[Introduction 1](#_Toc1786712247)

[Data cleaning 1](#_Toc929739797)

[Key Findings 2](#_Toc790814092)

[Correlation Between Budget and Gross Profit: 4](#_Toc297290122)

[Key Impacts on the Movie Industry during COVID-19: 6](#_Toc2129237945)

[Profit vs Budget 7](#_Toc719795207)

[Conclusion: 8](#_Toc1030422411)

# Introduction

This report analyzes the return on investment (ROI) across various movie genres to determine the most profitable categories for investment. Based on our analysis, **Action, Animation, Drama, and Comedy** exhibit the highest ROI, making them ideal candidates for increased funding. Additionally, we assess the impact of COVID-19 on movie production and release trends.

## Data cleaning

|  |  |  |  |
| --- | --- | --- | --- |
| **Columns** | **Missing Count** | **Imputation Strategy** | **Reasoning** |
| **rating** | 77 | Replaced with "Unknown" | Categorical, mode could be misleading. |
| **released** | 2 | Inferred from year if possible; otherwise "Unknown" | Small missing count. |
| **score** | 3 | Filled with **median** | Score is numerical and likely skewed. |
| **votes** | 3 | Filled with **median** | Votes are counted data and often skewed. |
| **writer** | 3 | Replaced with "Unknown" | Categorical. |
| **star** | 1 | Replaced with "Unknown" | Minimal missingness. |
| **country** | 3 | Filled with **most frequent country** | Country is categorical. |
| **budget** | **2171** | Filled with **median** or estimated from similar movies | High missingness; dropping loses too much data. |
| **gross** | **189** | Dropped rows with missing values | Revenue is crucial for analysis. |
| **company** | 17 | Replaced with "Unknown" | Small missing count, categorical. |
| **runtime** | 4 | Filled with **median** | Small missing count, numerical. |

# Key Findings

**Top 3 Highest Budgeted Genres:**

* **Action**: 88,022,413,068.00
* **Comedy**: 48,380,348,215.00
* **Drama**: 32,432,339,987.00

Top 3 Most Profitable Genres:

* **Action**: 155,413,442,686.00
* **Animation**: 57,830,975,796.00
* **Comedy**: 48,795,120,247.00

Genres with the Worst Profitability:

* **Music**: -20,389,986.00 (negative profit)
* **Musical**: -35,809,308.00 (negative profit)
* **Sport**: -19,432,371.00 (negative profit)

Genres with Positive Profit but Lower Budgets:

* **Animation**: Budget: 22,311,070,000.00; Profit: 57,830,975,796.00 (high return)

Genre Profitability Ratios (Updated):

The profitability ratio is calculated by dividing **Profit** by **Budget** for each genre.

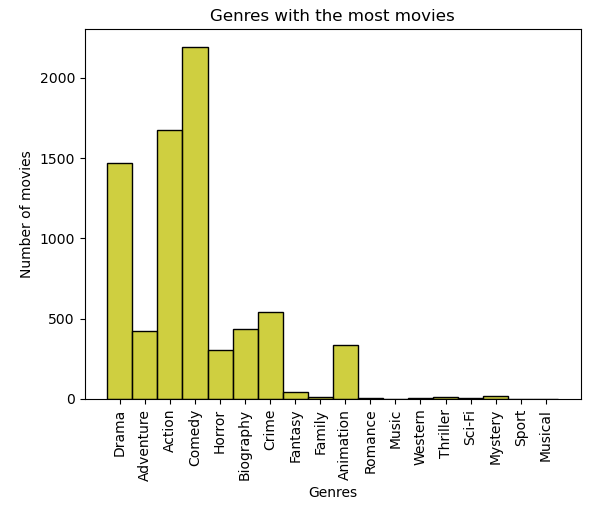
* **Action**: Profit/Budget = 155,413,442,686.00 / 88,022,413,068.00 ≈ **1.77**
* **Animation**: Profit/Budget = 57,830,975,796.00 / 22,311,070,000.00 ≈ **2.59**
* **Comedy**: Profit/Budget = 48,795,120,247.00 / 48,380,348,215.00 ≈ **1.01**
* **Drama**: Profit/Budget = 24,718,308,550.00 / 32,432,339,987.00 ≈ **0.76**

Insights:

* **Animation** continues to show the best profitability ratio at **2.59**, meaning it generates 2.6 times more profit than the budget spent.
* **Action** is still highly profitable with a ratio of **1.77**, yielding a substantial return on investment.
* **Comedy** breaks even with a ratio of **1.01**, indicating that the profits almost match the budget.
* **Drama**, with a ratio of **0.76**, shows a lower return on investment, meaning the profits are less than the budget, reflecting a less efficient use of funds.

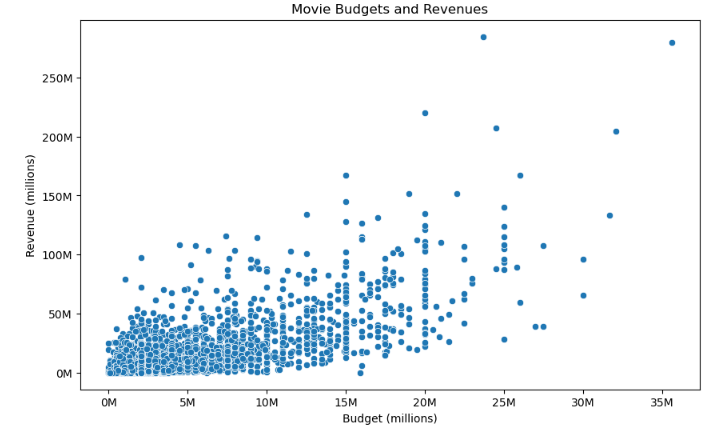
Genres with the Highest Number of Movies:

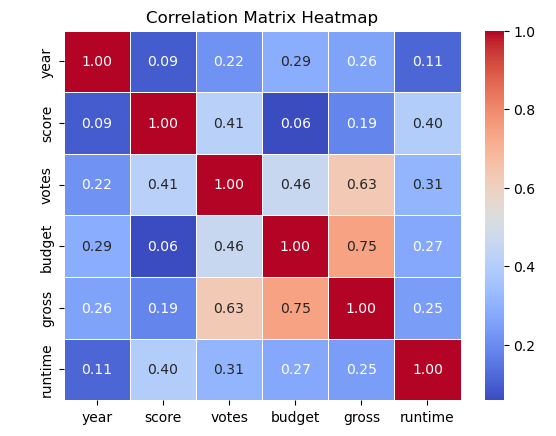
* **Comedy**: Highest number of movies, though with a **1.01** profitability ratio. Even though the return per movie is low, the high volume of movies ensures substantial profits.
* **Action**: Second highest number of movies, with a strong profitability ratio of **1.77**, making it a solid performer.
* **Animation**: While **Animation** has the lowest number of movies in the dataset, it is **included here** for its strong profitability ratio of **2.59**. Fewer movies, but higher profits per movie.
* **Drama**: Third in volume, but with the lowest profitability ratio of **0.76**, indicating that it is not as efficient in generating profits compared to its budget.



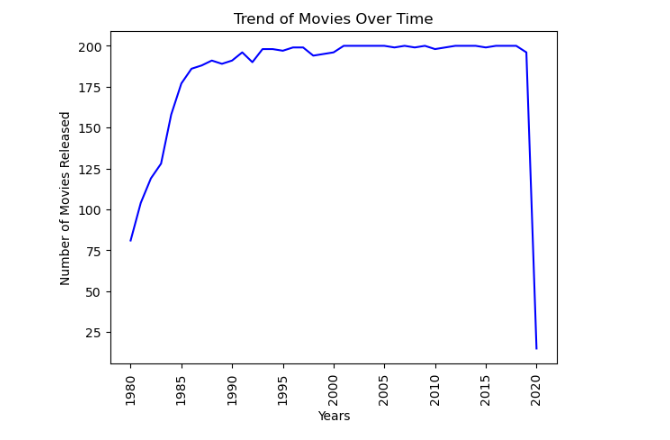
* **Comedy** and **Action** continue to dominate in terms of movie volume, but **Animation** deserves special mention. Even though **Animation** produces fewer movies than **Comedy** and **Action**, it generates significantly higher returns on each movie.
* **Drama**, despite having a substantial number of movies, lags in profitability.
* Genres like **Music**, **Sport**, and **Sci-Fi** have fewer movies, low budgets, and poor profitability, making them underperformers in the dataset.

## Correlation Between Budget and Gross Profit:





* The correlation coefficient between **Budget** and **Gross Profit** is **0.75.** This indicates a **strong positive correlation**, meaning that as the budget increases, the gross profit tends to increase as well. This suggests that higher-budget movies are more likely to generate higher gross profits.



During the **COVID-19 pandemic**, the number of movies produced drastically dropped due to factors such as lockdowns, production halts, and limitations on filming and distribution. This had a significant impact on industries like film production, with many studios postponing releases, reducing production budgets, or halting entirely for a period.

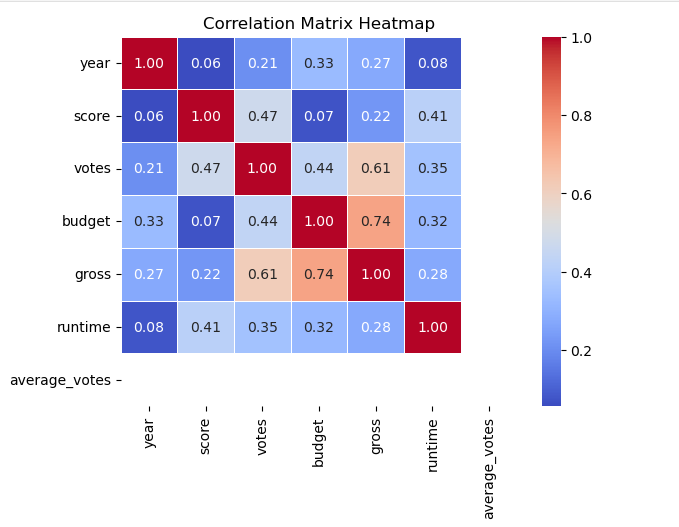
## Key Impacts on the Movie Industry during COVID-19:

* **Production Delays**: Many movies scheduled for production during the pandemic were delayed or canceled, causing a significant drop in the number of films being made.
* **Theatrical Releases**: Cinemas were forced to close in many parts of the world, limiting the release of films and reducing the profitability from box office revenues.
* **Shift to Streaming**: As theaters were closed, many studios shifted to streaming platforms for movie releases. This shift also changed how profits were earned (direct-to-streaming vs. theatrical release).
* **Budget Adjustments**: Due to the financial instability brought about by the pandemic, many productions had smaller budgets, and some genres experienced higher profitability per movie due to reduced competition.

## Profit vs Budget



Invest more in Action and Animation as the genres accumulated twice as much



The **0.74 correlation** indicates that there is a **predictable relationship** between **budget** and gross. Higher **budgets lead to higher revenues.**

### **Conclusion:**

Our analysis highlights key insights into the return on investment across various movie genres, helping guide future investment decisions in the film industry. **Action and Comedy** dominate in terms of movie volume and overall profitability, while **Animation** stands out as the most efficient genre, delivering the highest returns per movie despite fewer releases.

The correlation between **budget and gross profit (0.75)** confirms that higher-budget films yield higher revenues, reinforcing the importance of strategic investment in high-performing genres. Additionally, the COVID-19 pandemic significantly impacted movie production, shifting industry trends towards streaming platforms and altering profitability structures.

Given these findings, **Action and Animation emerge as the strongest candidates for investment**, offering substantial returns and financial stability. Meanwhile, **Drama, Music, Sport, and Musical genres underperform**, requiring a reevaluation of production strategies to improve profitability.

Going forward, stakeholders should **prioritize funding for Action and Animation films**, optimize production costs for Comedy, and explore innovative distribution strategies to adapt to the evolving entertainment landscape.