

A blue slot machine is the central focus, with several large stacks of US dollar bills (primarily \$100 bills) piled up to its right and in front of it. The scene is set on a dark, reflective surface under dramatic blue and purple lighting, with a blurred background of colorful neon lights.

Do Genres Impact Movies Revenue?

A data-driven analysis to discover the “Golden Egg Genre.”

by Matteo Ladislai

2024-01-11

Contents

Executive Summary	2
Research Question	3
The Start of The Analysis	5
Low-budget movies	6
Mid-budget movies	9
High-budget movies	12
Conclusion: A Nuanced Relationship	14
Appendix: Dive Into The Nitty-Gritty	16

Contacts

- Email: matteo.ladislai@hotmail.com
- LinkedIn: <https://www.linkedin.com/in/matteo-ladislai/>

Executive Summary:

The Main Findings in <200 Words

Does the genre impact a movie's revenue? **The answer lies somewhere between Yes! and No!**

Most of the genres don't make a significant difference in the movie's revenue. However, some genres consistently yield higher (and lower) profit percentages.

In the low-budget movies (<\$20M), Horror and Supernatural genres emerge as unexpected heroes. In a typical bad-case scenario, they still more than double the profit percentage compared to the whole group. The Sports genre is the worst-performing one. It achieves (in a good case scenario), 3 times less profit percentages than the group.

The Supernatural and Buddy genres dominate the mid-budget movies. They both achieve at least 50% more profit percentages than all genres. Meanwhile, the Drama genre falls short. It experiences 58.6% less profit even in a typical good-case scenario.

The same can't be said for the high-budget movies (>\$100M). This analysis tells us that brand recognition may play a more substantial role than the genre itself. Established fan bases, like the ones of Marvel or DC, are more important than the genre itself.

This creates a solid baseline for the film-making industry. However, further research (with the help of industry experts) is needed.

Research Question

Sure, movies are a form of art. But creating a movie requires many resources and studios are, after all, businesses. So the question arises; Is there a “Golden Egg Genre” — a genre that consistently yields a higher profit percentage?

The relation between genres and revenue is essential for informed decision-making in an industry where the stakes are high and the competition fierce. This is the question this analysis tries to solve.

Methodology:

Context: What is the profit percentage of a film?

In this analysis, profit percentage measures how much revenue a movie generates at the box office compared to its initial budget. If a movie has a budget of \$10M and a box office of \$15M, the profit percentage is 50%.

The box office is the gross revenue collected at theaters all over the world.

Data Source

Budget numbers for movies can be both difficult to find and unreliable. Studios often try to keep the information secret and use accounting manipulation to inflate the budgets.

The online place with the most reliable data is Wikipedia. This has been confirmed also by the website The Numbers.

The data of interest (box office and genre) are not in a tabular format anywhere on the web. Hence, the data analyzed are scraped from almost 5,000 pages of Wikipedia — all the American movies from 2004 to 2023.

Limitations

As discussed before, information about box office and revenue are manipulated by studios. This introduces an element of uncertainty into our analysis. Even if it is mitigated as much as possible.

Moreover, the gross revenue only includes the revenue made in theater sales. We aren't considering money made outside of the theaters — mediums such as public TV, video rentals, Netflix, or Hulu.

For each budget group, we consider just the genres that surpass a certain threshold. That's because analyzing all the genres may probably lead to being tricked by chance since some genres have too few movies.

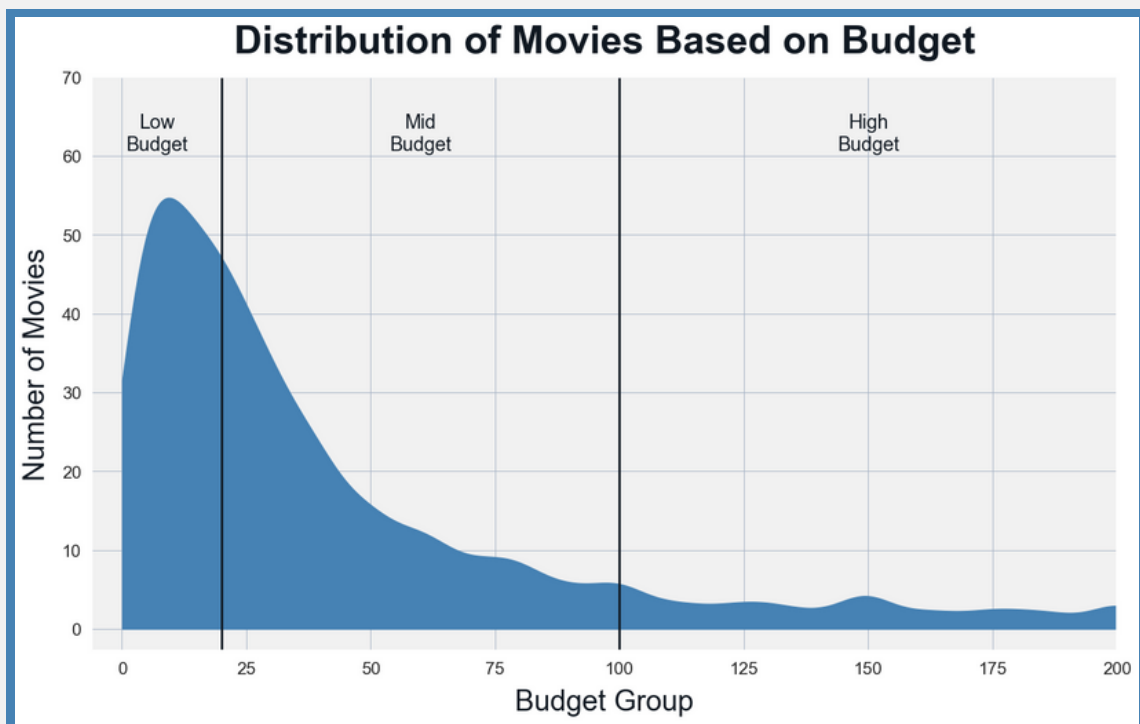
The Start of The Analysis

Popular movies vary greatly in the resources used during production. Some, like Toy Story, undergo a five-year animation process with a \$30 million budget. Others, such as Tangerine, are shot using iPhones on a modest \$100,000 budget.

For a more accurate and comprehensive analysis, movies are categorized based on their budget:

- Low-budget movies: below \$20M.
- Mid-budget movies: from \$20M to \$100M.
- High-budget movies: Above \$100M.

The graph below shows the distribution of movies based on the Budget.

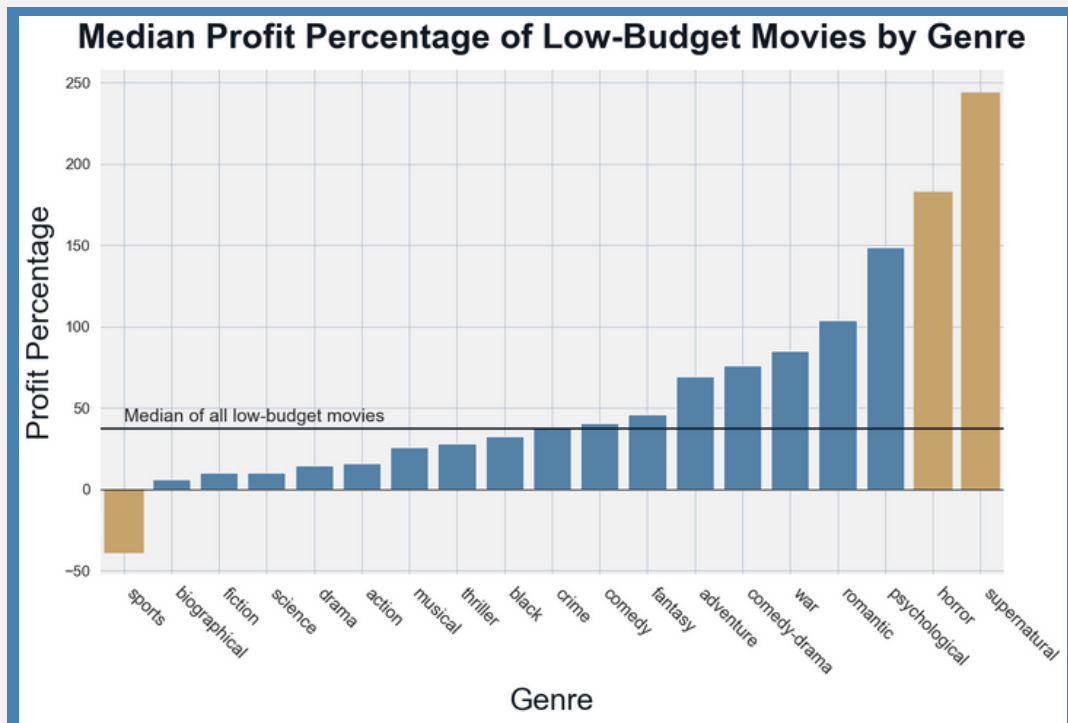


Low-budget movies: Does the genre impact their revenue?

Before anything, it's imperative to see if there really is some (significant) relationship between genre and profit percentage. Our initial step involves an ANOVA test, detailed in the appendix.

Based on it, **we can confidently say that genre does influence profit percentage.** However, understanding the behavior of each genre requires a deeper examination.

The following graph illustrates the median of the low-budget movies by genre. Think of the median as the typical profit percentage a movie of a certain genre gets.

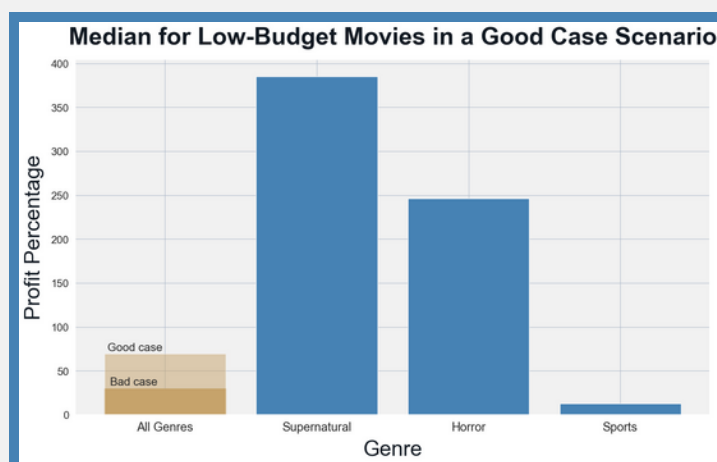
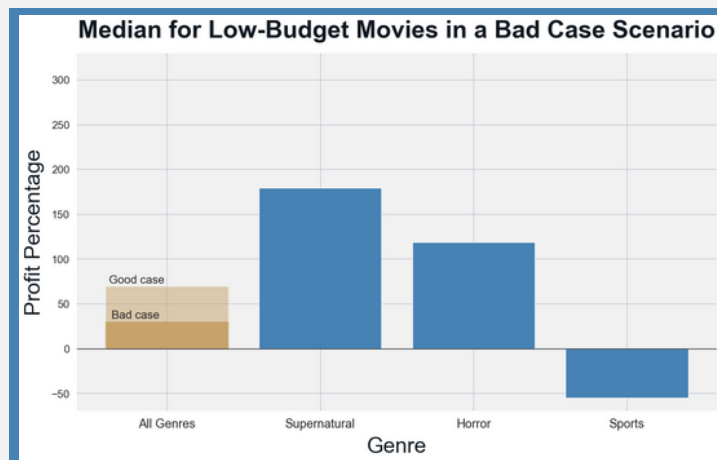


You can see that the median profit percentage varies a lot. Yet, the variation may just be due to chance or that particular genre has better quality movies. This will be explored shortly. For now, just know that the genres in yellow are the ones that really influence the profit percentage.

Chance may trick us

Some genres have very different median values compared to the overall group. Yet, caution is warranted; **we must ensure these differences aren't mere products of chance.** Two additional tests are conducted to validate our findings. (See the Appendix for more details.)

To validate the genuine impact of a genre on profit percentage, we explore bad-case and good-case scenarios. Consider the graphs below;



Even in a typical bad-case scenario, "Supernatural" and "Horror" outperform the typical good-case scenario for all genres. Specifically;

- Bad case scenario for Horror: 118.9 Profit percentage.
- Bad case scenario for Supernatural: 179.5 Profit percentage.
- Good case scenario for all genres: 69.4 Profit percentage.

The horror genre nearly doubles the profit percentage, and the supernatural genre almost triples it!

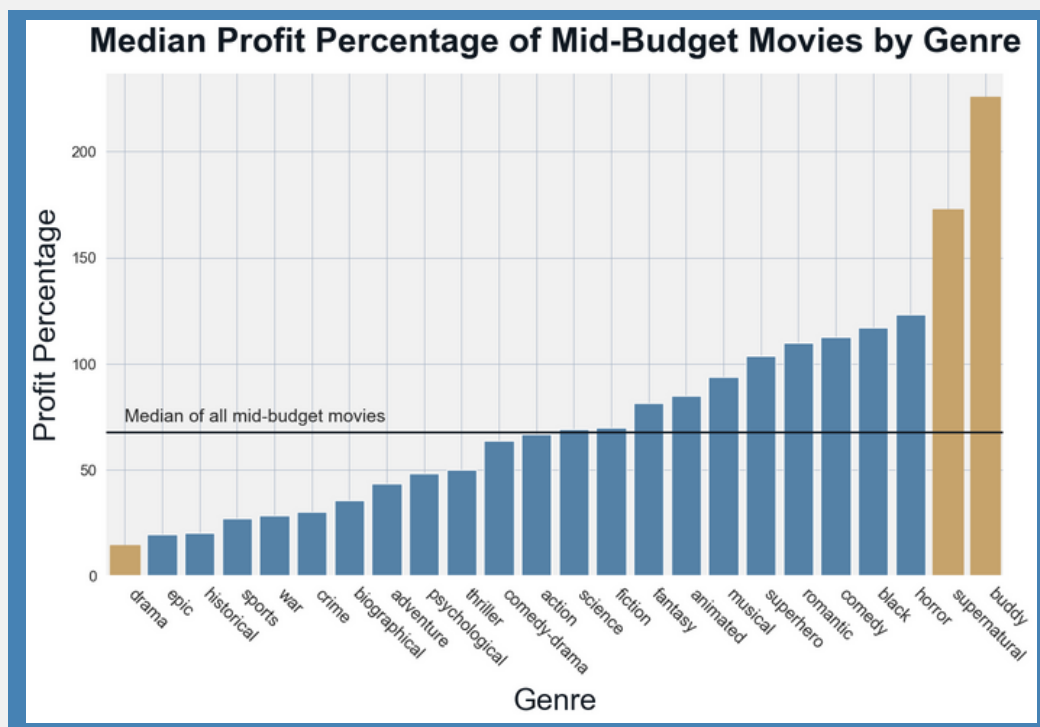
On the other hand, even in a typical good-case scenario, the "Sports" genre falls short. **It gets nearly three times less profit percentage compared to the overall group's bad-case scenario.**

- Good case scenario for Sports: 12.5 Profit percentage.
- Bad Case scenario for all the genres: 30 Profit percentage.

Mid-budget movies

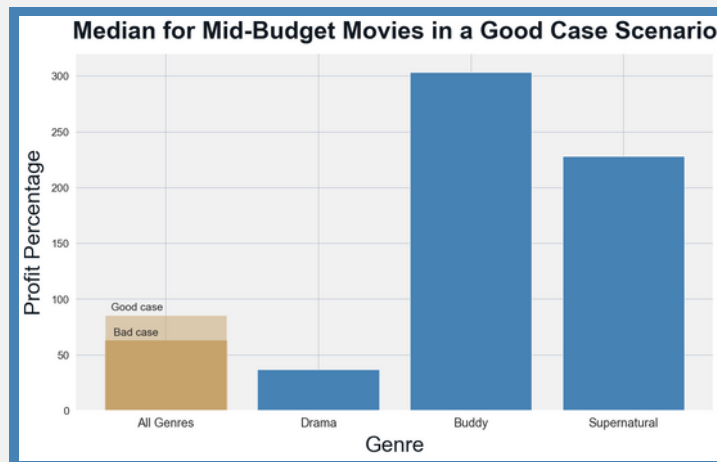
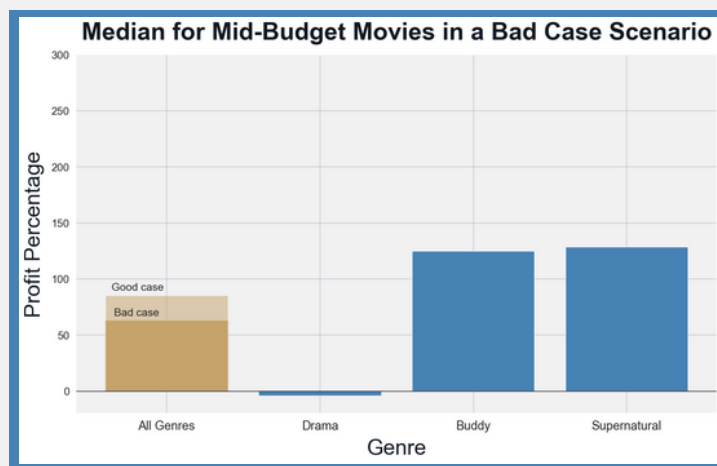
Now, the analysis builds upon the groundwork laid in the preceding section. The steps are repeated to discover whether genre influences profit percentage in mid-budget movies.

One of the few differences is the genres considered. Certain genres, such as "Sports," are excluded due to their limited representation in this budget range. The following graph illustrates the median of the mid-budget movies by genre.



As for low-budget movies, we have to make sure that chance is not tricking us.

The analysis reveals that three genres stand out as significant influencers: Drama, Supernatural, and Buddy. These are the graphs for the bad and good case scenarios.



In a typical bad-case scenario, "Supernatural" and "Buddy" outperform the typical good-case scenario for all genres;

- Bad case scenario for Buddy: 125.1 Profit percentage.
- Bad case scenario for Supernatural: 128.6 Profit percentage.
- Good case scenario for all genres: 82.6 Profit percentage.

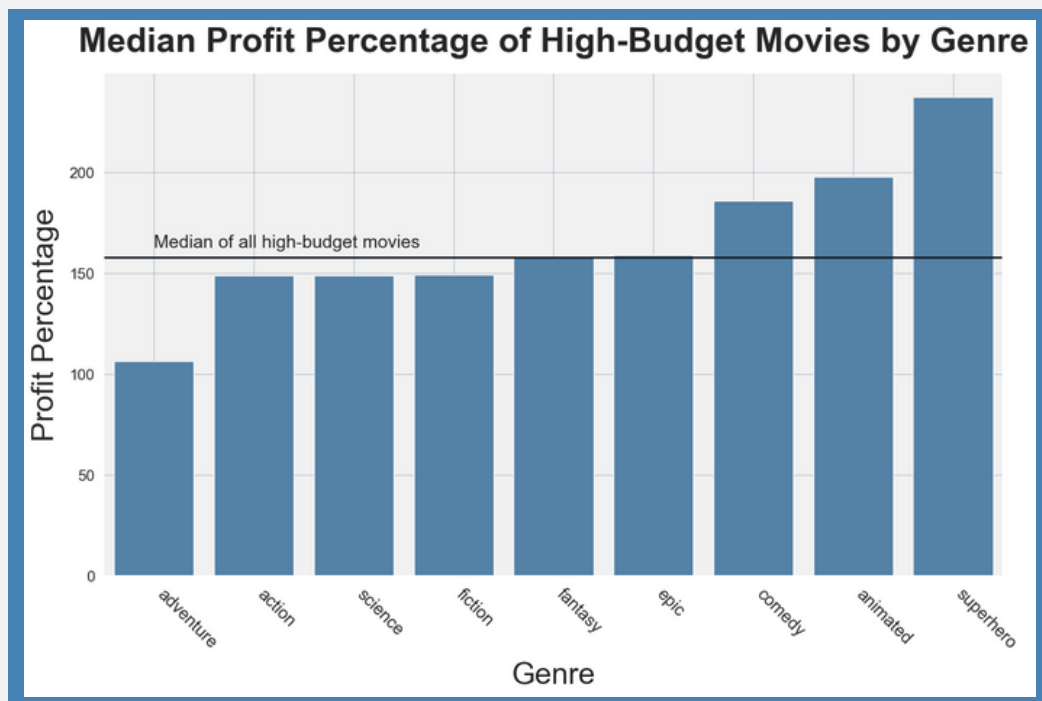
The Supernatural genre secures at least 51% more profit percentage compared to the overall group, while the Buddy genre attains at least 55.7% more.

Instead, the Drama genre gets at least 58.6% less profit percentage compared to the overall group in a typical bad-case scenario.

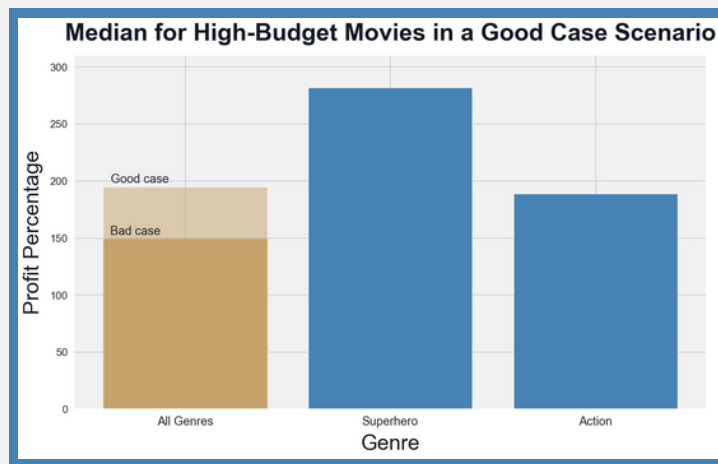
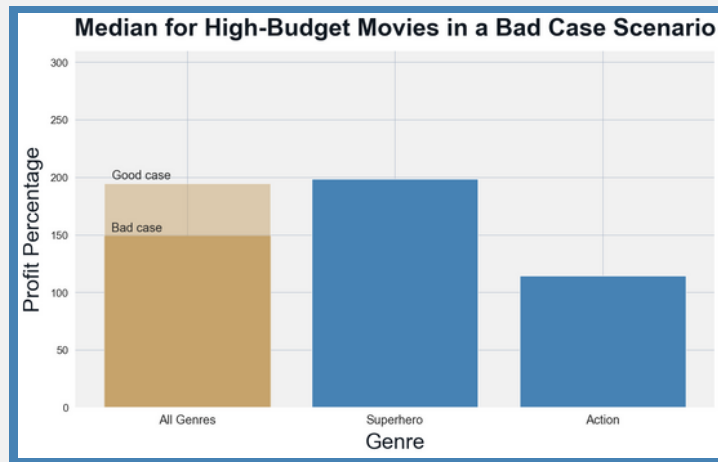
High-budget movies

Things are different for the high budget movies;

First of all, the genres considered are just 9. This highlights that the \$100M mark is hard to reach by movies. Below there's the median of the high-budget movies by genre.



In contrast to mid and low-budget counterparts, **high-budget movies have a smaller variance in profit percentages**. The graphs below illustrate bad and good case scenarios for the two most influential genres.



At first glance, the "Action" genre underperforms compared to the overall group. But this falls short when you consider a good case scenario for the "Action" genre; it's very similar to that of the group.

The "Superhero" genre's bad-case scenario surpasses the group's good-case scenario. Yet, the difference is minimal (just 2.4%). **This hints that, probably, the name Marvel or DC is the only thing that makes the superhero genre earn a little more than others.**

Conclusion:

A Nuanced Relationship

There are 2 answers to the question, “Does The Genre Impact a Movie’s Revenue?”

The short answer: No.

The long answer: It depends. Not all genres influence profit percentage, and not all genres influence it in the same way.

In the low-budget movies (<\$20M) the “Horror” and “Supernatural” genres seem to dominate. Their bad case scenario is still double the profit percentage of the good case for whole genres. This is probably caused by two factors:

- **Cost-Effectiveness.** Low-budget films typically cannot afford A-list actors or settings. Horror and Supernatural can leverage suspenseful storytelling, practical effects, and minimal sets, reducing production costs while maintaining audience engagement.
- **Broad Audience Appeal.** Horror and Supernatural genres have a broad appeal. Moreover, fans of the Horror and Supernatural genres often exhibit strong genre loyalty. This is vital to improve viewership.

Further research is needed to explain why the “Sport” genre performs half as well as the whole group. One of the reasons may be that viewers might prefer watching real sports events rather than low-budget sports movies, especially when the latter may lack production quality.

In the mid-budget movies (\$20M to \$100M) the stars are the “Supernatural” and “Buddy” genres. (But it’s worth mentioning that the “Horror” genres narrowly failed our bad-case/good-case test.)

The reasons for this Supernatural success are explained above.

- About the Buddy genre. The reason for his success is probably the very nature of this genre; The nature of "Buddy" movies often focuses on character development and interpersonal relationships rather than action sequences or costly visual effects. This can also allow studios to bring in popular stars to increase viewership.

As for the Sports genre in low-budget films, industry expertise is highly needed to explain why the Drama genre performs so badly in the mid-budget group. The Drama genre gets at least 58.6% less profit percentage compared to the overall group in a typical bad-case scenario.

On the other hand, high-budget movies tell a different story. The only genre that seems to influence profit percentage is the “Superhero.” And even that, the effect is probably due more to the established fan bases (like Marvel or DC) and less to the genre itself.

Again, industry expertise is necessary for more thorough explanations. But one probable reason is that high-budget movies, by nature, involve substantial financial investments. As a result, filmmakers and studios may prioritize diverse genres to cater to a broader audience. Moreover, if that movie is not a Blockbuster, the market may reach a point of saturation.

Appendix:

Dive Into The Nitty-Gritty

Data Source

The data analyzed are scraped from nearly 5,000 pages of Wikipedia. They contain all the American films released from 2004 to 2023.

[On this link, there's the Jupyter Notebook with all the code I used to scrape the data.](#)

Data Cleaning

As is often the case with web-scraped data, the initial dataset was pretty messy. Before analyzing it, it was necessary a phase of data cleaning.

[In this Jupyter Notebook, you find the exact process of data cleaning.](#)

Tests

For those curious about the inner workings of statistical tests – the ANOVA test, Pearson correlation, and confidence intervals...

[In this Jupyter Notebook, you'll find all that and all the code I used to bring this analysis to life.](#)

NOTE: Each Jupyter Notebook comes with documentation so the code is easier to navigate.

A special thanks goes to [Wikipedia](#). Without their contribution, there would be no dataset to analyze, no insights to glean, and certainly nothing new to discover.

Moreover, my biggest THANK YOU! goes to the three people who shaped my skills the most:

- [Jake VanderPlas](#).
- [Wes McKinney](#).
- [Claus O. Wilke](#).

Thank You!

Contacts

- Email: matteo.ladislai@hotmail.com
- LinkedIn: <https://www.linkedin.com/in/matteo-ladislai/>