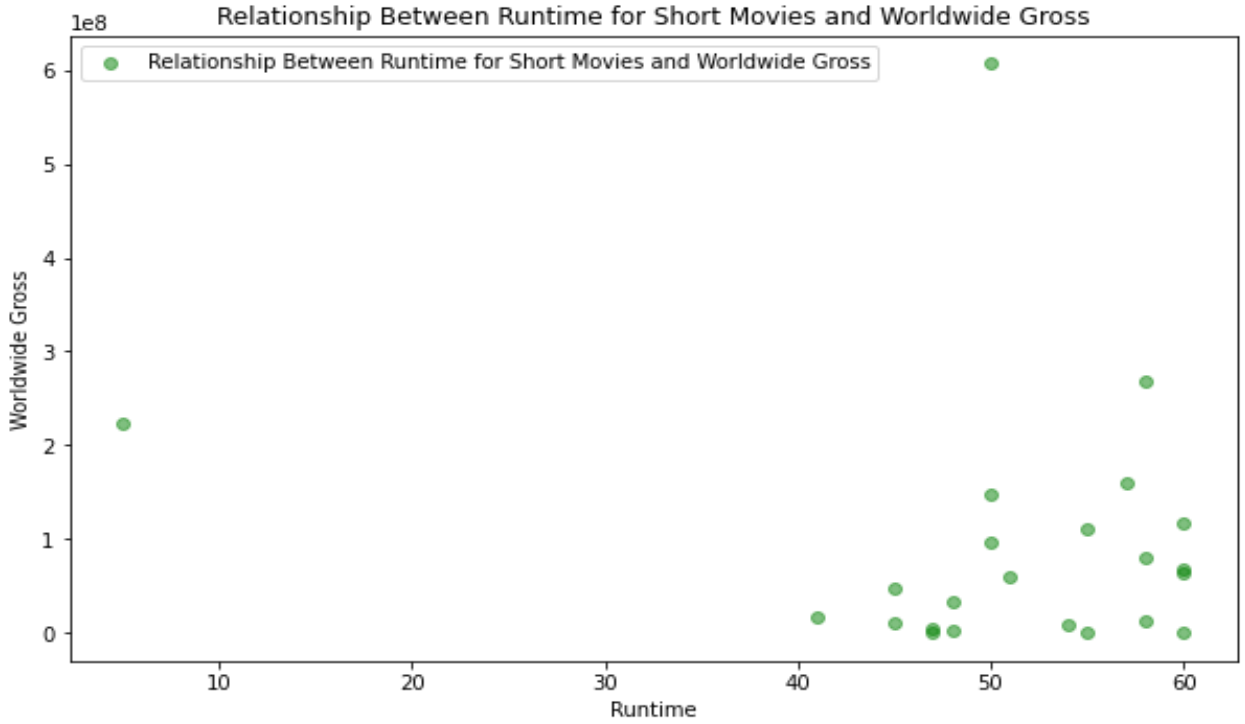
  We created a scatter plot to provide a visual representation of this relationship. We noticed that there was a concentration of data points around the middle range of runtime values, with some extending upwards along the y-axis. This implies that while there is a general distribution of movie durations, movies with moderate runtimes are more common and some of them achieve higher box office earnings. The plot is as shown below:

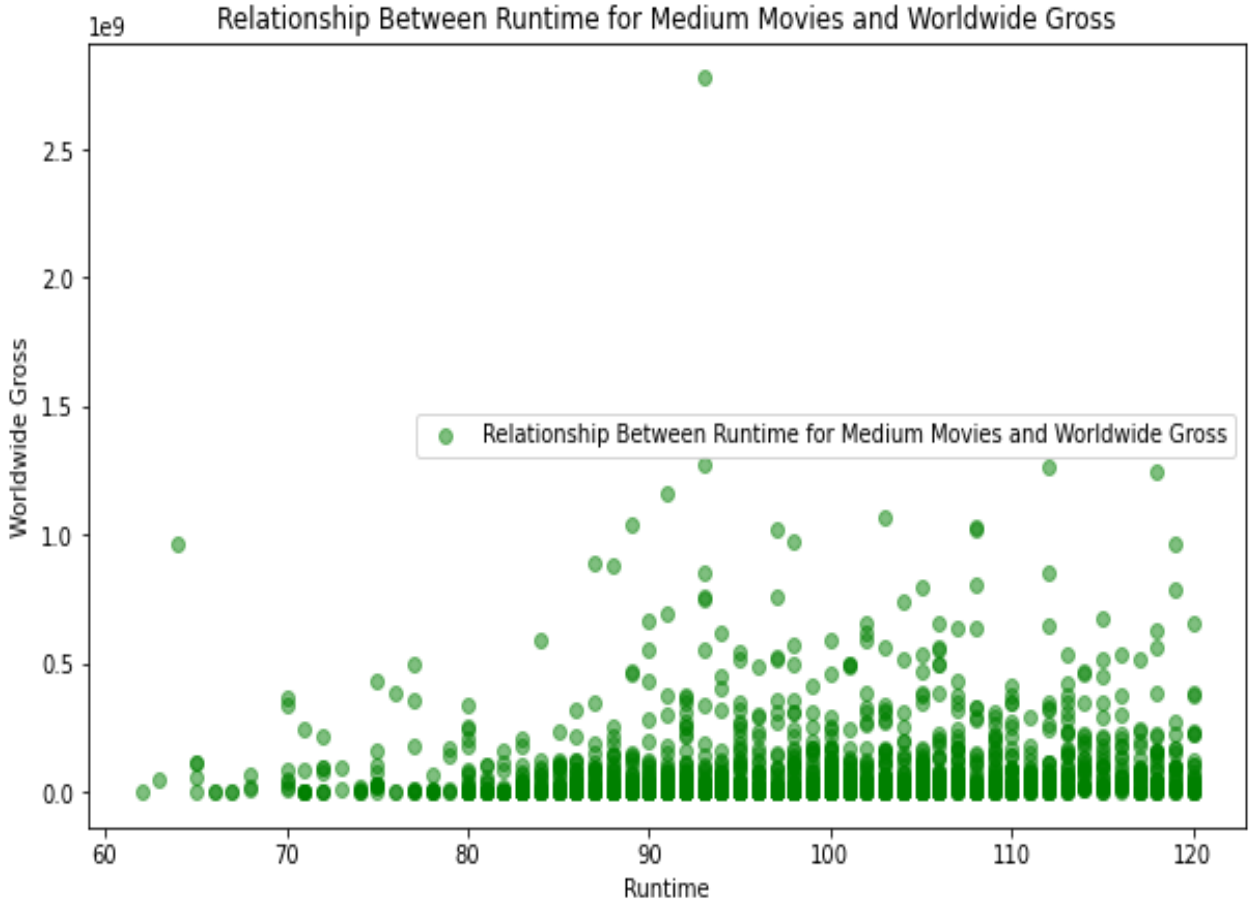


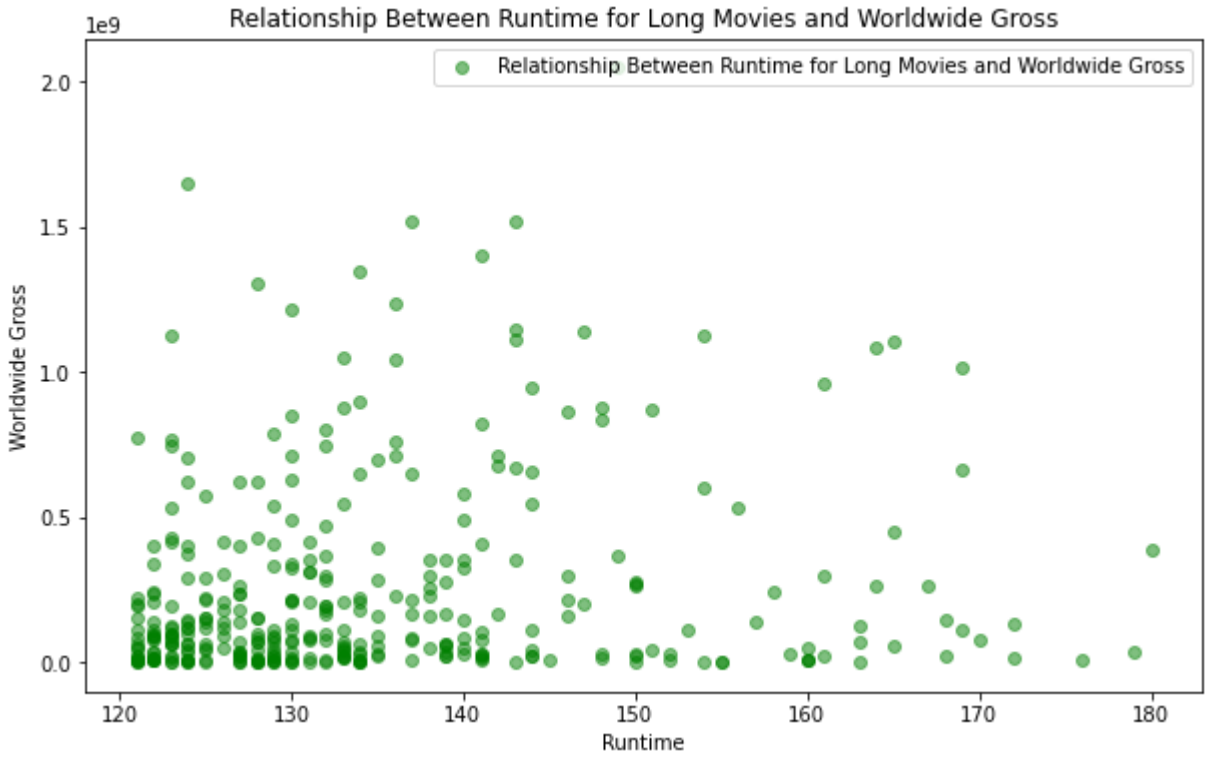
- Categorizing Runtimes

  We grouped the runtimes into three groups in order to gain further insights. There were small films which had a duration of 5-60 minutes, medium films with a duration of 61-120 minutes and long films which had a duration of 120-180 minutes. We then created scatter plots for each of these groups of films. Our subsequent scatter plots showed a concentration of movies with moderate

runtimes at the middle with some having high box office earnings. This resonates with our initial findings. The scatter plots are shown below:







Conclusions

It is important to convey the insights gained through our exploration of the determinants of a movie's box office success. We will revisit the key findings and translate them into actionable recommendations. We will also outline the next steps that Microsoft's new movie studio can take to generate additional insights. These items will pave way for the studio's entry into the competitive world of original content creation.

Recommendations

This analysis led to three recommendations for Microsoft to consider when getting into the movie business:

- **Allocate sufficient resources to production budgets.** Investing in production is important in order to maximize returns from the films. Producing high-quality content plays a significant role in the success of the movies.
- **Consider seasonal strategy.** Given that movies released during Summer tend to have a higher average worldwide gross, consider allocating more resources towards the summer season when the audiences are more active and engaged.
- **Diversity in runtimes.** Consider producing a mix of short, medium and long films to cater to various target audiences. This allows for broader audience appeal and accommodates different viewer preferences.

Next Steps

Further analyses that could yield additional insights in order to make an informed decision are listed below:

- **Analyze how genre can affect the performance of the movies.** Investigate whether there are differences in reception across the genres and how this affects box office performance. Are there emerging genres gaining popularity?
- **Predicting audience engagement.** This model could analyze social media trends, audience sentiment and audience discussions to predict the level of audience engagement and interest in a movie before its release.
- **Market research.** Identify trends, emerging opportunities and audience demographics to understand the state of the movie industry. Analyze the competition to find key players and gaps that the business can fill.