L. A. M. E.

Laila Gamaleldin · Andrew Reed · Matthew Alexander · Emery Shelley

Seasons of Success: Effects of Seasonality on Film Profitability

INTRODUCTION

Over the years, the terms “summer blockbuster” and “Oscar season” have become part of the vocabulary of even non-movie-buffs. So, our group was interested in determining whether these tropes may actually hold water in the data. Specifically, we set out to find what role season plays in a movie’s success or failure.

Our data came from two Kaggle repositories, and contained film data from IMDB and Rotten Tomatoes from the past 60 years. After joining the datasets on title and removing entries with null or zero values, we were left with about 1900 data points, each with release_date, revenue, and budget (IMDB), and tomatometer_rating (Rotten Tomatoes).

HYPOTHESIS

We were interested in testing the hypothesis that (a) the success of films differs based on the season in which it is released (b) that these differences have grown over time, specifically that the the success of films belonging to one of the classic genre-season combos has become more predominant over time, specifically that

- Summer action, comedy, and kids films
- Winter/fall dramas, and
- Fall horror films have become more profitable

DATA VISUALIZATION

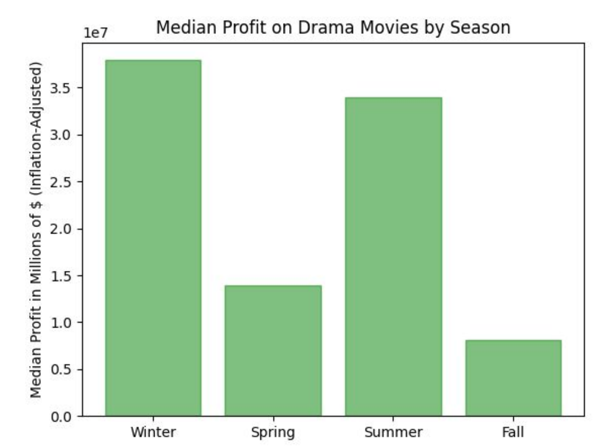

FIGURE 3 (below): *Median inflation-adjusted profit for action films, split by season*

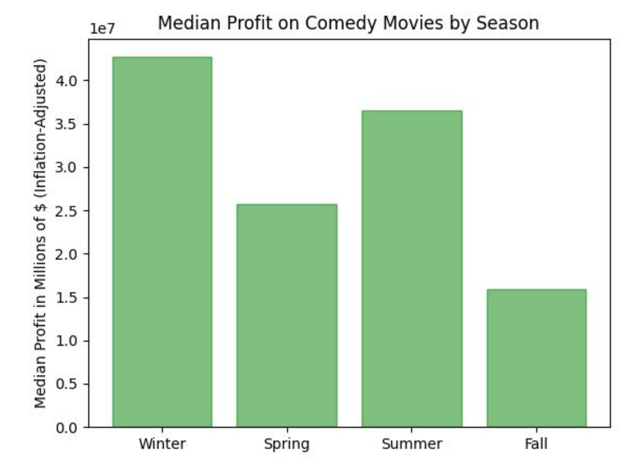

FIGURE 5 (below): *Median inflation-adjusted profit for kids films, split by season*

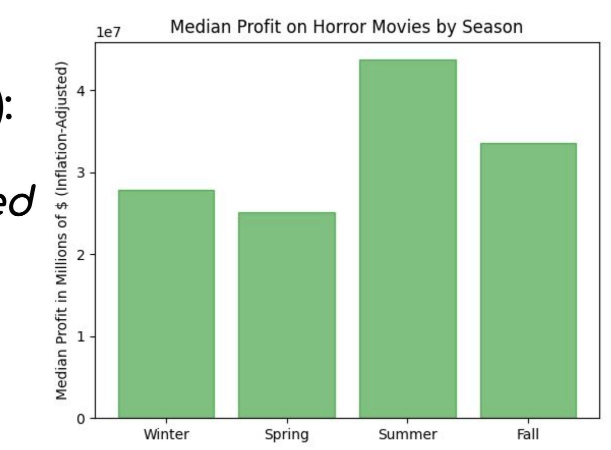

FIGURE 2 (above): *Median inflation-adjusted profit for drama films, split by season-*

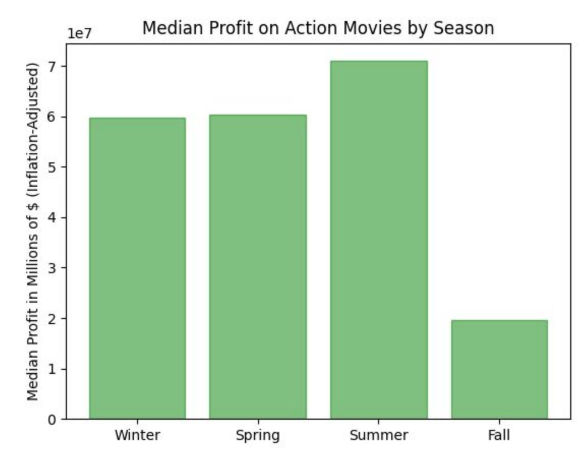

FIGURE 4 (above): *Median inflation-adjusted profit for comedy films, split by season*

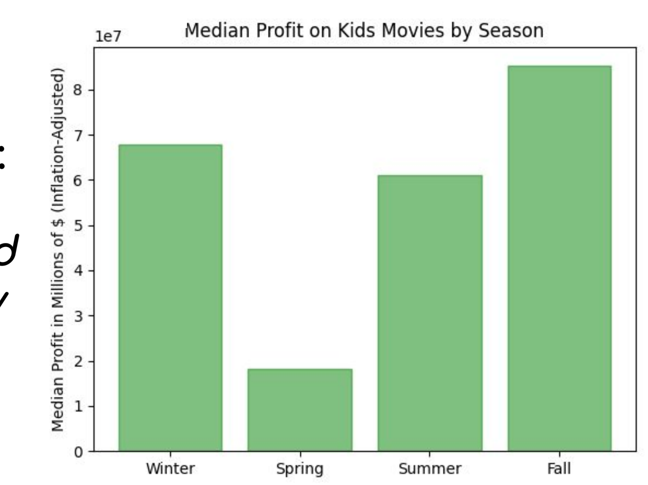

FIGURE 6 (above): *Median inflation-adjusted profit for horror films, split by season*

CONCLUSIONS

- Over the entire population of films, our chi squared tests suggested that, over time, the distribution of profit by season has become less uniform (our p-values consistently decreased for each our tests)
- Additionally, the chi squared tests on the five genres suggest that the distribution of profit across season is non-uniform for drama, action, and kids films, and more uniform for comedy and horror films
- Given these results, we went ahead with our multiple regression to try to further examine the nature of these non-uniform distributions
- There is statistically significant evidence to support that, all else equal, over time summer-released action and kid movies showed increased profitability year to year on average.
- There was negative statistically significant evidence to support that, all else equal, winter-released dramas had decreased profits over time.
- We do not show statistically significant evidence to support increased profitability for fall-released dramas, fall horrors, and summer comedies (all other genre interaction terms show no statistical significance).
- Despite it not reaching p-value < .05, there are three additional interaction terms that are nearly significant and worth potential notice.
- Claims only suggest possible correlation not causation. Statistical significance does not guarantee correlation and non-statistical significance does not guarantee no correlation

METHODOLOGY

In order to approach this hypothesis correctly, we made several adjustments to our data. This included adding dummy variables for each season, as well as for each genre. We also added an inflation-adjusted profit calculation in an effort to measure and compare success across different years.

Ultimately, our columns of interest included original_title, budget, revenue, year_released, season dummy variables (isWinter, isSpring, isSummer, isFall), genre dummy variables (isDrama, isAction, isComedy, isHorror, isKids), tomatometer_rating, and inflationAdjustedProfit (created using historical CPI data).

Five Year Rolling Average Profit (Inflation Adjusted)

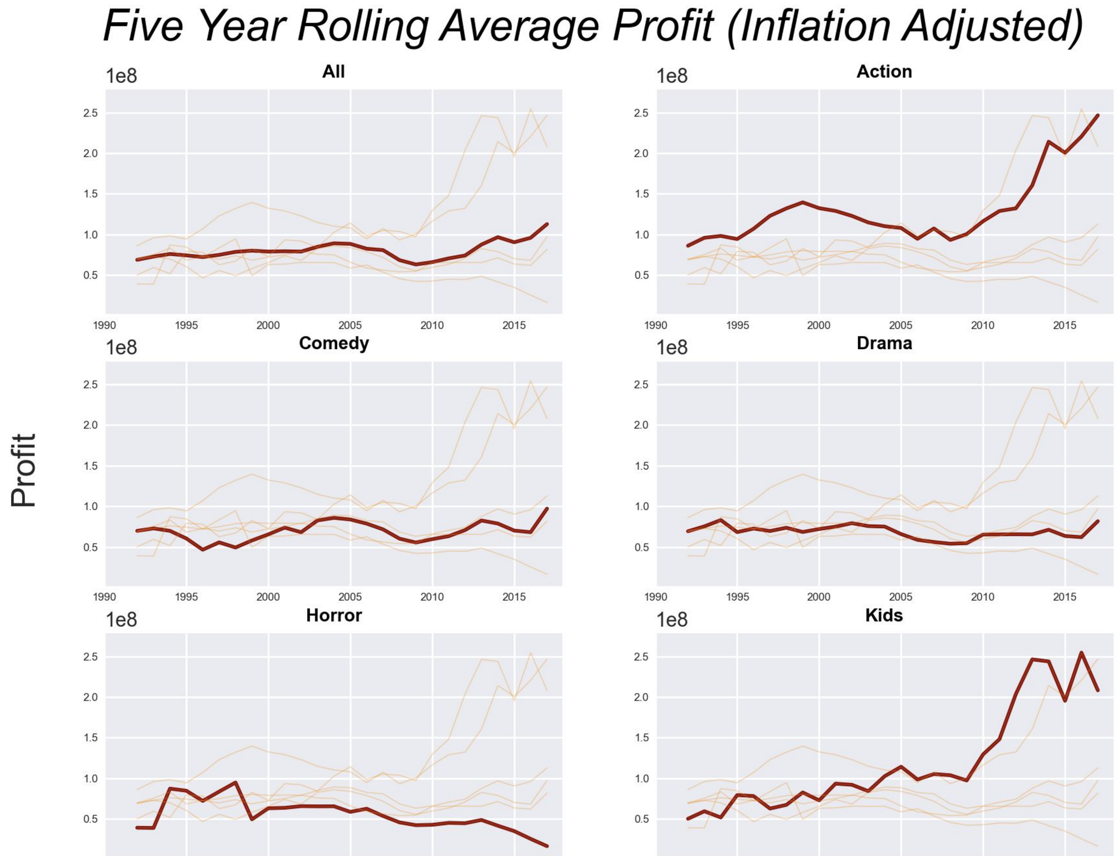

FIGURE 1: *The Five Year Rolling Average Profit (Inflation Adjusted)*

RESULTS

CHI-SQUARED TESTS

We ran eight chi-squared tests on our dataset. All of these tests split the data on season and whether the film turned a profit. The first three (not represented in figure 7) were on the entire dataset split into three time periods (pre 1980 1980-2000, and 2000 onward). Of these three tests, the first two yielded statistically insignificant p-values and the final test returned a p-value of **0.001**. Therefore, it would appear that the distribution of film profitability against season has become less uniform over time. An additional five chi-squared tests were run on films from our five chosen genres. Of these tests, drama, action, and kids films resulted in statistically significant p-values (**0.003**, **0.02**, **0.012** respectively). Comedy and horror films resulted in statistically insignificant p-values. (These tests used a significance level of 5%)

FIGURE 7 (below): *Contingency tables for chi-squared tests*

Observed Values	Winter	Spring	Summer	Fall
Action				
Inflation Adj Profit > 0	86	102	149	90
Inflation Adj Profit ≤ 0	33	27	44	50
Drama				
Inflation Adj Profit > 0	155	123	170	169
Inflation Adj Profit ≤ 0	69	62	52	107
Comedy				
Inflation Adj Profit > 0	123	118	172	116
Inflation Adj Profit ≤ 0	42	42	46	64
Horror				
Inflation Adj Profit > 0	40	43	36	59
Inflation Adj Profit ≤ 0	13	9	14	11
Kids				
Inflation Adj Profit > 0	22	20	50	25
Inflation Adj Profit ≤ 0	5	11	15	13

MULTIPLE REGRESSION

Drama Films post-1960 (n=906)

Variable	Coefficient	p-Value
Winter-Year Interaction	-2.145e+06	0.002**
Spring-Year Interaction	1.976e+05	0.804
Summer-Year Interaction	4.709e+04	0.945
Fall-Year Interaction	3.789e+05	0.088†

Action Films post-1960 (n=581)

Variable	Coefficient	p-Value
Winter-Year Interaction	-1.857e+06	0.714
Spring-Year Interaction	-2.813e+06	0.536
Summer-Year Interaction	7.872e+06	0.044*
Fall-Year Interaction	-1.677e+06	0.691

Comedy Films post-1960 (n=723)

Variable	Coefficient	p-Value
Winter-Year Interaction	-1.222e+06	0.161
Spring-Year Interaction	8.439e+05	0.389
Summer-Year Interaction	5.765e+05	0.474
Fall-Year Interaction	-1.005e+06	0.265

Horror Films post-1960 (n=225)

Variable	Coefficient	p-Value
Winter-Year Interaction	1.394e+06	0.178
Spring-Year Interaction	-6.903e+05	0.490
Summer-Year Interaction	-1.675e+06	0.091†
Fall-Year Interaction	9.109e+05	0.348

Kid Films post-1960 (n=171)

Variable	Coefficient	p-Value
Winter-Year Interaction	-1.823e+06	0.517
Spring-Year Interaction	5.071e+06	0.219
Summer-Year Interaction	4.716e+06	0.045*
Fall-Year Interaction	-4.596e+06	0.066†

† = Nearly Significant
* = Significant
** = Highly Significant

FIGURE 8: *The tables above state the results of our multiple regression for each of the genres we are examining, laying out the coefficients and p-values for each season-year interaction term.*

The multiple regression contained 11 independent variable. Year, budget, critic review (Rotten Tomato Score) were included so they could be controlled for. 4 season dummy variables and 4 season-year interaction terms were also included.

CHALLENGES/FURTHER WORK

- In our dataset, revenue is the total international box office gross, however, since season is dependent on hemisphere, some films may have been considered “summer” films when it wasn’t really summer in every nation it was screened
- Determining what time frame to run the regression on could be seen as p-value hacking (used chi-squared as a guide).
- Sometimes films belonged to more than one genre
- One possible follow up could involve creating a model to predict profit, in part, based on season
- At times, our project was limited by the small number of films in given time periods. As more films are released in the next 20 years, it may be worthwhile to go back and redo this experiment with more data

SIGNIFICANCE

A clear analysis and depiction of the effect of seasonality could have great impact on the film industry, possibly factoring into decisions about ideal release dates for certain films.

A deeper dive into changes in the most profitable genre-season combinations over time could potentially allow studios and production agencies to see, capitalize on, and possibly even influence these trends early on.