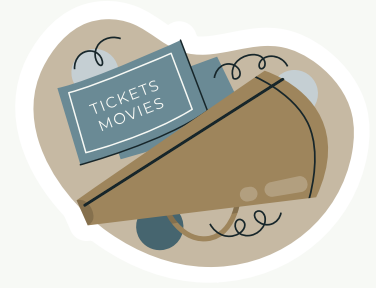


# Evolution of Movies Over the Years

Academic Project | Team Project | 3 Group Members



# Table of contents



01

About the  
Project

02

About the  
Dataset

03

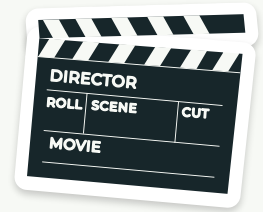
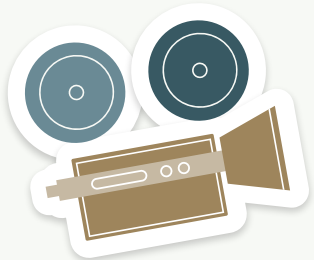
Findings and  
Insights

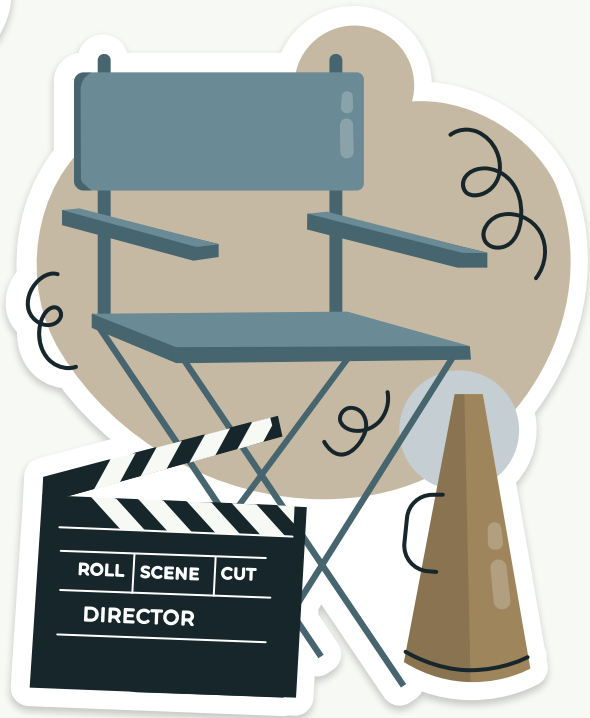
04

Conclusion

05

Recommendation





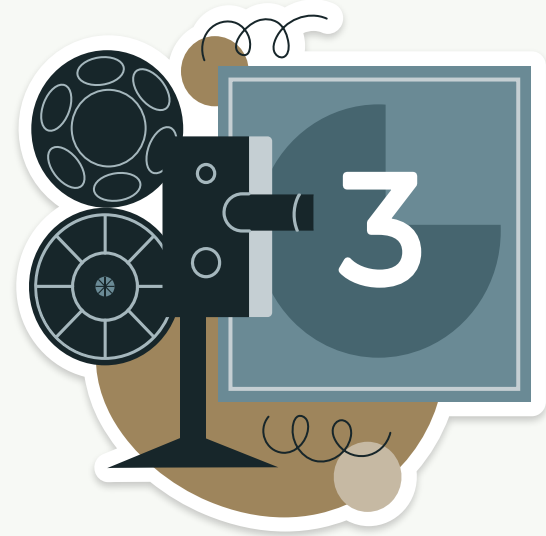
# 01

## About the Project



# Project Background

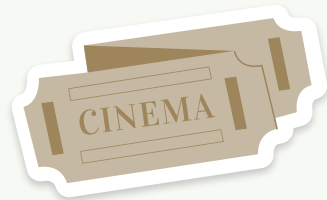
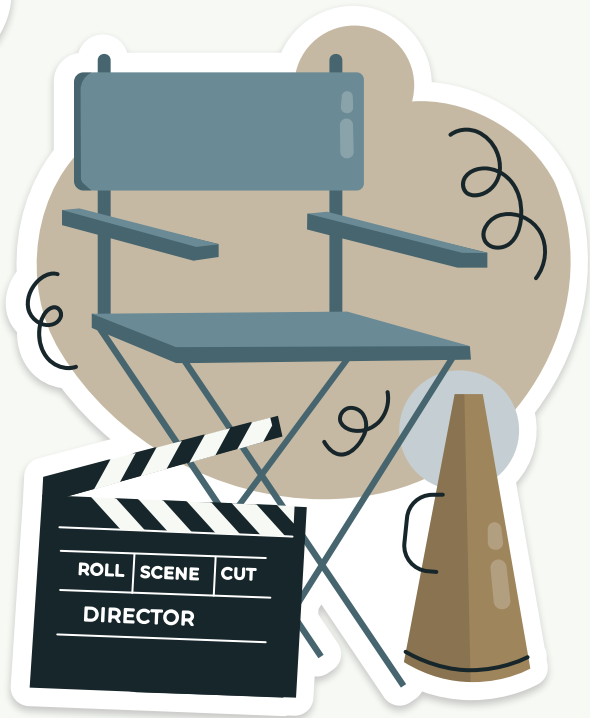
The world of cinema is a captivating realm, filled with stories, emotions, and cultural influences. But behind entertainment lies a complex industry driven by audience preferences, production decisions, genres, and box office revenue. To dive deeper into this, we turn to a rich dataset. With the analysis of such data, we can explore fascinating trends, identify audience preferences, and gain insights into the recent trends in the film industry



# Problem Statement

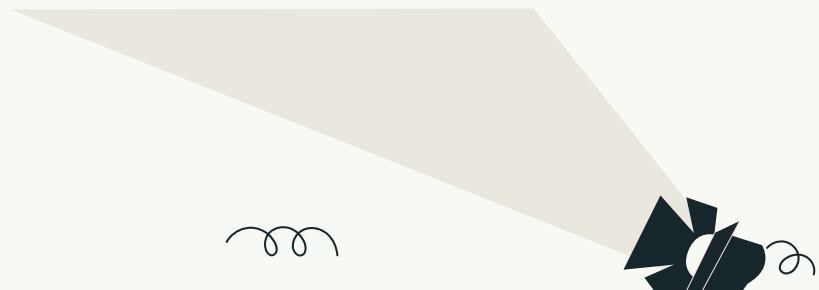
How has trends in movie industry changed with time.  
Are there seasonal trends?  
Do certain genres gain or lose popularity over time?





# 02

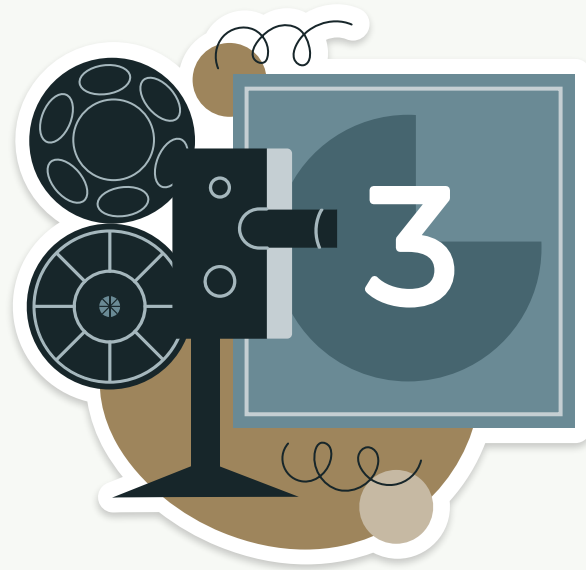
## About the Dataset



# About Movies Data



This data is **publicly available on Kaggle**, its name is **TMDB Movies (900k movies + daily updates)**. The dataset contains details about all the Hollywood movies released, planned, and rumored to be released between 1865 and 2115. It has collected information like status, revenue, budget, casts, genre, etc. related to the movies. The original size of the dataset was **906373 rows × 25 columns**, but after preprocessing done by us for analysis and visualization, the numbers have changed to **524753 rows × 36 columns**.



# Steps Followed



## Data Pre-processing

Using Python, we performed pre-processing steps.

## Conclusion

Wrap the whole analysis with findings, insights and suggestions.

01

## Define Problem Statement

Based on the objectives of our analysis

02

03

## Analysis & Visualization

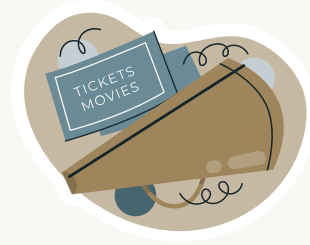
Decides what are the opportunity in data. Visualize it using Tableau

04





# Data Pre-processing



## Remove Duplicates

To improve the quality and accuracy of data. After that  
**906373 rows × 25 columns.**

## Formatting Dates

In US dates format. To have consistency in data.

## Deal with Missing Values

Drop the rows with null values in columns like title, release\_date etc.

## Tree-based imputation

To predict the missing values in revenue and budget columns.

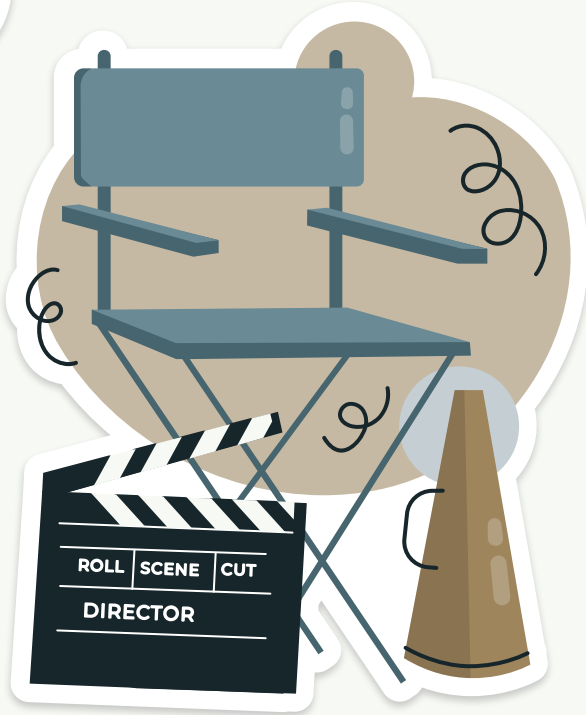
## Feature Engineering

On features like genre, directors etc. to improve the performance of data.

## Correlations

To identify features that are correlated in the dataset, useful during analysis.





# 03

## Findings and Insights



# Top 10 Most Popular Movies

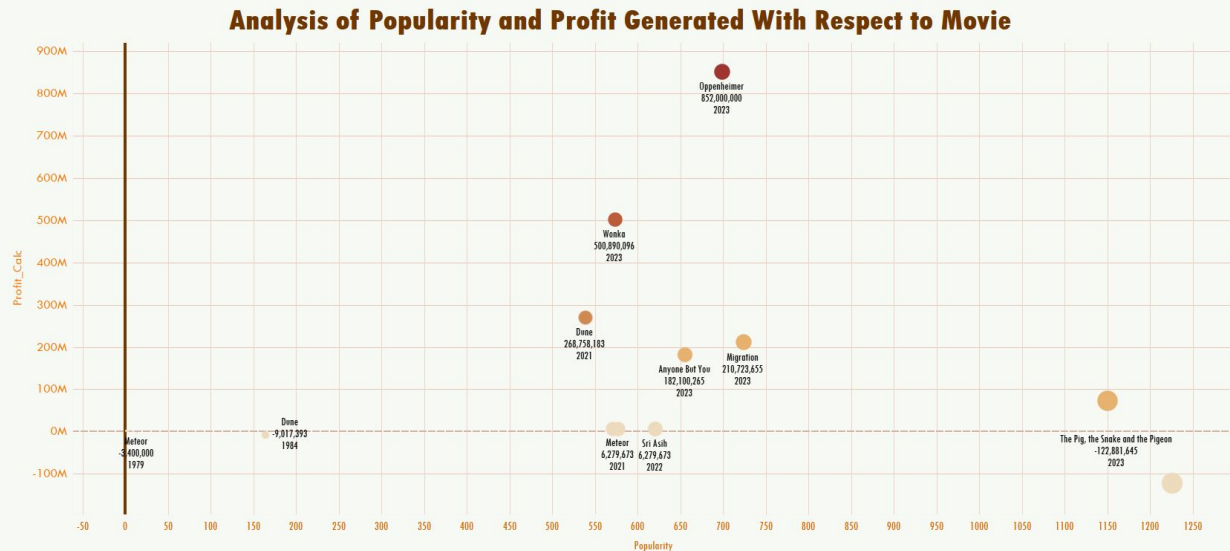
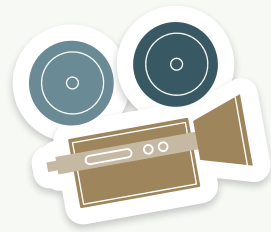


Figure 1. : Most popular and profit generated movies till 2023

**Oppenheimer (2023)** turns out to be the most profitable movie till date. Earning a **profit of around \$825 millions, even with less popularity(10%)**. The **Pig the Snake and the Pigeon (2023)** is the **most popular movie (17%)** of the era, but has not earned enough profit in box office, probably because of the **language barrier and wrong release month**. This analysis also supports the 38% correlation of popularity of the movie and profit earned by it.



# Impact of Budget on Runtime

- Back *in the old days* around 1911 and 1971, movies *used to be long*, but *now they are much smaller and more attention-seeking*.
- The *high-budget movies* used to be longer compared to low to moderate-budget movies in older days, that *had suffered huge losses in their box office collections*.
- *Low-budget movies* have as shorter *timeframe and a higher success rate* than others.
- The analysis shows that there *is a relationship between box collection, runtime, and the budget* category of the movies.

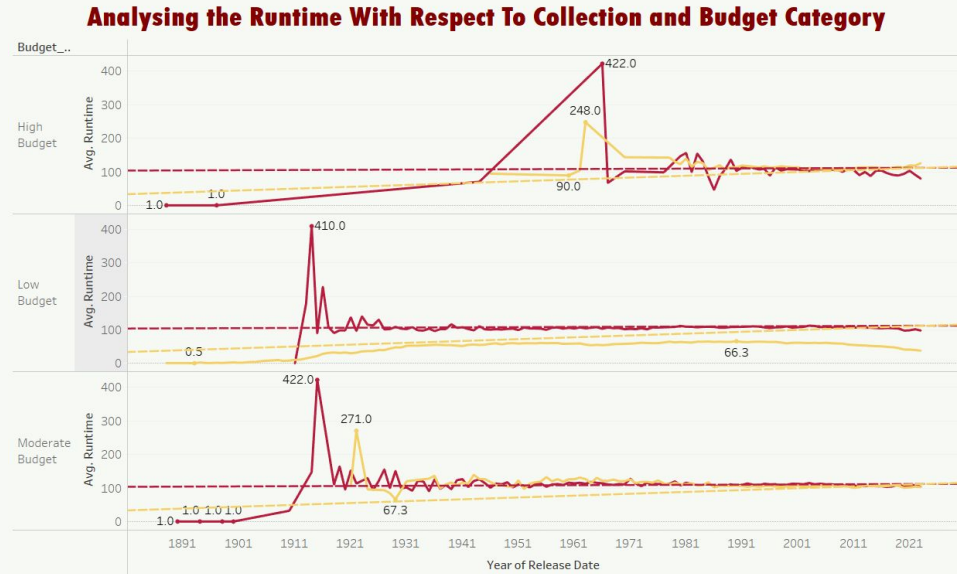
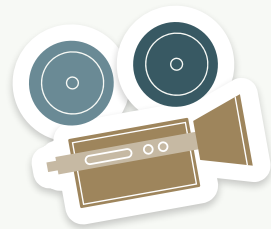


Figure 2. : Runtime and Budge Correlation





# Impact of Runtime on Revenue

- Movies with **runtime between 70–75 minutes** generate higher revenue and has high popularity.
- Whereas **shorter duration movies** less than with runtime **less than 50 minutes** earns lesser revenue due to the **lack of character development or unclear plot**.
- Or even **too long movies beyond 75 minutes** earns comparatively lesser because **it becomes either overwhelming or is unnecessarily dragged**.

Analysis of Runtime, Revenue and Popularity With Respect to Movie Recommendation

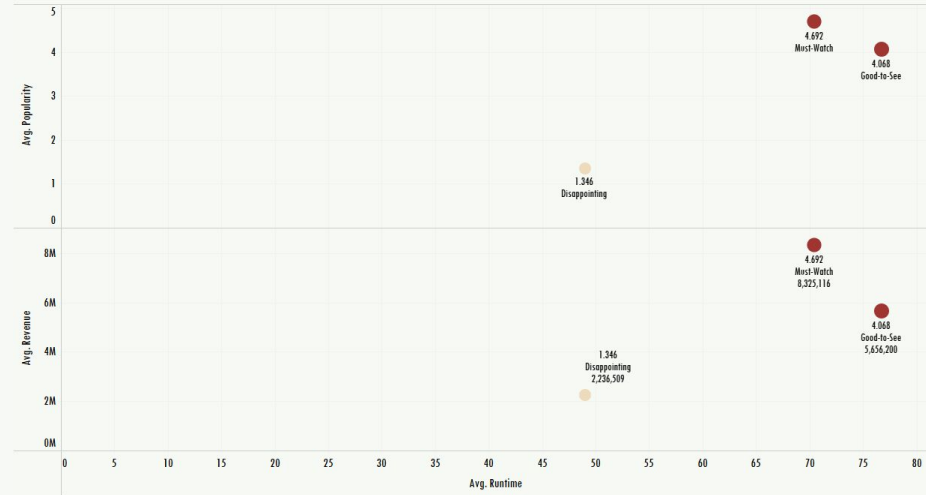


Figure 3. : Revenue and Runtime Correlation



# Budget and Popularity

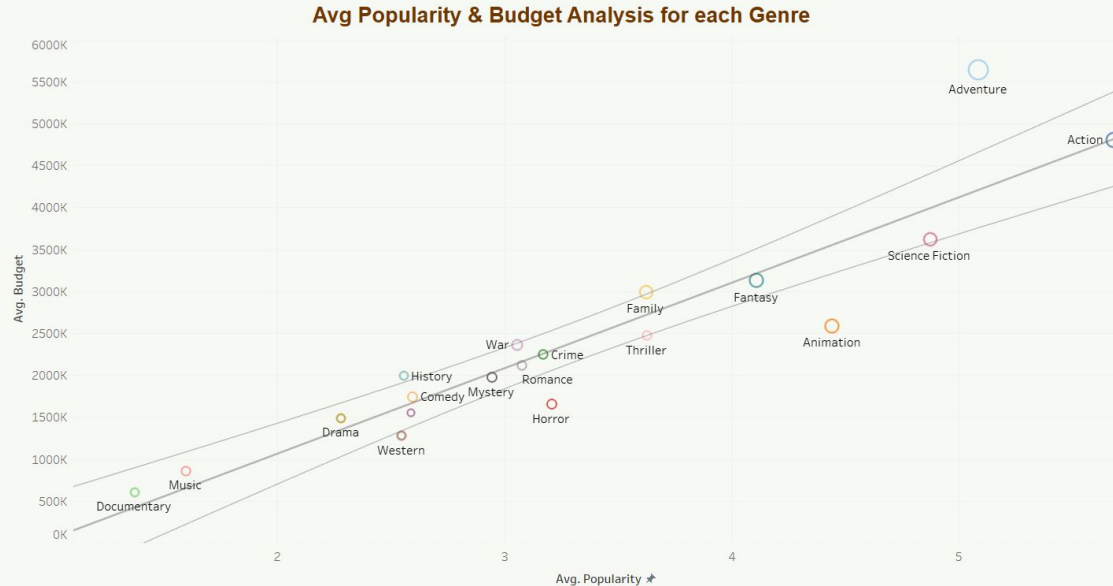
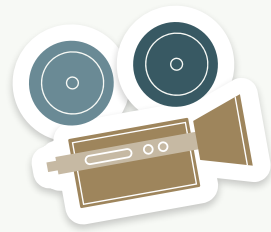


Figure 4. : Budget and Popularity analysis

It shows **linear relationship between budget and popularity**, higher budgets may signify higher popularity. Although most genres follow the linear trend, **some like animation, horror, and adventure are outliers**. **Animation and horror** movies have more popularity amongst the audience, despite having a smaller budget than **thriller and fantasy films**. **The most popular and high budget genre is Adventure, followed by Action and Fantasy**. **The analysis shows positive relation between Avg.Popularity and Budget.**

# Most Popular Genre



## Comedy 28%

The most popular genre that has generated an average **revenue of around \$13.6 and profit of \$8 millions** till 2023.



## Action 14%

With an average **revenue of around \$6.98 and profit till is negative**. These movies have **high budget, but many times revenue generated by them is very low, based on the popularity** of the movie.



## War 13%

With an avg. **revenue of \$6.5 millions and profit of \$2 millions** till 2023.

### Analysis of Revenue With Respect to Major Genre of the Top 10 movies until 2023

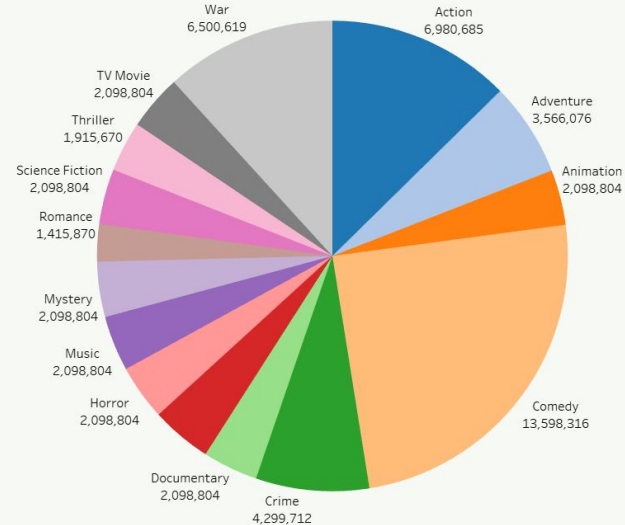


Figure 5. : Popular Genre



# Trends in Genre Choice

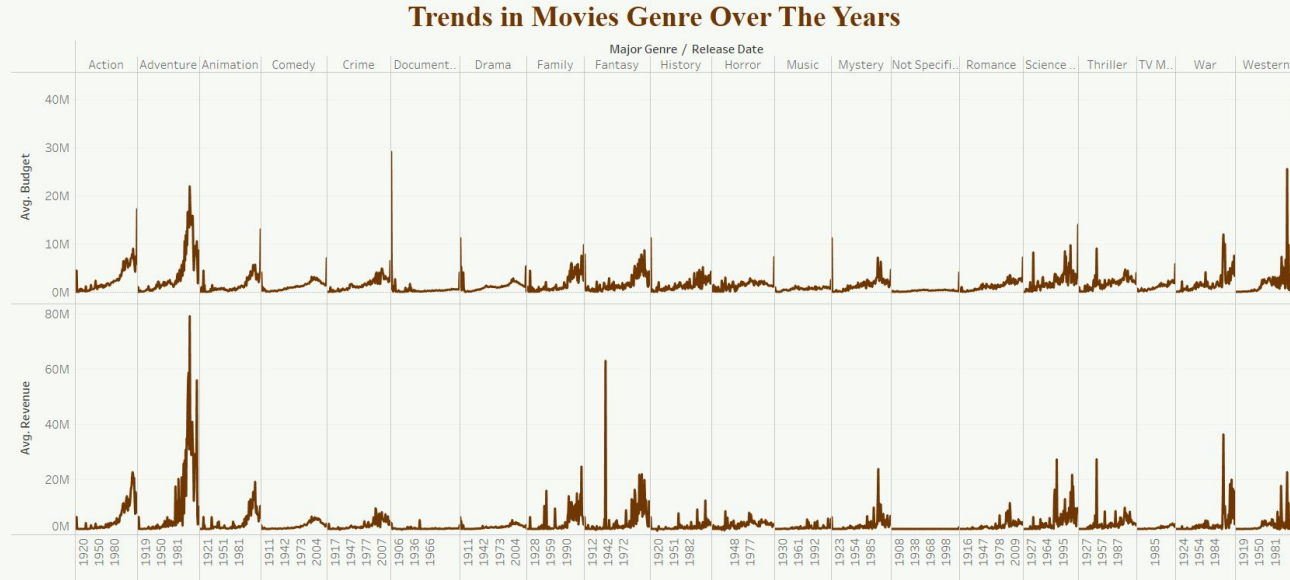


Figure 6. : Trends in Genre Over the years.

**Action Adventure and Animated movies appears to have increasing popularity over time. Comedy genre have a fluctuating popularity, but with a slight upward trend. Drama is consistently popular throughout the years, but with a possible slight decline in recent years. The popularity of other genres like Crime, History, Horror, Sci-fi, Thriller, War etc., appears to be fluctuating over time, with no clear upward or downward trends.**



# Genre Vs Avg. Runtime

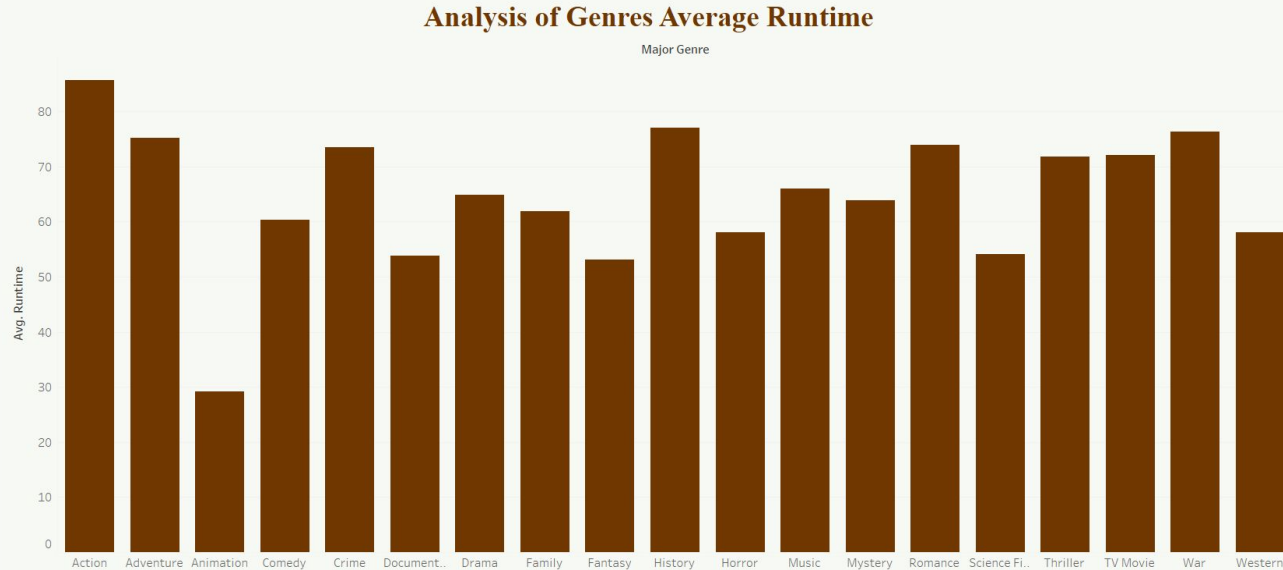


Figure 7. : Runtime analysis, based on genre

Revenue is impacted by runtime, apart from that genre of the movie. According to our analysis, **action movies are the longest with average runtime greater than 80 mins** sometimes. Followed by **historical and war based movies with the longest screen time**. That is because such movies require a lot of plot development and planning, having shorter version might lead to disappointments.



# Budget Vs Shooting Location

Analysis of Average Budget based on the Shooting Location

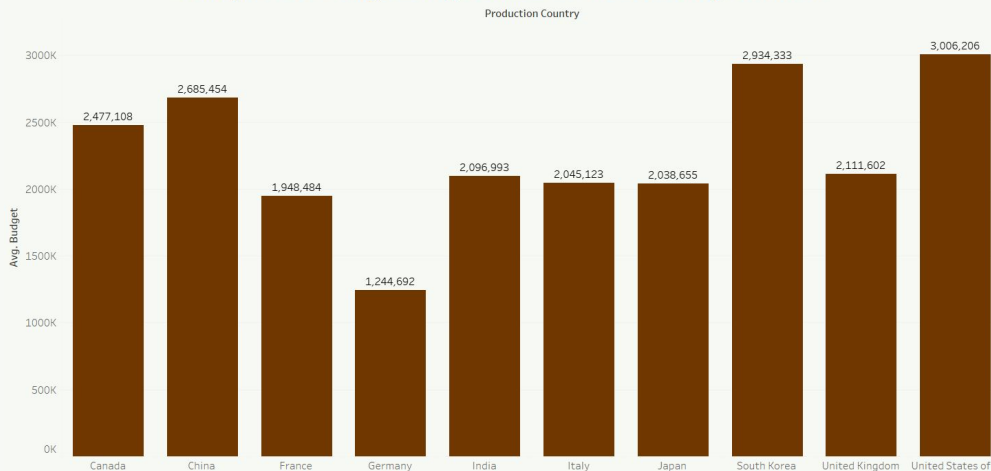
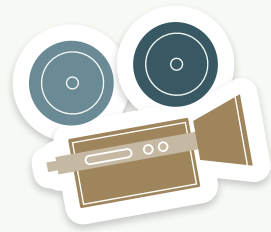


Figure 8. : Shooting location budget

- **USA seems** to be the most expensive location for shooting, with **average budget of \$3millions**
- Followed by **South Korea with \$2.93 millions and China with \$3.68 millions.**
- **Germany** being **the least expensive** of all, with \$1.24millions average budget.
- **USA** has the highest budget because of **prime location, Hollywood etc.**



# Production Companies



- **Marvel Studio** has been leading in movie industry because of its major revenue coming *from their popular Avengers Sequel. Company has earned around \$5 billions total till now, from the Sequel.*
- **Lucasfilm Ltd** has earned revenue of around **\$2.06 billions from legendary Star Wars movies**, making it one of the popular production house.
- **Amblin Entertainment, Universal Picture**, is popular for their Jurassic Park Sequel. This **adventurous movie has been part of peoples entertainment for long time now.**

Analysis of Revenue and Popularity of Production Companies based on Top 6 Movies

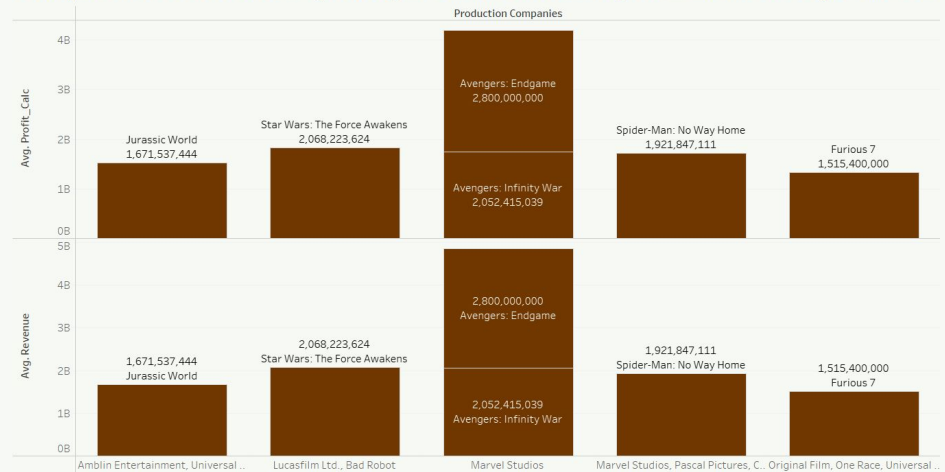


Figure 9. : Revenue Generated by Production companies



# Monthly Trends In Genres

## Analysis of Revenue and Profit Genrated by the movies based on Release Month

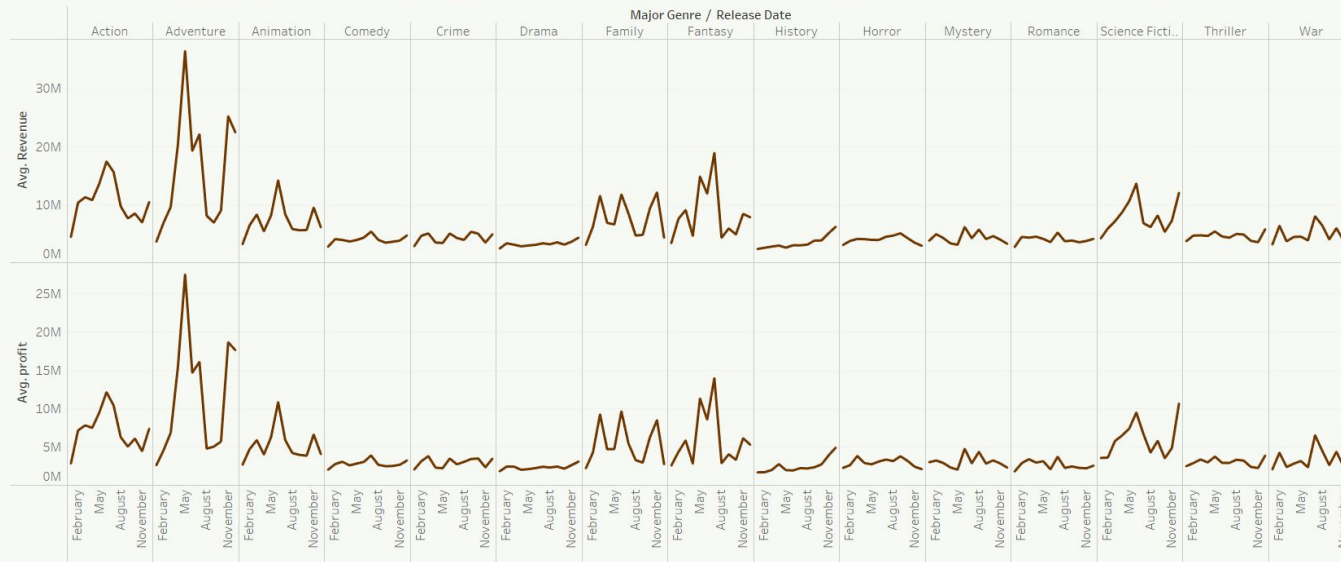


Figure 10. :  
Trending genres in  
different months of  
a year

This graph shows that **May, June, and July see the highest profit**, most likely due to higher revenue during the summer holidays. **Action, adventure, and fantasy movies often see much higher profits** than the average profit of all genres throughout the year. Whereas genres like history, TV movies, war, and crime movies earn significantly less.

# Revenue, Votes and Budget

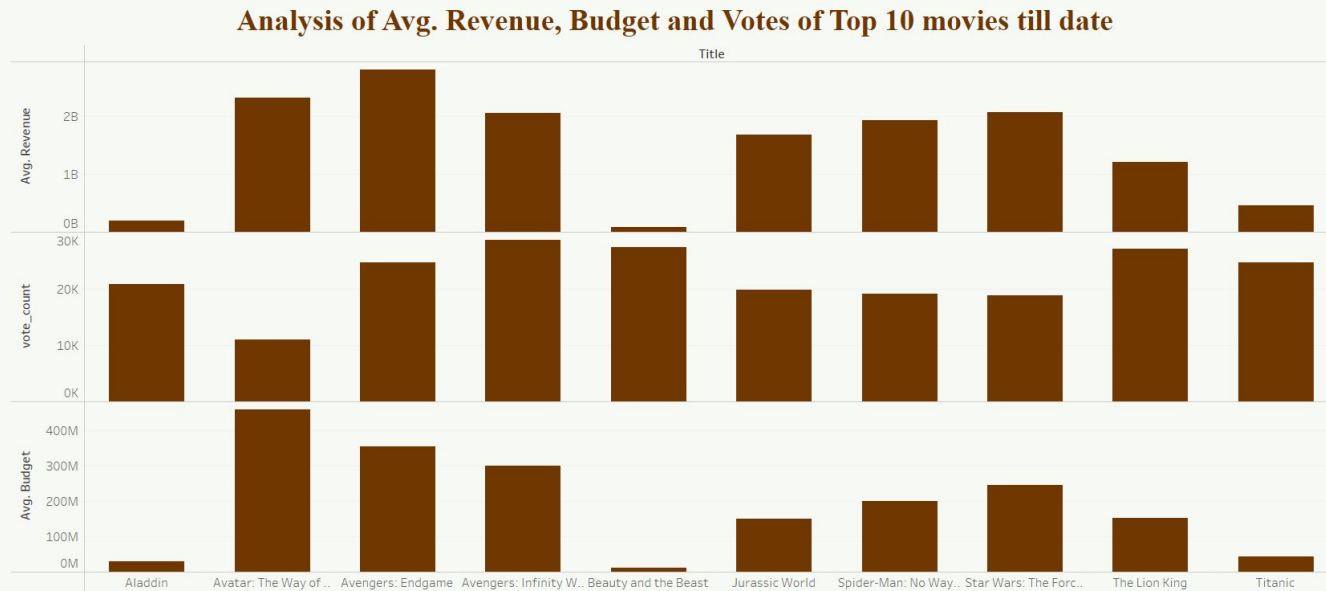


Figure 11.: Relation between Avg. Revenue, Budget and votes received by top 10 movies till date

The analysis shows **positive correlation between budget and revenue generated by the movie, however their relationship to people votes are volatile**. High budget doesn't mean that people will always like the movie, often people are disappointed after they spend on a movie because of its buzz. As in case of Avatar: The way of water, but this relationship is positive many times too as in case of movies like Avenger, Star War, Spider-Man etc.



# Genre Trends in 2023

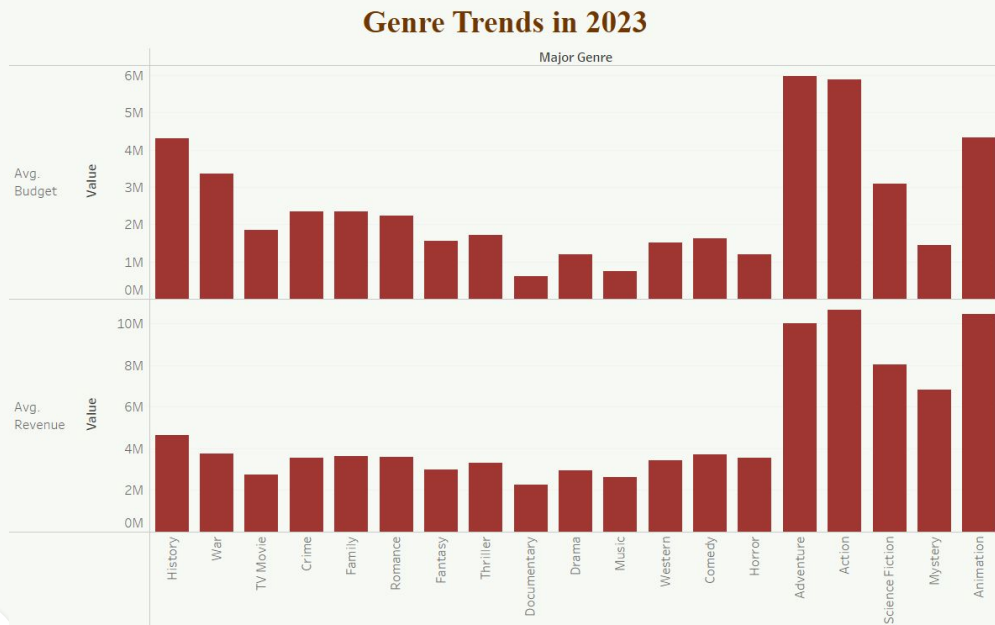
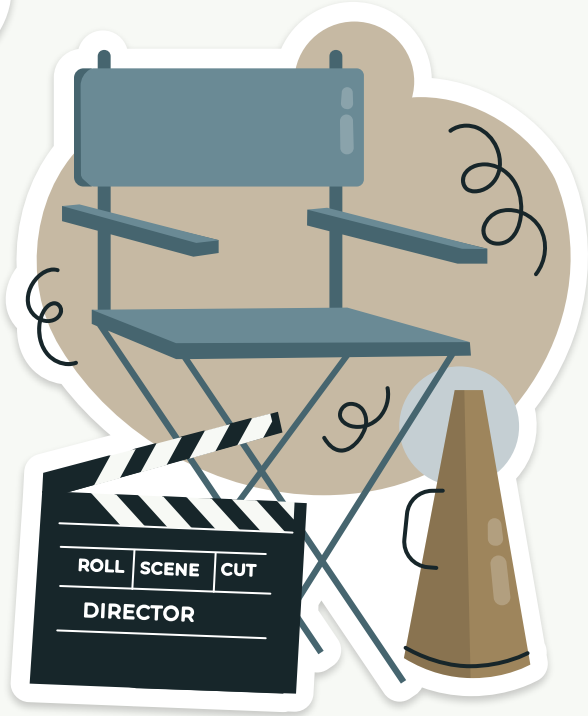


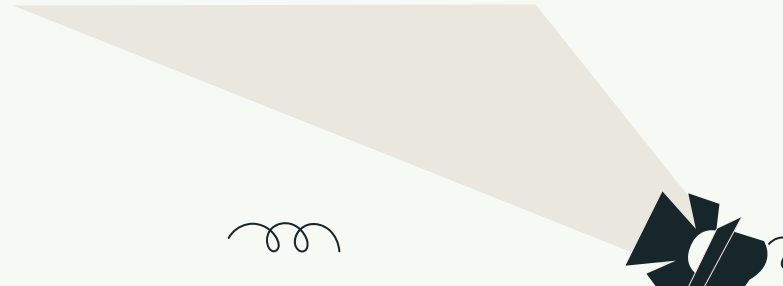
Figure 12. : Genre trends in 2023

- **Average Revenue** for movies **Action, Adventure and Animated** movies were **highest around \$10 millions**.
- **Actions and Adventure** movies has the **highest average Budget** of around **\$6 millions**
- According to our analysis, **Action, Adventure and Animated movies were in trend last year.**
- **Historical and War based** movies had high budget, but **failed to get people attention**, leading to low revenue.



04

Conclusion





# Factors that Impact Revenue



## Budget and Popularity

High budget movies tends have more buzz, that makes people more elevated and excited about it. Ultimately, then going to watch it in theaters, but their relation is volatile.



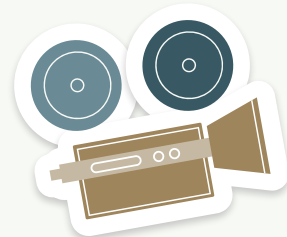
## Runtime

Movies of around 70–75 minutes has gained more attention, votes and popularity amongst people.



## Genre

Currently, Adventure, Action, Animated movies are in trend. However, this is the most fluctuating factor, that changes with time.





# Conclusion



Action or sci-fi movies are high budget due to special effects, equipments and technology used. Whereas, Comedy or dramas requires lesser investments



Longer films can offer more character development, but audience preference has changed to more attention-grabbing movies. Runtime that effectively tells the story without dragging it too long or rushing it.

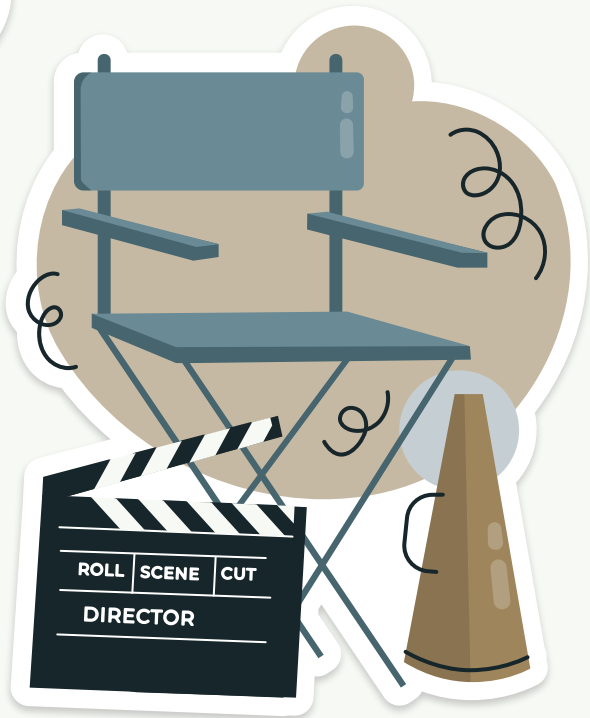


For high-profit earnings, production companies should analyze trending genres with higher profit margins (e.g., action, animation, adventure) and try to incorporate elements of those genres into their movies.



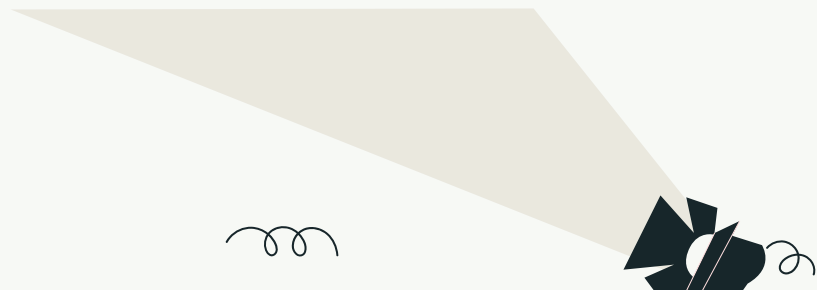
A shorter runtime allows for a low budget that could be redirected towards other aspects of production to make it more popular among the audience.





05

## Recommendations



# Recommendation



The releases date must be decided strategically, with the aim of higher profit earnings. Companies can go for summer blockbusters to capitalize on vacation periods.



Counter programming by releasing niche genres during lower-profit months to potentially fill a market gap. For genres like horror, thriller, and mystery, promote these genre movies using niche marketing during the off-season to increase revenues and fill the market gap.



Movies in the "Must-to-Watch" category are highly recommended. Runtime of these movies is between 70–75 mins. Such movies have higher popularity and have good recommendation ratings.




For high-budget movies, large-scale campaigns can increase their popularity, while smaller films could focus on targeted online marketing or film festival appearances to increase their popularity and profit earnings at the same time.





# References

- TMDB Movies (900k movies + daily updates) dataset
  - Evolution of Movies over the years Visualization on Tableau
  - Evolution of Movies over the years Case Study
  - TMDB analysis (basic stuff) | Kaggle
- 



Thank  
You

