

The background features a large, stylized blue 'V' shape composed of two thick diagonal lines meeting at the bottom. This 'V' is set against a white background with faint, light blue geometric lines forming a grid-like pattern. Several solid blue circles of varying sizes are scattered around the central elements. The main title is centered within a dark blue horizontal bar that spans across the middle of the image.

# MICROSOFT VISUAL STUDIO

**BLUEPRINT**

REPORTED BY:  
BRADLEY OUKO

# SUMMARY

- Microsoft wants to enter into the lucrative business of creating original video content by coming up with a new movie studio. Based on scientific research on the domain, the following were the insights found:
  - The movie genres **Action, Adventure & Drama** are the most popular.
  - Summer months (**June, July, and August**) are peak performance months in terms of in terms of both average worldwide gross and average rating.
  - The safest budget range is **\$170M - \$255M**, with the highest profitability index of **96.96%**.

# CONTENTS



BUSINESS PROBLEM



METHODS



CONCLUSION



DATA



RESULTS



1

**BUSINESS  
PROBLEM**

# BUSINESS PROBLEM

- Microsoft wants a piece of the cake in the movie industry, however, it is limited in the domain knowledge. The research was done based on the following research questions:
  - Which are the **top performing movie genres**?
  - What are the **preferences of the target audience**?
  - How to break **market barriers**?

# BUSINESS PROBLEM

- The following are the **data questions** answered in this analysis:
  1. What types of films are currently performing best at the box office?
    - What are the characteristics of top-performing movies based on box office gross?
  2. Which movie genres have been the most popular and successful over time?
    - What are the trends in genre preferences?
  3. How does the movie budget impact box office revenue, and can smaller budget films be profitable?
    - How does the movie budget affect box office revenue?
  4. Are there seasonal trends in movie performance, and when is the best time to release a movie?



DATA



- The research retrieves information from IMDb, Box Office Mojo and The Numbers.

## Data Description

### 1. **IMDb**

#### IMDb Basics:

- Dataset: imdb.title.basics
- Description: 'movie\_id', 'primary\_title', 'original\_title', 'start\_year', 'runtime\_minute', 'genres'
- Relationship to Data Analysis Questions: Provides information on movie genres and basic movie details required for genre analysis.

#### IMDb Ratings:

- Dataset: imdb.title.ratings
- Description: 'movie\_id', 'averagerating', 'numvotes'
- Relationship to Data Analysis Questions: Allows us to explore audience preferences based on user ratings and reviews.





## 2. **Box Office Mojo** by IMDbPro

- Dataset: bom.movie\_gross
- Description: Contains box office gross information for movies;  
i.e 'title', 'studio', 'domestic\_gross', 'foreign\_gross', 'year'
- Relationship to Data Analysis Questions: Essential for analyzing box office success and financial performance of movies.



### 3. **THE NUMBERS**

- Dataset: tn.movie\_budgets
- Description: Contains movie financials;  
i.e. 'id', 'release\_date', 'movie', 'production\_budget', 'domestic\_gross', 'worldwide\_gross'
- Relationship to Data Analysis Questions: Enables us to analyze the impact of movie budgets on box office revenue.

The background is white with several light blue circles of varying sizes. A large, bold, blue 'Z' shape is positioned in the center-left. A horizontal blue bar with a gradient and horizontal lines is overlaid on the right side of the 'Z'.

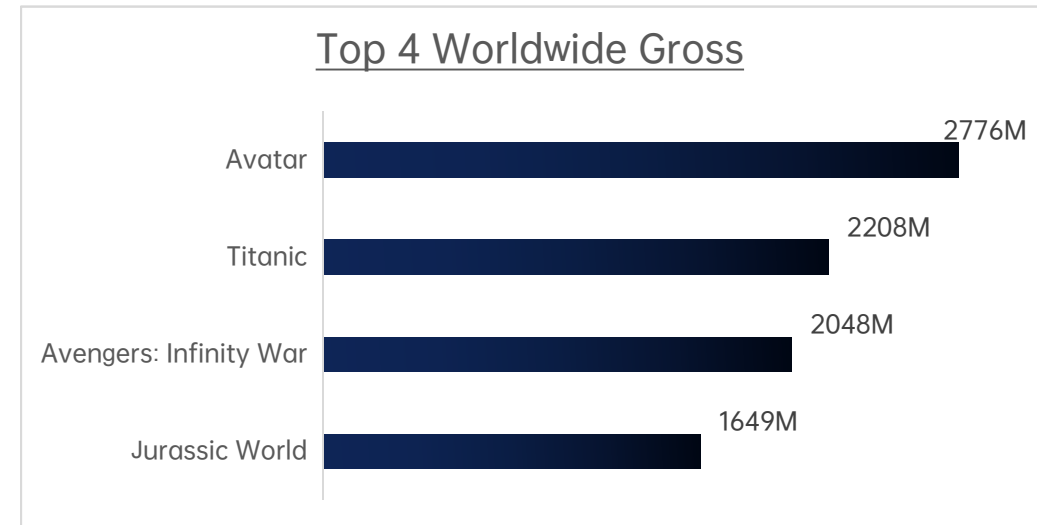
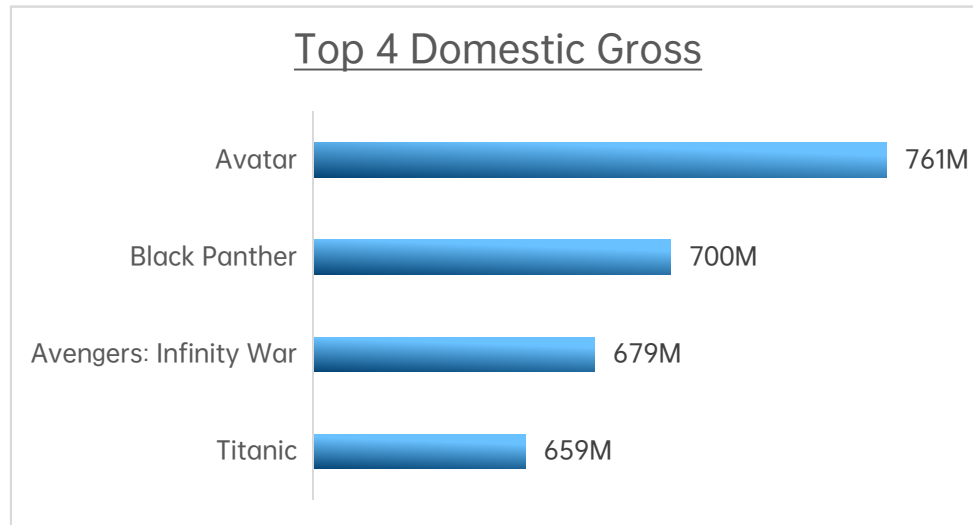
# METHODS

# DATA PREPARATION

- The following describes the conventional data cleaning process used to remove inconsistencies and operate on standardized data:
  1. **Data Loading**
    - The datasets were loaded into DataFrames in the working environment.
  2. **Data Cleaning**
    - Irrelevant variables & missing values were dropped; Relevant datasets were merged.
  3. **Handling Outliers**
    - Outliers i.e. genre outliers were put into consideration.
  4. **Feature Engineering**
    - Involves creating new variables that best suit the overall analysis;  
i.e. 'profit', 'release\_month' and 'foreign\_gross'

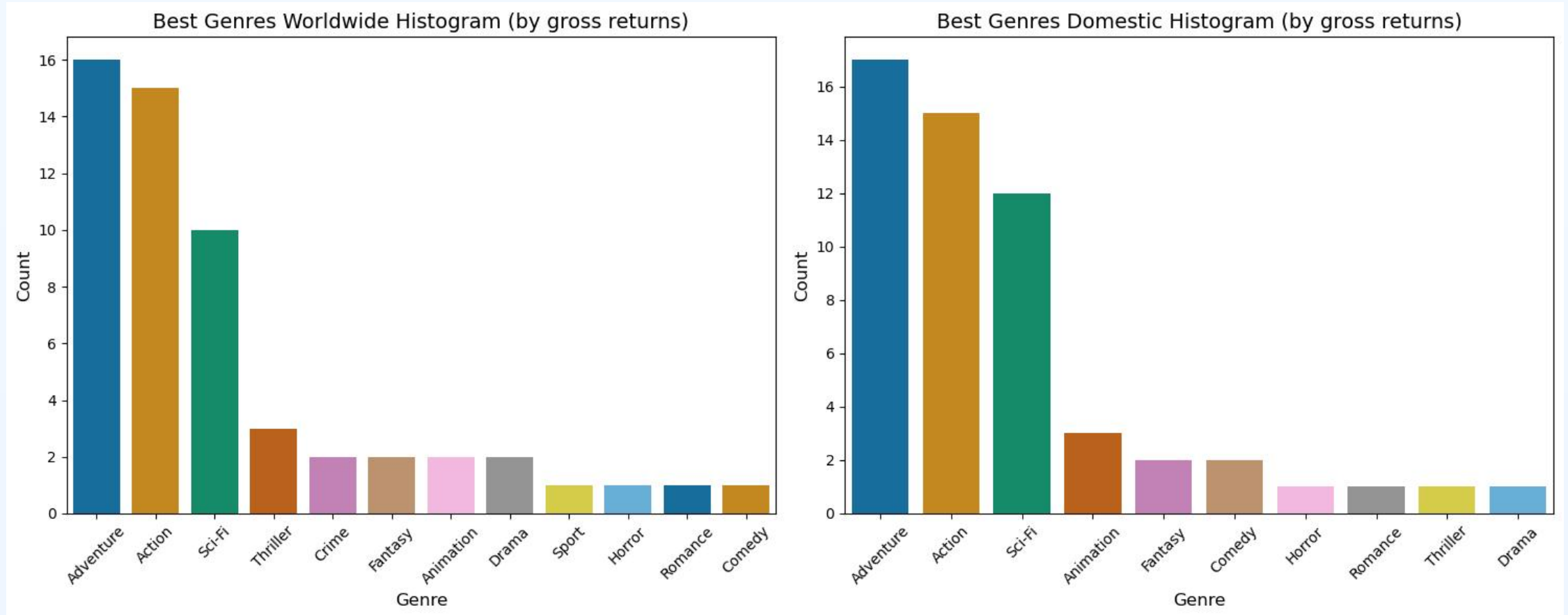
# ANALYSIS

- Exploratory Data Analysis was carried out in order to provide answers to our business questions defined earlier.
1. What types of films are currently performing best at the box office (based on box office gross)?



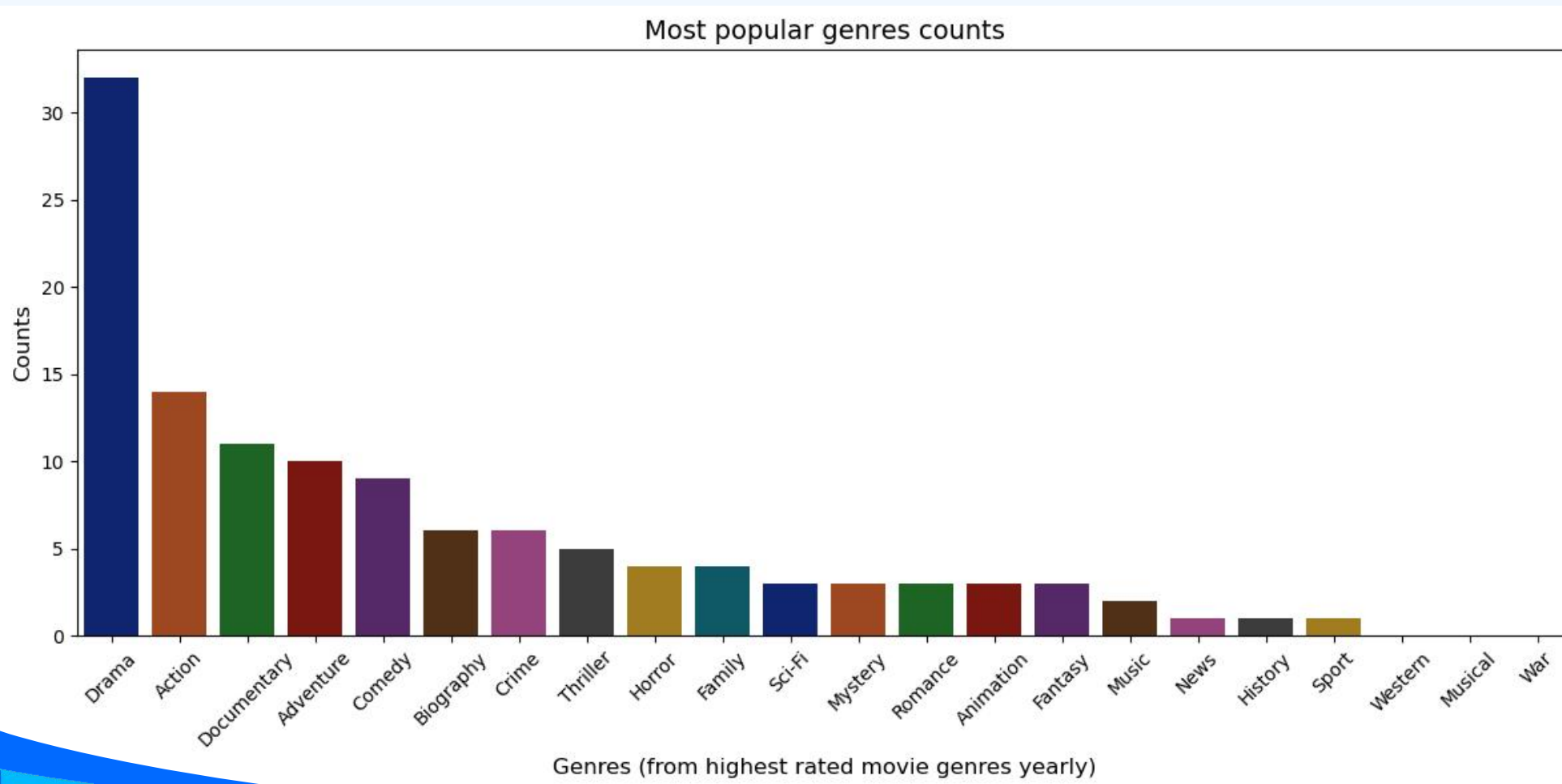
# ANALYSIS

- The most occurring genre in the top 20 best performing movies?



# ANALYSIS

2. Which movie genres have been the most popular and successful over time?



- Drama
- Action
- Documentary
- Adventure
- Comedy

# ANALYSIS

3. How does the movie budget impact box office revenue, and can smaller budget films be profitable?

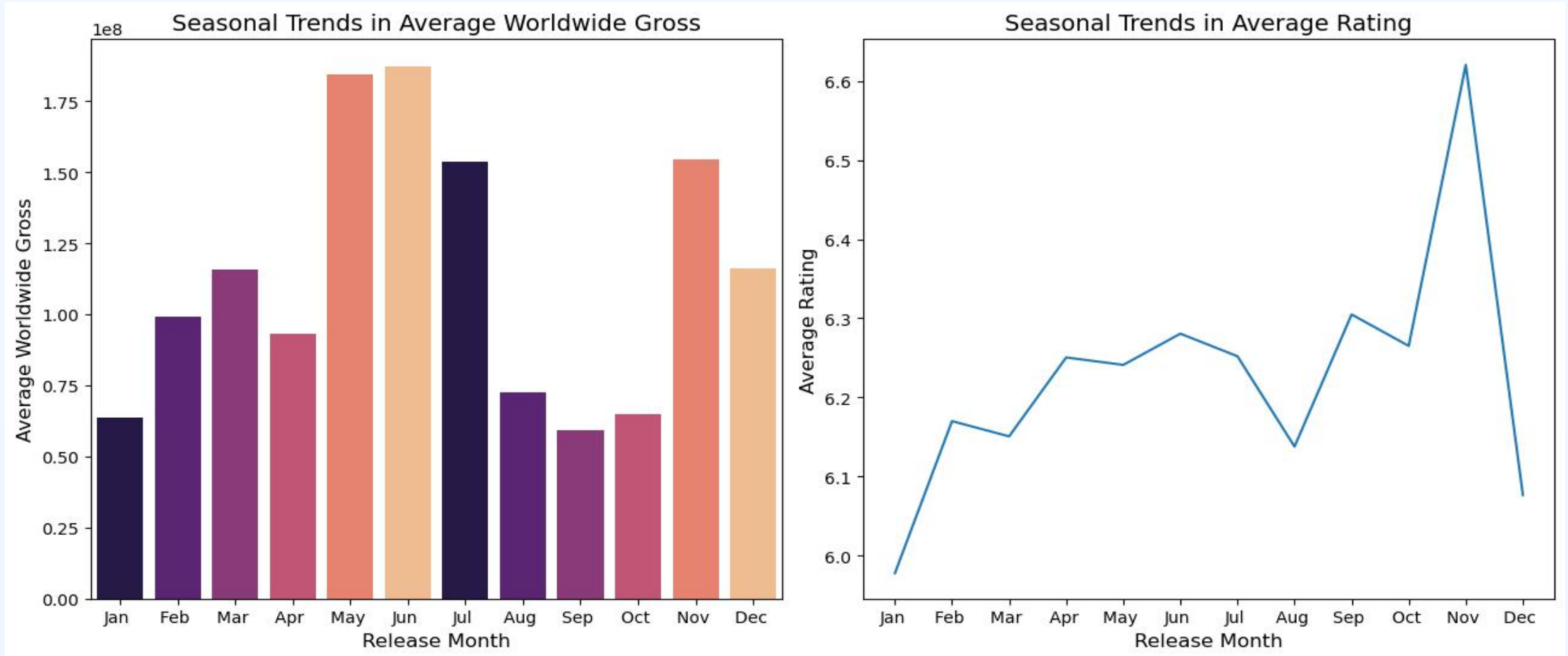
- The relationship between movie budget and worldwide gross was determined by calculating Pearson's correlation coefficient. (corr = 0.78 ; strong positive linear correlation)

<b>budget_range</b>	<b>profitability(%)</b>
1.4K - 85M	58.7
85M - 170M	91.0
170M - 255M	97.0
255M - 340M	88.9



# ANALYSIS

4. Are there seasonal trends in movie performance, and when is the best time to release a movie?



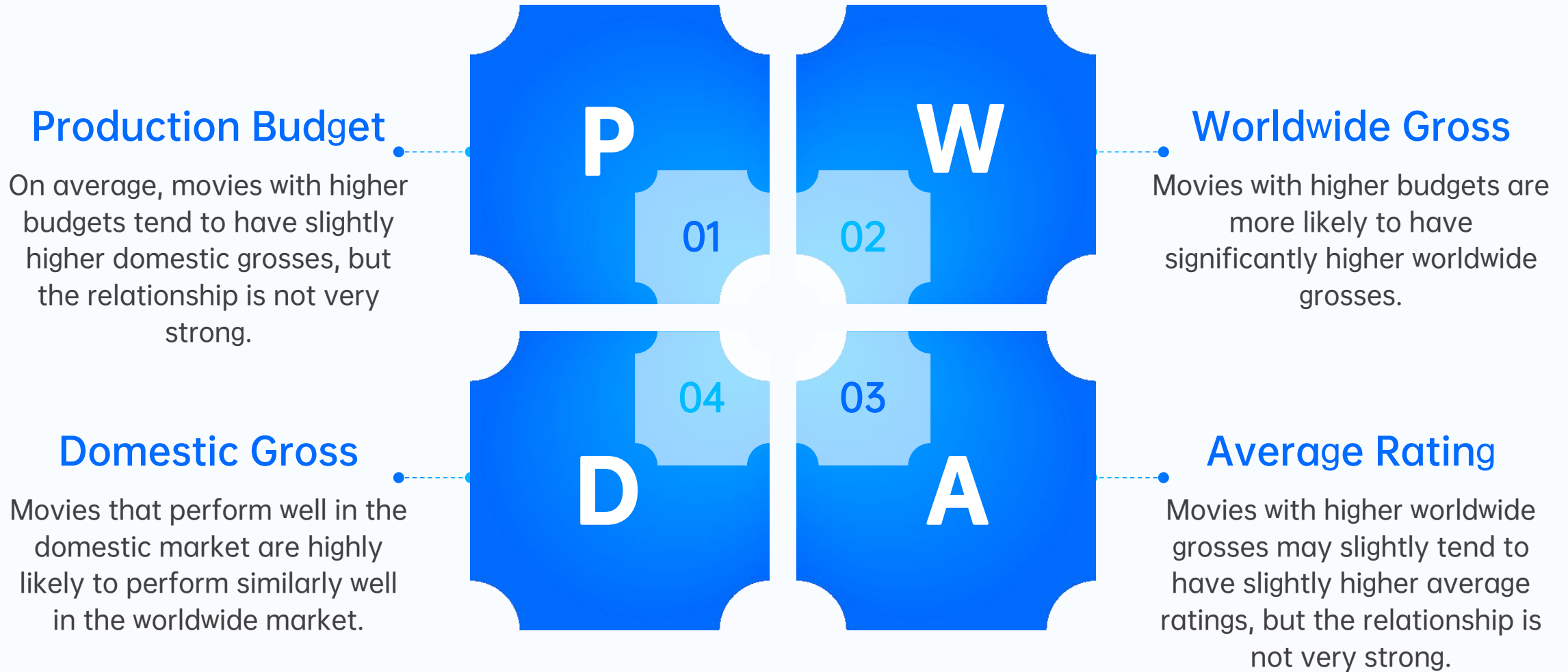


# 4 RESULTS

# RESULTS

- Months with Higher Average Worldwide Gross:
  - June has the highest average worldwide gross among all months, followed closely by May and December.
  - These months could be considered the best time to release a movie for maximum revenue, as they tend to generate higher box office gross on average.
- Seasonal Patterns:
  - Summer months (June, July, and August) show strong performance in terms of both average worldwide gross and average rating.
  - This could be due to the summer vacation period and increased movie going during this time.
  - December, being the holiday season, also exhibits high average worldwide gross and average rating, making it a potential lucrative period for movie releases.
- Top 5 best performing genres in terms of gross are:
  - Adventure
  - Action
  - Sci-fi
  - Thriller
  - Drama

# RESULTS



*The above information is based on the correlation coefficients between variables.*



# CONCLUSION

# CONCLUSION

## 1. Recommendations for the Business:

- Focus on Popular Genres

The genres that are likely to perform well are:

- Action
- Adventure
- Animation
- Comedy
- Documentary
- Drama
- Fantasy
- Scifi
- Thriller

# CONCLUSION

- Optimize Release Timing

Target months with historically higher average worldwide gross and audience ratings to maximize revenue and audience reach.

The best months of release are:

- August
- December
- July
- June
- May
- November

# CONCLUSION

- Budget Allocation

Analyze the relationship between movie budgets and box office revenue to make informed decisions about budget allocation for different movie projects. This can help optimize resources and manage financial risks.

The safest budget range is: \$170M - \$255M, with the highest profitability index 96.96%.



# CONCLUSION

## 2. Limitations of the Analysis:

a.) Data Quality: The accuracy and completeness of the analysis depend on the quality of the data used. Incomplete or biased data could impact the reliability of the results.

b.) External Factors: The movie industry is influenced by various external factors, such as competitor strategies, economic conditions, and cultural events, which are not directly captured in the analysis.

c.) Subjective Nature: Movie success involves subjective elements such as artistic creativity, storytelling, and audience emotional connection, which cannot be fully captured through data analysis alone.

# CONCLUSION

## 3. Future Improvements:

To enhance the project in the future, consider the following:

- a.) Validation and Testing: Validate the analysis results using external datasets or real-world performance. Testing the model on a separate hold-out dataset can help assess its generalizability.
- b.) Incorporate Audience Sentiment Analysis: Analyze audience sentiments and reviews using natural language processing techniques to gain deeper insights into audience preferences and reactions.
- c.) Market Research: Conduct market research and audience surveys to complement data analysis and gather qualitative feedback on movie preferences and expectations.



**“IT IS NOT ABOUT IDEAS.  
IT’S ABOUT MAKING  
IDEAS HAPPEN.”**

— SCOTT BELSKY  
CO-FOUNDER OF BEHANCE



# THANK YOU



[bradley.ouko@student.moringaschool.com](mailto:bradley.ouko@student.moringaschool.com)



[@Misfit911](https://github.com/Misfit911)