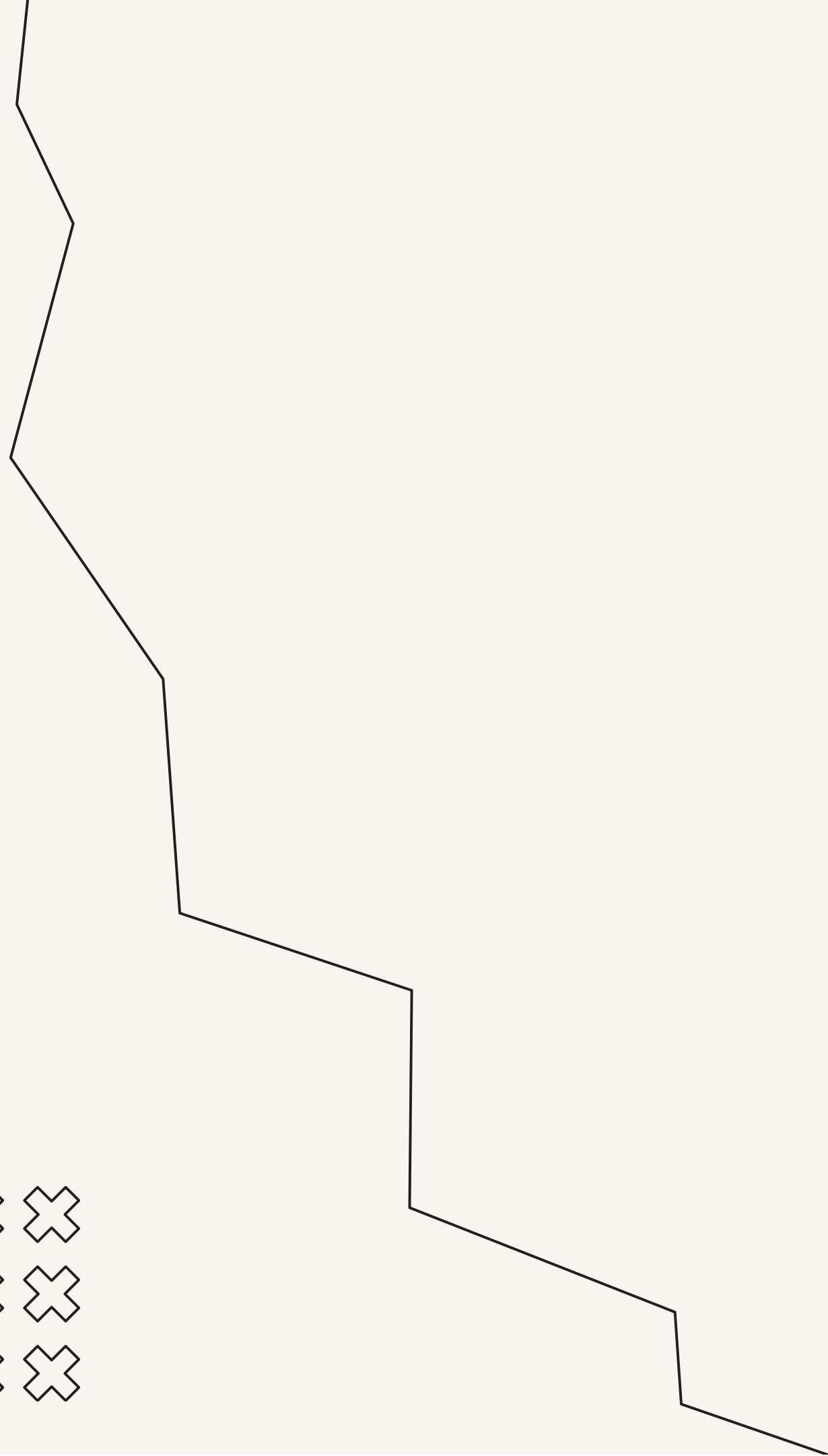
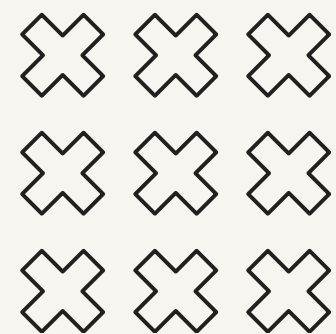




PRESENTATION BY:  
CELINE CHEGE



# Overview and Business Understanding



## Challenges :

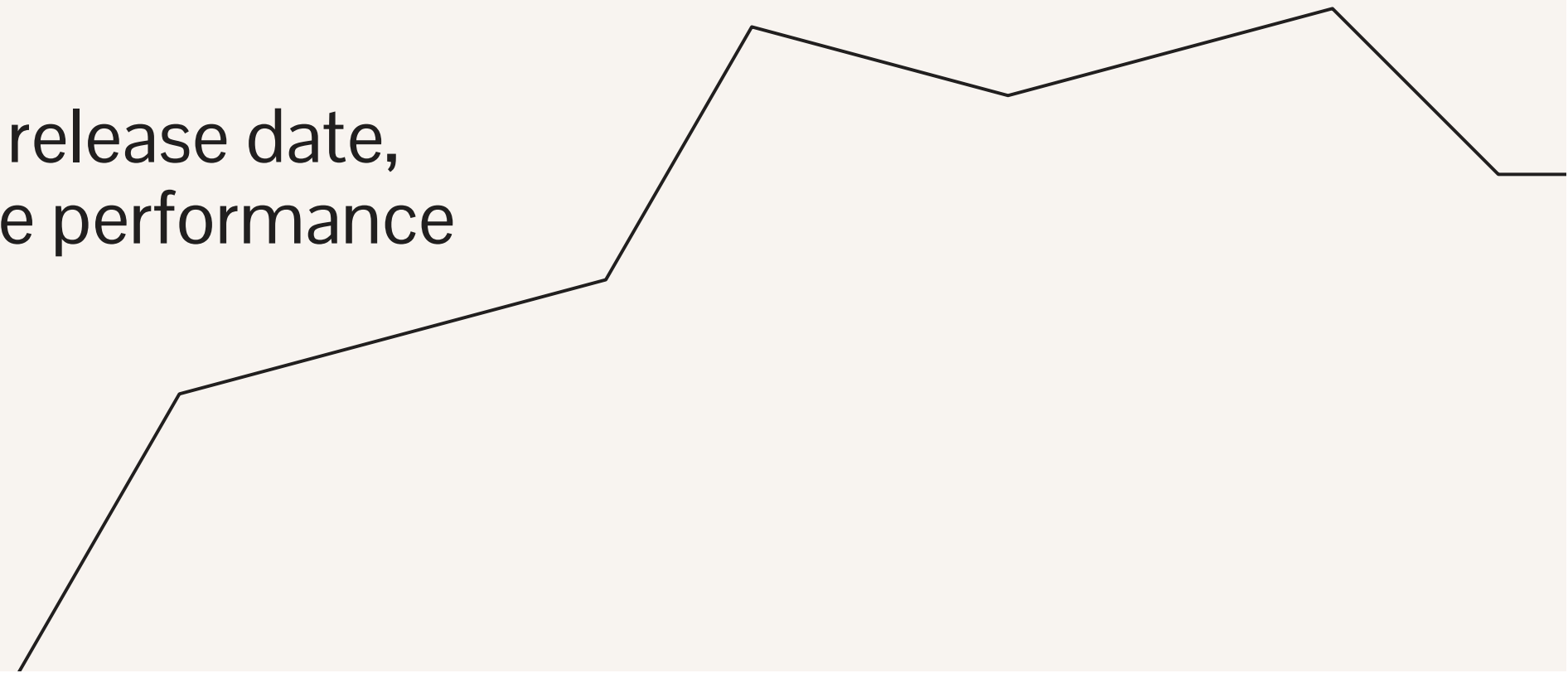
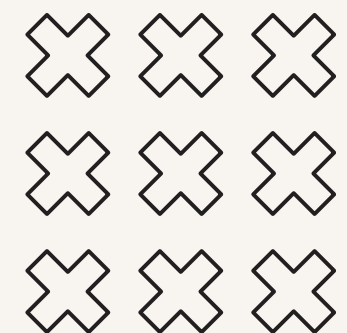
Lacking industry expertise, Microsoft aims to discern trends and patterns in successful films to guide production decisions.

## Business Objective :

Microsoft seeks to enter the movie industry with its own studio, leveraging data-driven insights for success.

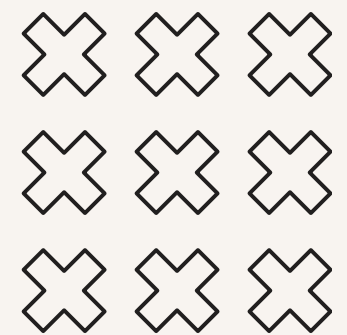
## Key Focus Areas :

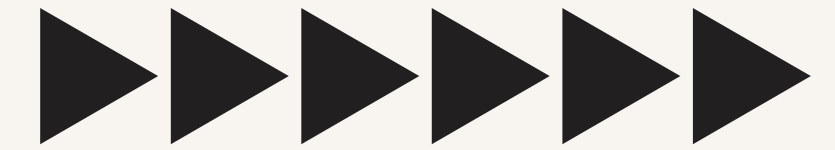
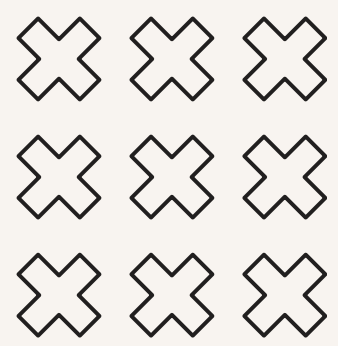
Genre, budget, and rating analysis Impact of release date, runtime, and production budget on box office performance



# Data Set 1

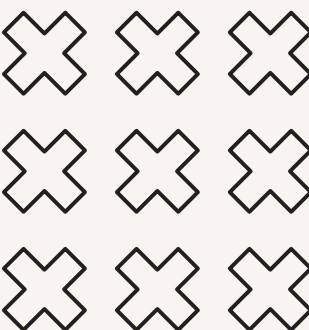
(Box Office)

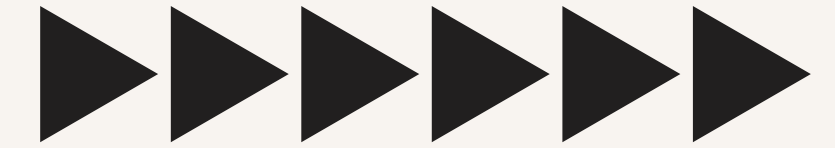
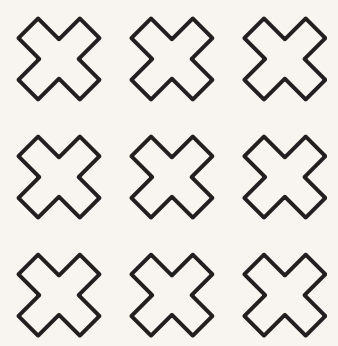




# Data Understanding

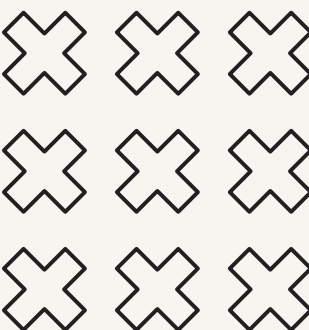
- Loaded the bom\_movie\_gross dataset .
- Contains movie titles, studios, domestic and foreign gross earnings, and release years.
- Summary statistics:
  - Mean domestic gross: \$28,745,850.
  - Mean release year: 2013.96.





# Data Cleaning

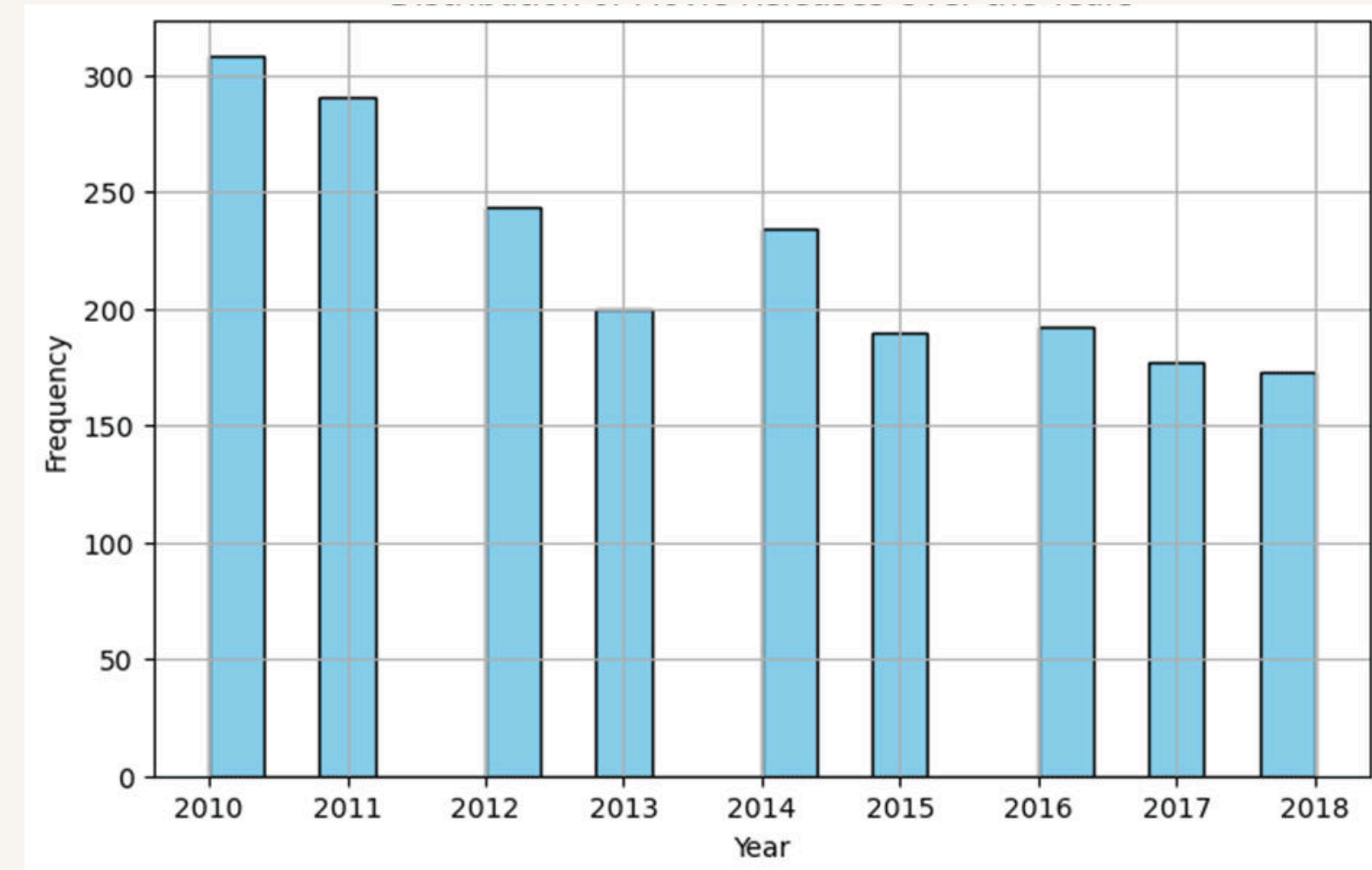
- Identified 3386 unique movie titles and 257 unique studios.
- Encountered missing values:
  - 5 for studio.
  - 28 for domestic gross.
  - 1350 for foreign gross.
- After removing null values, cleaned dataset comprises 2007 rows.





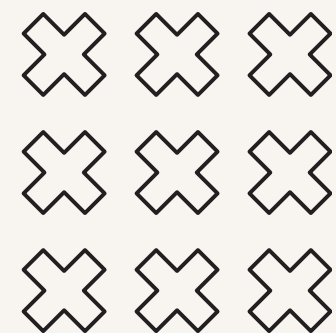
# Data Analysis and Visualization

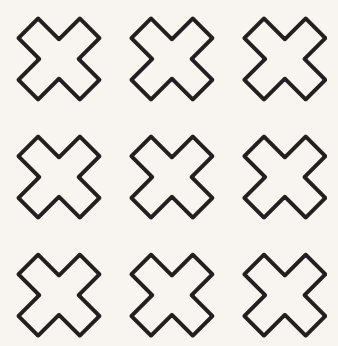
- Conducted analysis on distribution of movie releases over the years.
- Utilized histogram visualization for insights.
- Histogram showcases variation in annual movie releases.
- Provides understanding of industry trends and patterns.



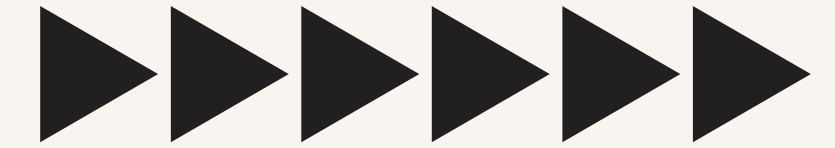
# Data Set 2

(Movie Budget)

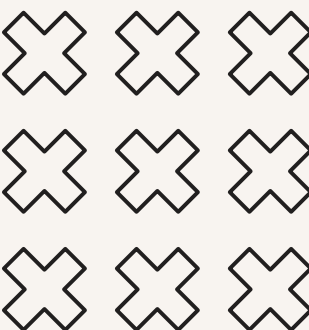




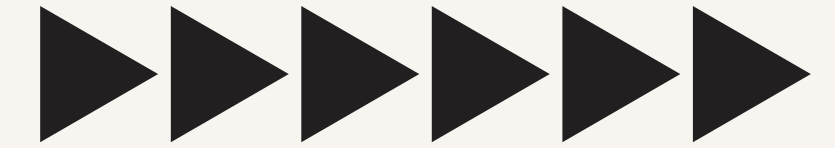
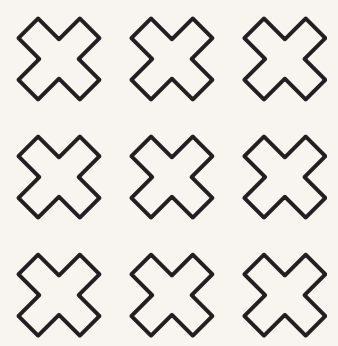
# Data Understanding



- Loaded the movie budgets dataset, encompassing details on production budgets, domestic and worldwide gross earnings, and release dates.
- Conducted data exploration to understand the structure and characteristics of the dataset, revealing 5782 entries and 6 columns.
- Utilized summary statistics to gain insights into the overall financial landscape of the movie industry, highlighting mean production budget, domestic gross, and worldwide gross.

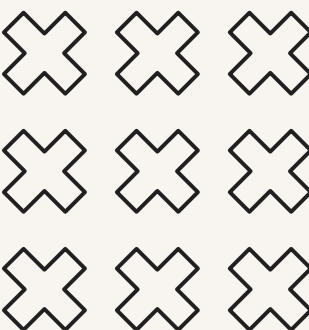






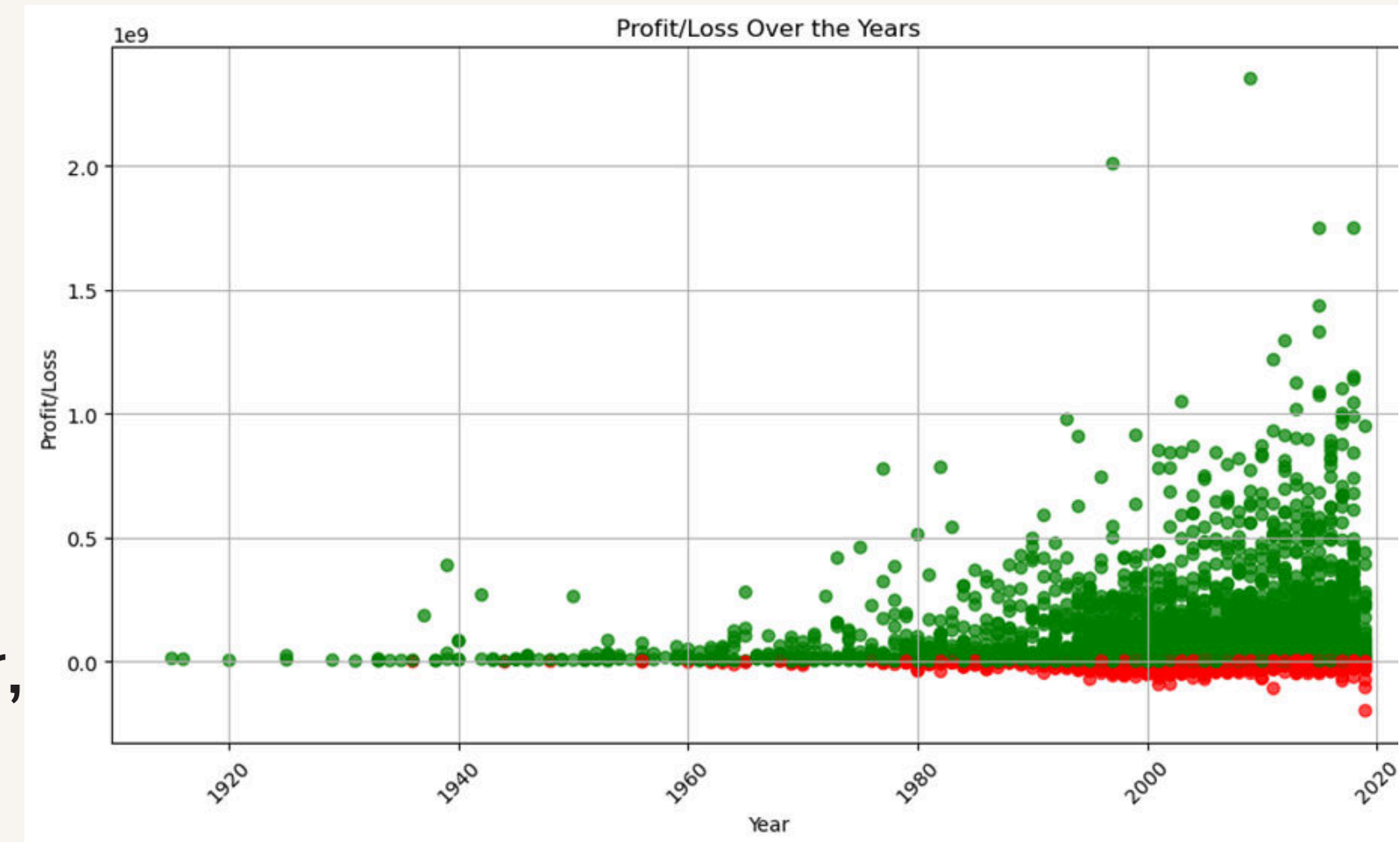
# Data Cleaning

- Identified and addressed missing values, ensuring data integrity.
- Converted string data types to integers for production budget, domestic gross, and worldwide gross columns.
- Verified the conversion and confirmed the absence of missing values.



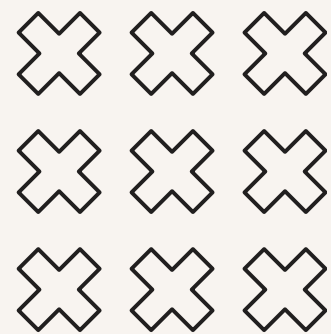
# Data Analysis and Visualization

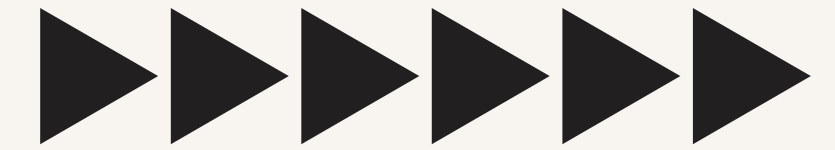
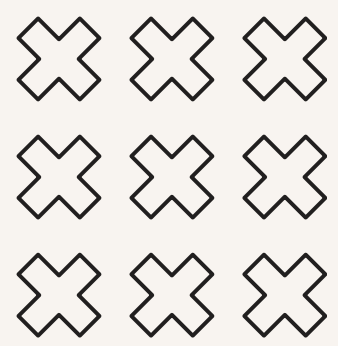
- Analyzed summary statistics to gain insights into the financial landscape of the movie industry, complemented by visualizations.
- Explored the distribution of movies over the years to understand trends and patterns visually.
- Examined the frequency of movies released each year, identifying potential areas for further analysis.



# Data Set 3

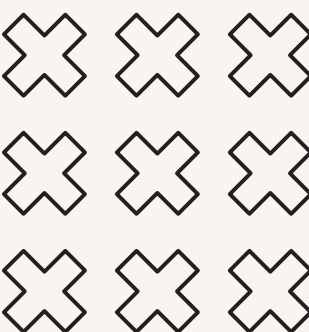
(Movie Reviews)

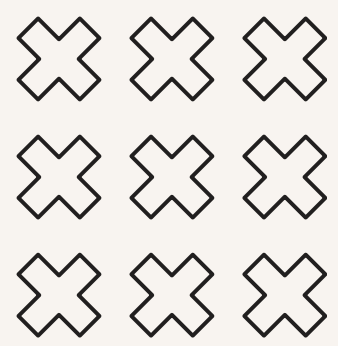




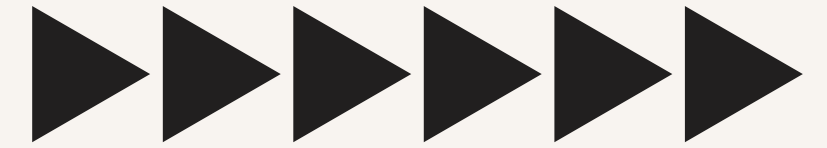
# Data Understanding

- The dataset comprises 54,432 entries and 8 columns, with various data types.
- Initial examination reveals missing values in 'review', 'rating', 'critic', and 'publisher' columns.
- Notable findings include diverse rating formats and top critics such as Roger Ebert and James Berardinelli.

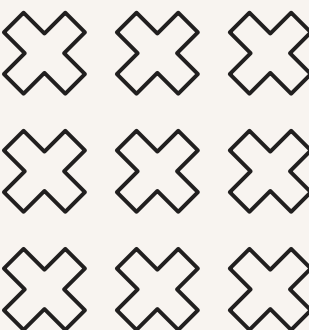




# Data Cleaning



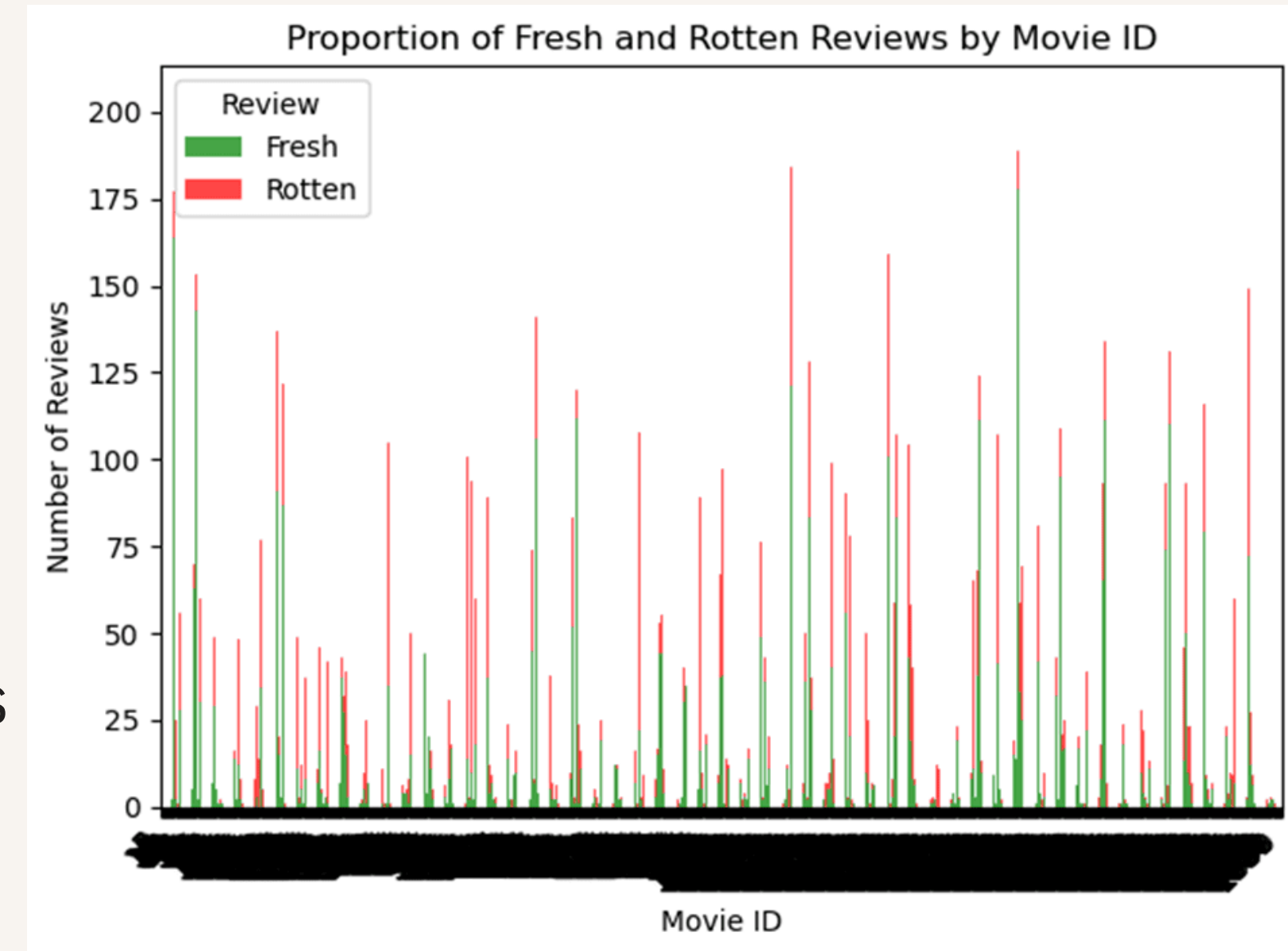
- After handling missing values and duplicates, 34,022 entries remain in the dataset.
- The cleaned data is free of missing values and duplicates, ensuring data integrity for further analysis.





# Data Analysis and Visualization

- Analysis reveals diverse rating formats and identifies top critics such as Roger Ebert and James Berardinelli.
- Proportion analysis shows 20,765 'fresh' and 13,223 'rotten' reviews.
- A stacked bar chart visually depicts the proportion of fresh and rotten reviews by movie ID, providing insights into critical sentiment and movie reception trends.

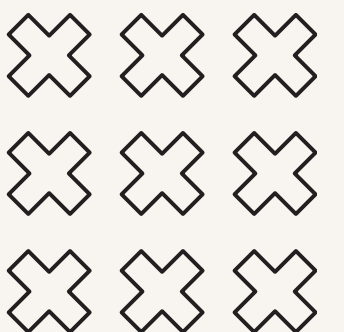


# Recommendations

1) Genres with High Box Office Performance: Microsoft should prioritize creating movies in genres such as Action, Adventure, and Sci-Fi, as these genres have consistently performed well at the box office.

2) Impact of Production Budget on Box Office Success: Microsoft should prioritize creating movies in genres such as Action, Adventure, and Sci-Fi, as these genres have consistently performed well at the box office.

3) Release Date Strategy: Microsoft should strategically plan the release dates of its movies to optimize box office revenue.



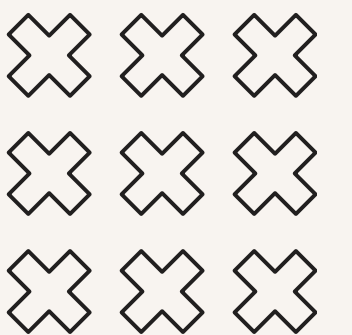
# Cont'....

## 4)Global Market Potential:

Microsoft should explore opportunities in the global market by analyzing the performance of movies in international markets.

## 5)Critical Reception:

Microsoft should prioritize creating high-quality content that resonates with both audiences and critics. Positive reviews from reputable critics can significantly impact a movie's success.







# Questions?

## Thank You!

