

An Analysis of Movie Success through Profits and Reviews 1929-2016

Mitchell Joram

7/2/2023

Introduction

Movies are the cornerstone of entertainment. We've all seen good movies, bad movies, and movies so bad they're good. A question I've always asked is what factors contribute to a movie's success?

I will be conducting an analysis on a movie data set to determine what types of movies could be the most successful by profit and reviews.

The Data

I located this data set from Github user ribeiromatheus.

The data is sourced from IMBD featuring movies released from 1927 to 2016. The data used in this analysis is a subset of movies made only in the United States.

Data Attributes and Descriptions

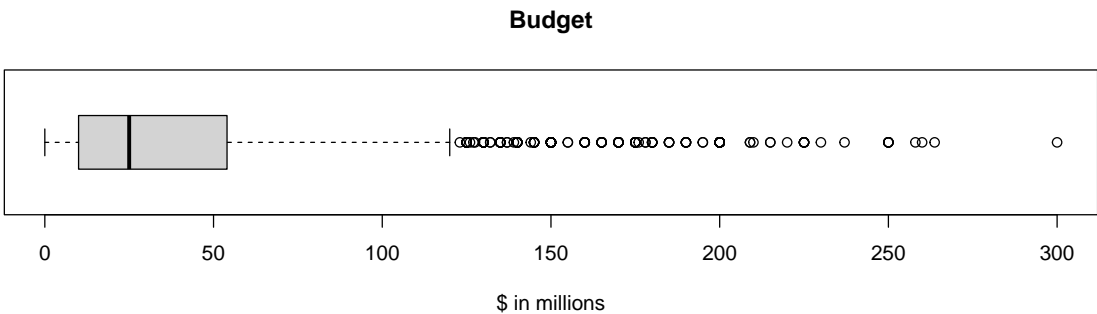
Attribute	Data Type	Description
movie_title	String	Title of movie.
title_year	Integer	Year of movie release.
director_name	String	Director of movie.
genres	String	Genre(s) of movie. Movies with multiple genres are separated by ' '. '
duration	Integer	Duration of movie in minutes.
budget	Float	Amount of money to produce movie.
gross	Float	Amount of money movie earned in box office.
language	String	Language the movie is originally scripted in.
country	String	Country movie was produced in. Analyzing only "USA".
content_rating	String	Movie rating given by the Motion Picture Association, The Motion Picture Production Code, or other.
color	String	The movie was filmed in "Black and White" or in "Color".
movie_imdb_link	String	IMDB URL link.
num_user_for_reviews	Integer	Number of IMDB user reviews.
imdb_score	Float	IMBD score out of 10 averaged across all user reviews.

Budget, Gross, and Profit

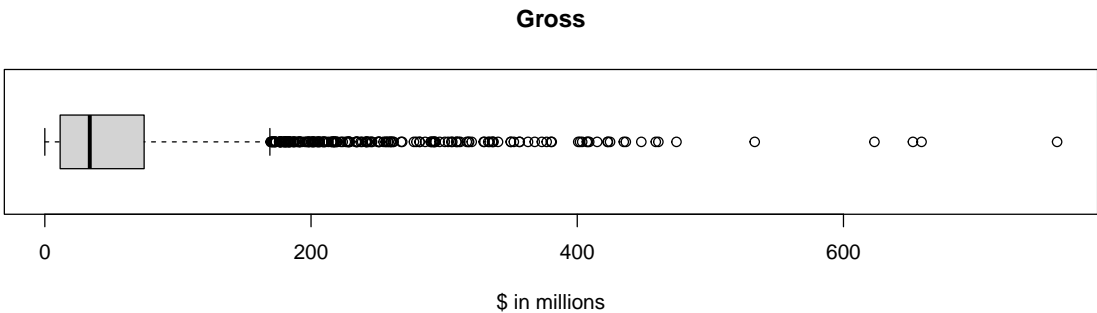
One way to measure a movie’s success is it’s ability to make a profit. A movie that sells a lot of tickets must mean something of value.

Obviously movies are expensive to make. Movie studios provide productions with a budget to produce the movie. It is the movie’s capitalist goal to make enough gross revenue through the box office to pay off the studio for its budget and a profit. For there to be a positive profit, the gross revenue must exceed the budget costs.

$\text{Profit} = \text{Gross} - \text{Budget}$. $\text{Profit} > 0$ if $\text{Gross} > \text{Budget}$.



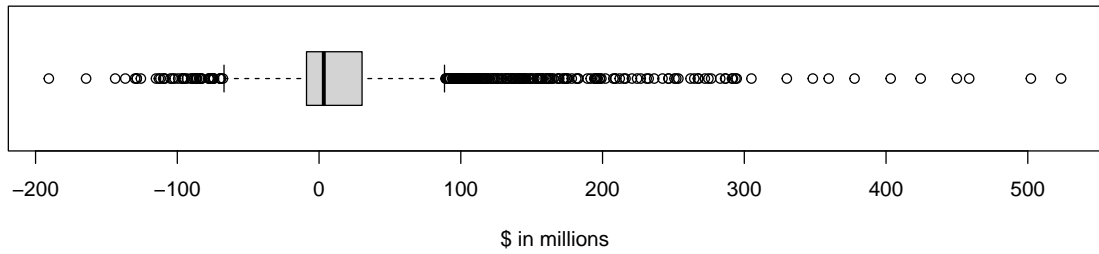
	Budget in millions
Min	0.000218
First Quartile	10.000000
Median	25.000000
Third Quartile	54.000000
Maximum	120.000000
Mean	39.791653



	Gross in millions
Min	0.000703
First Quartile	11.501093
Median	33.771174
Third Quartile	74.608545
Maximum	169.076745

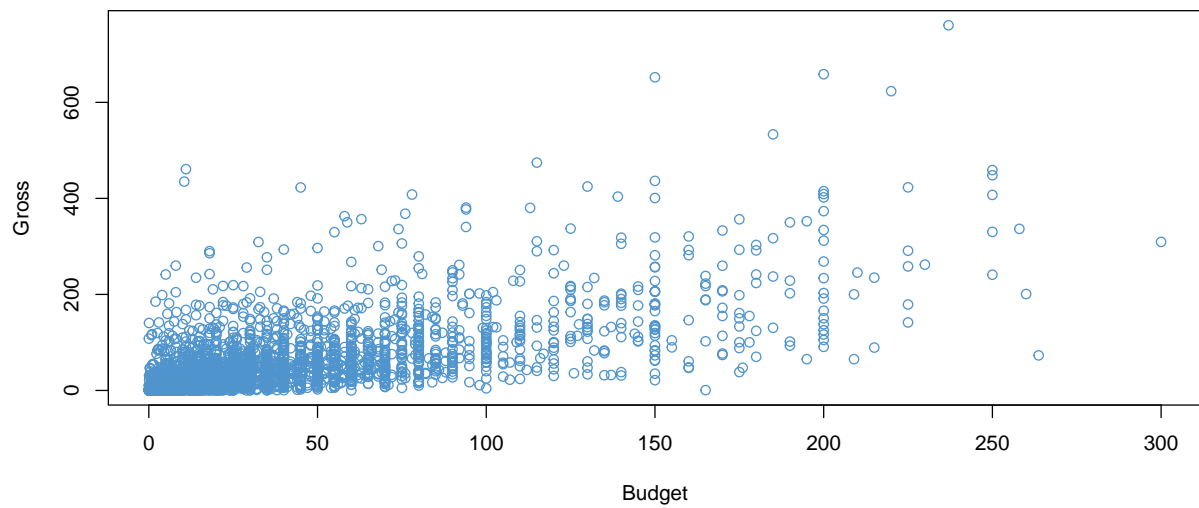
	Gross in millions
Mean	57.117131

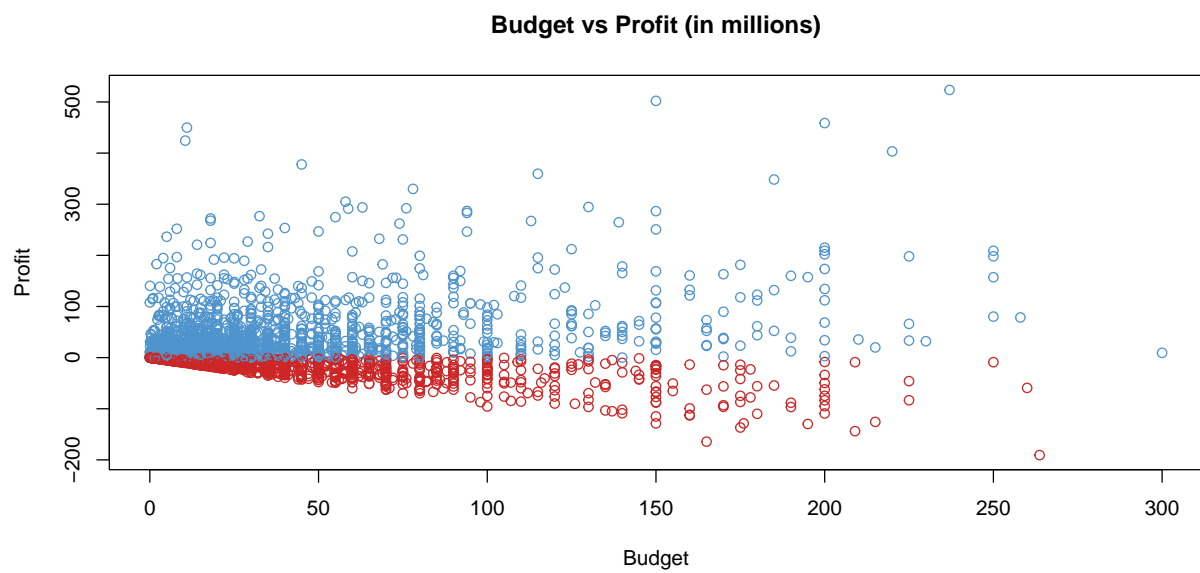
Profit



	Profit in millions
Min	-67.047980
First Quartile	-8.851518
Median	3.341469
Third Quartile	30.341670
Maximum	88.495848
Mean	17.325478

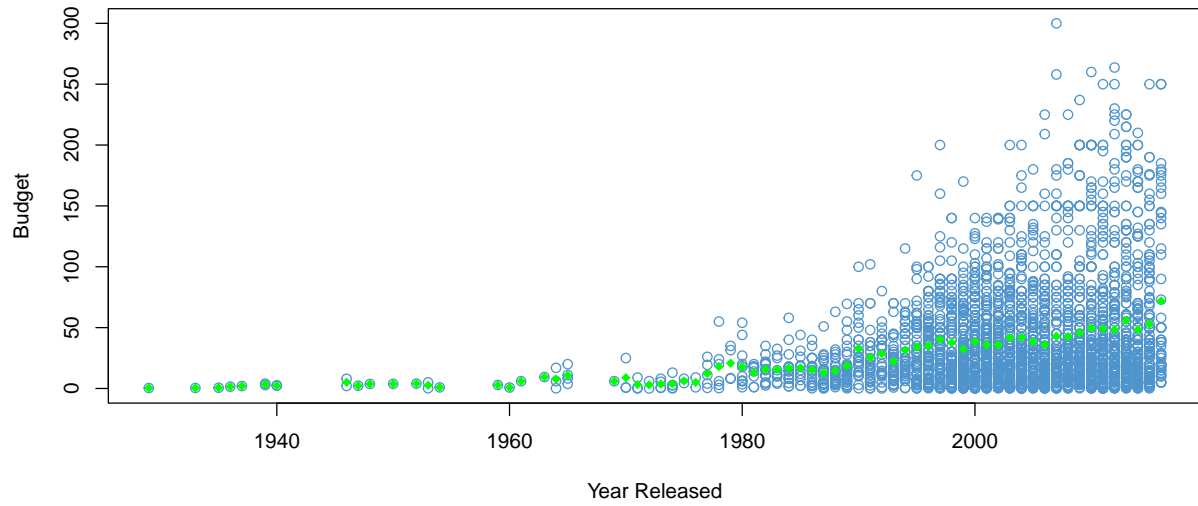
Budget vs Gross (in millions)



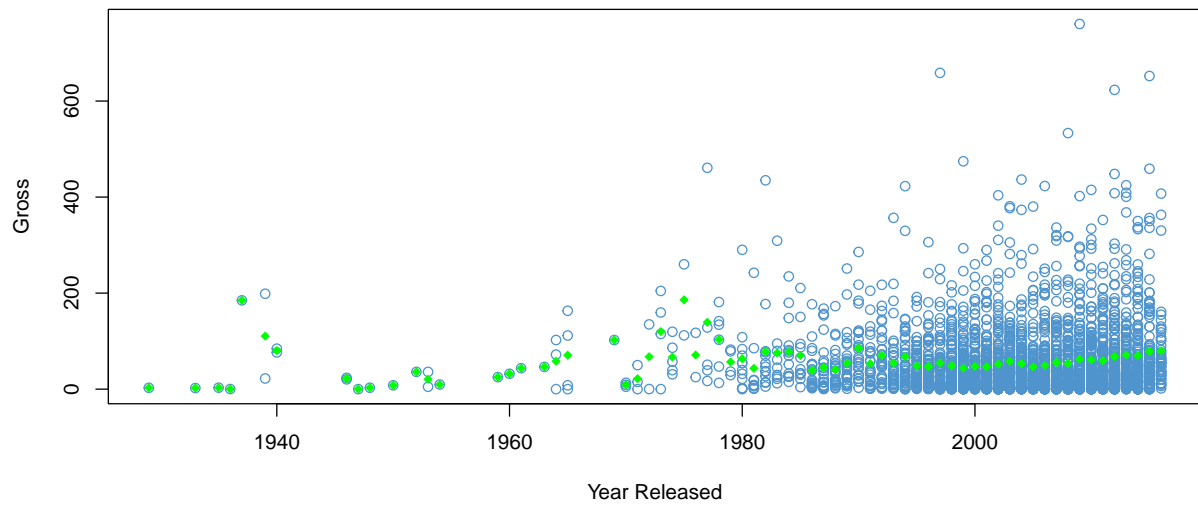


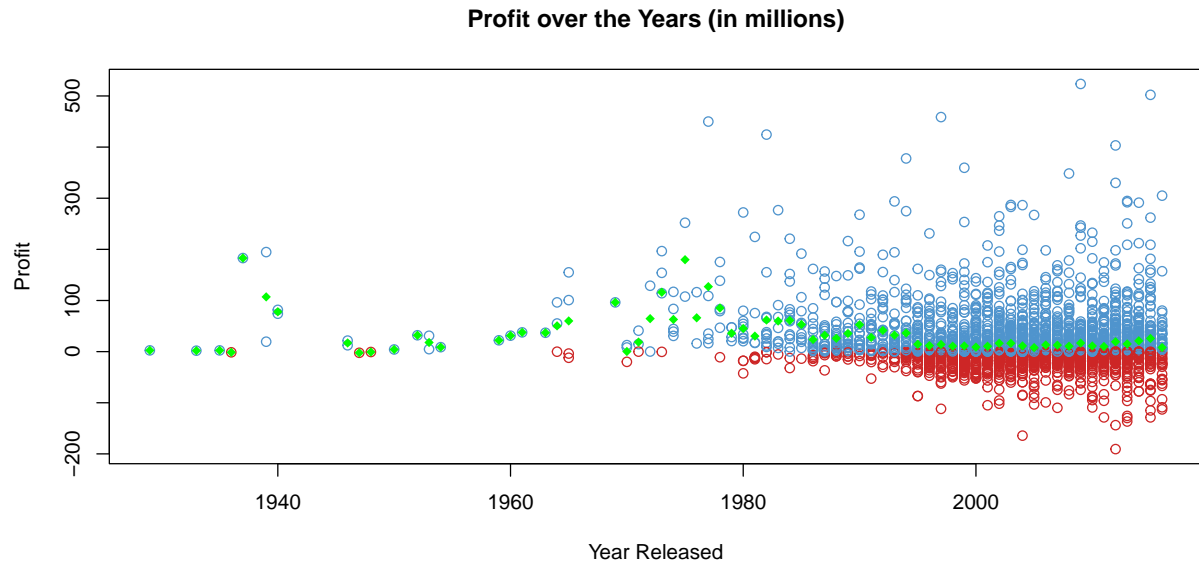
Over the Years

Budget over the Years (in millions)



Gross over the Years (in millions)





The average budget every year has increased since 1929. Average gross increased until it fairly stagnated in 1980, but appears to be slightly above the average year budget. This is supported by the average profit being positive throughout the figure, except in 1948 with -\$744000.00 in average profits.

Top 10 Movies by Budget

Rank	Movie Title	Budget
1	Pirates of the Caribbean: At World's End	300000000
2	John Carter	263700000
3	Tangled	260000000
4	Spider-Man 3	258000000
5	Pirates of the Caribbean: On Stranger Tides	250000000
6	The Dark Knight Rises	250000000
7	Avengers: Age of Ultron	250000000
8	Batman v Superman: Dawn of Justice	250000000
9	Captain America: Civil War	250000000
10	Avatar	237000000

Top 10 Movies by Gross

Rank	Movie Title	Gross
1	Avatar	760505847
2	Titanic	658672302
3	Jurassic World	652177271
4	The Avengers	623279547
5	The Dark Knight	533316061
6	Star Wars: Episode I - The Phantom Menace	474544677
7	Star Wars: Episode IV - A New Hope	460935665
8	Avengers: Age of Ultron	458991599
9	The Dark Knight Rises	448130642
10	Shrek 2	436471036

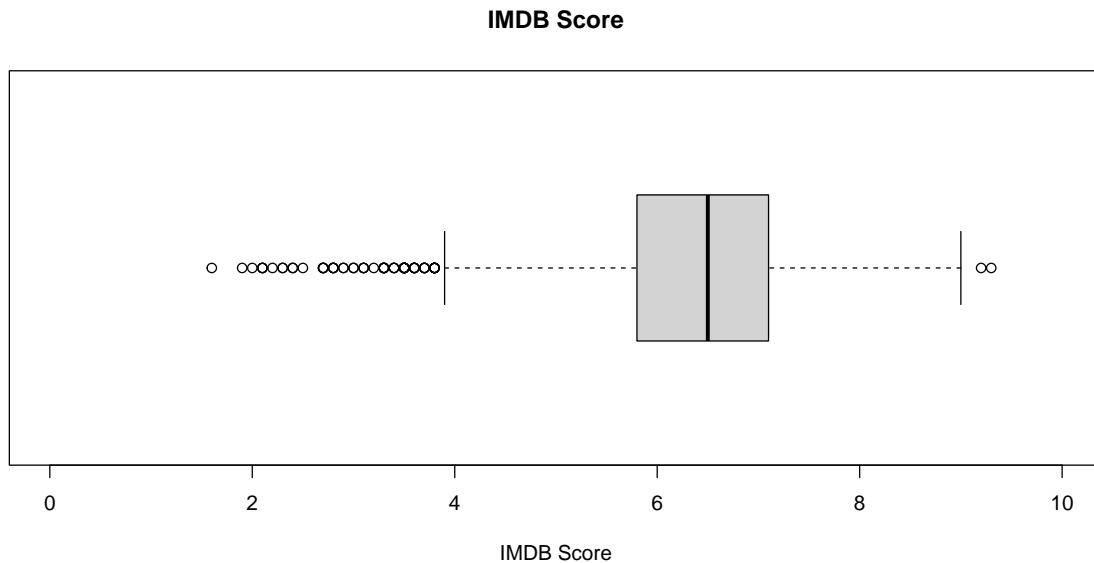
Top 10 Movies by Profit

Rank	Movie Title	Profit
1	Avatar	523505847
2	Jurassic World	502177271
3	Titanic	458672302
4	Star Wars: Episode IV - A New Hope	449935665
5	E.T. the Extra-Terrestrial	424449459
6	The Avengers	403279547
7	The Lion King	377783777
8	Star Wars: Episode I - The Phantom Menace	359544677
9	The Dark Knight	348316061
10	The Hunger Games	329999255

IMDB Score

Another way to determine a successful movie is through it's reviews. A good movie must have good reviews.

IMDB holds one of the greatest databases on movie reviews. Each movie holds an "IMDB score" which determines how much a movie was enjoyed by the users out of 10 points.



	IMDB Score
Min	3.900000
First Quartile	5.800000
Median	6.500000
Third Quartile	7.100000
Maximum	9.000000
Mean	6.384742

Top 10 Movies

Rank	Movie Title	IMDB Score
1	The Shawshank Redemption	9.3
2	The Godfather	9.2
3	The Godfather: Part II	9.0
4	The Dark Knight	9.0
5	Schindler's List	8.9
6	Pulp Fiction	8.9
7	The Lord of the Rings: The Return of the King	8.9
8	Star Wars: Episode V - The Empire Strikes Back	8.8
9	Forrest Gump	8.8
10	Fight Club	8.8

Directors

It is the Film Director's job to oversee the film production. They are responsible for managing the actors and crew, help them interpret the script, make creative choices, and ensure the visual storytelling aligns with the desired narrative.

Behind every quality movie is a great director. Movies that are well done had the adequate guidance from their directors.

Top 10 Directors by Total Budget

Rank	Director	Total Budget
1	Michael Bay	1461000000
2	Steven Spielberg	1417900870
3	Ridley Scott	1139000000
4	Gore Verbinski	1071000000
5	Robert Zemeckis	1070000000
6	Christopher Nolan	1005000000
7	Ron Howard	954000000
8	Tim Burton	947000000
9	Bryan Singer	938000000
10	Martin Scorsese	914300000

Top 10 Directors by Total Gross

Rank	Director	Total. Gross
1	Steven Spielberg	4014061704
2	Michael Bay	2195443511
3	James Cameron	1909725910
4	Christopher Nolan	1813227576
5	George Lucas	1741418480
6	Robert Zemeckis	1619309108
7	Tim Burton	1557078534
8	Sam Raimi	1443167519
9	Clint Eastwood	1378321100
10	Francis Lawrence	1358501971

Top 10 Directors by Total Profit

Rank	Director	Total Profit
1	Steven Spielberg	2596160834
2	George Lucas	1386641480
3	James Cameron	1167725910
4	Christopher Nolan	808227576
5	Francis Lawrence	755501971
6	Michael Bay	734443511
7	Peter Jackson	664837575
8	Jay Roach	661636412
9	Clint Eastwood	610221100
10	Tim Burton	610078534

Top 10 Directors by Avg. Budget

Rank	Director	Avg. Budget
1	Nathan Greno	260000000
2	Dan Scanlon	200000000
3	Lee Unkrich	200000000
4	Mark Andrews	185000000
5	David Yates	180000000
6	Robert Stromberg	180000000
7	Andrew Stanton	179233333
8	Carl Rinsch	175000000
9	Joss Whedon	170000000
10	Rupert Sanders	170000000

Top 10 Directors by Avg. Gross

Rank	Director	Avg. Gross
1	Lee Unkrich	414984497
2	Chris Buck	400736600
3	Joss Whedon	369202360
4	Tim Miller	363024263
5	George Lucas	348283696
6	Kyle Balda	336029560
7	Colin Trevorrow	328092532
8	James Cameron	318287652
9	Pete Docter	313113780
10	Pierre Coffin	309775640

Top 10 Directors by Avg. Profit

Rank	Director	Avg. Profit
1	Tim Miller	305024263
2	George Lucas	277328296
3	Richard Marquand	276625409

Rank	Director	Avg. Profit
4	Irvin Kershner	272158751
5	Kyle Balda	262029560
6	Colin Trevorrow	252717532
7	Chris Buck	250736600
8	Pierre Coffin	237275640
9	Lee Unkrich	214984497
10	Joss Whedon	199202360

Top 10 Directors by Average IMDB Score

Rank	Director	Avg. IMDB Score
1	Irvin Kershner	8.800
2	Charles Chaplin	8.600
3	Tony Kaye	8.600
4	Alfred Hitchcock	8.500
5	Damien Chazelle	8.500
6	Milos Forman	8.500
7	Ron Fricke	8.500
8	Christopher Nolan	8.425
9	Marius A. Markevicius	8.400
10	Richard Marquand	8.400

Genre

Every movie follows a particular theme. These themes are categorized into genres, from “History” to “Western” to “Sci-Fi”. Each genre taking the viewer on a different journey. The genre does not define the success of the film, but can still finds trends in movie genres.

Budget, Gross, and Profit

Genre	Avg. Budget	Avg. Gross	Avg. Profit
Action	70599708	85998943	15399235
Adventure	81035488	107766059	26730571
Animation	85607853	120847003	35239151
Biography	29319224	46384707	17065482
Comedy	35271599	54884655	19613056
Crime	34008911	45020472	11011561
Documentary	4570617	17377025	12806408
Drama	29803898	42802240	12998342
Family	65452201	95365815	29913614
Fantasy	66304956	91030456	24725500
History	47916590	56194121	8277531
Horror	22448905	37814598	15365692
Musical	36945783	64960320	28014537
Music	28677295	51853277	23175982
Mystery	37650059	50438513	12788454
Romance	31500906	49493715	17992808

Genre	Avg. Budget	Avg. Gross	Avg. Profit
Sci-Fi	71593625	94917500	23323875
Sport	34382082	47211301	12829219
Thriller	42575055	53584836	11009781
War	47937141	62302690	14365549
Western	47562475	52571093	5008618

Some genres like Adventure, and Animation just tend to make more gross revenue by average than Documentary. Yet, Adventure, and Animation cost way more to produce by average with Documentary costing the least. The Western and History genres are the least profitable by average.

IMDB Score

Genre	Avg. IMDB Score
Action	6.237723
Adventure	6.412893
Animation	6.640491
Biography	7.103205
Comedy	6.133546
Crime	6.499461
Documentary	6.797674
Drama	6.716608
Family	6.192120
Fantasy	6.234390
History	7.134940
Horror	5.869435
Musical	6.561905
Music	6.422772
Mystery	6.404861
Romance	6.343427
Sci-Fi	6.299483
Sport	6.541600
Thriller	6.316547
War	6.990588
Western	6.689583

This supports the claim that genre does not determine the success of a film. The vast majority of all movies sit in the score range of 6-7. Although the genres Biography and History average above 7 and Horror averages below 6, the difference is so small compared to the other genres that this is insufficient evidence of how the genre can determine a film's overall review success.

If the difference in scores across multiple genres were points apart, then we might be able to determine if genre affect overall success.

Rating

Today every movie is given a rating by the Motion Picture Association. In the past, it was done by The Motion Picture Production Code, or some other association. Movies back then were only given the ratings of "Approved" or "Passed, and"Denied" for public release. Nowadays, there are different ratings depending on

the content the film captures. More mature content is given rating's that suits the general age of the target audience.

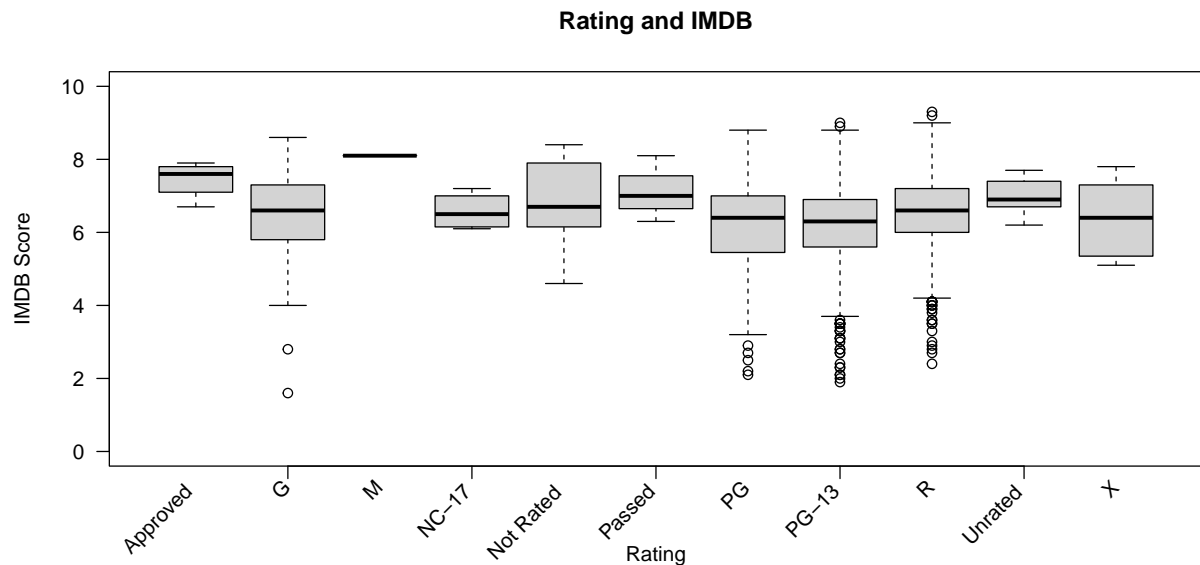
Rating	Description
G	General audiences - All ages admitted.
PG	Parental guidance suggested - Some material may not be suitable for children.
PG-13	Parents strongly cautioned - Some material may be inappropriate for children under 13.
M	Suggested for mature audiences - Parental discretion advised.
R	Restricted – Under 17 requires accompanying parent or adult guardian.
X	No one under 17 admitted. Later associated with pornography, but this data set contains no pornographic films.
NC-17	No one under 17 admitted. Replaced the X rating in 1990.
Approved	Approved for release by The Motion Picture Production Code.
Passed	Approved for release.
Unrated	Film has not been submitted for a rating or is an uncut version of a film that was submitted.
Not Rated	Film has not been submitted for a rating or is an uncut version of a film that was submitted.

- Approved, Passed, M, and X are rating no longer given to movies by any film association.

Budget, Gross, and Profit

Rating	Avg. Budget	Avg. Gross	Avg. Profit
Approved	5035444	55644039.8	50608595.3
G	55436014	98380290.0	42944275.9
M	6000000	102308900.0	96308900.0
NC-17	927500	639489.2	-288010.8
Not Rated	1604203	5032593.8	3428391.3
PG	52748204	80945461.0	28197256.7
PG-13	53011031	71565386.3	18554355.7
Passed	2315928	11003537.3	8687609.0
R	25097755	36286935.0	11189179.6
Unrated	1799324	8291790.0	6492465.5

IMDB Score



	Approved	G	M	NC-17	Not Rated	Passed	PG	PG-13	R	Unrated	X
Min	6.700000	4.000000	8.1		6.100	4.600000	6.300000	3.200000	3.700000	4.200000	6.200000
First Quar- tile	7.100000	5.800000	8.1		6.150	6.150000	6.650000	5.450000	5.600000	6.000000	6.700000
Median	7.600000	6.600000	8.1		6.700000	7.000000	6.400000	6.300000	6.600000	6.900000	6.400
Third Quar- tile	7.800000	7.300000	8.1	6.500	7.000	7.900000	7.550000	7.000000	6.900000	7.200000	7.400000
Maximum	7.900000	8.600000	8.1		7.200	8.400000	8.100000	8.800000	8.800000	9.000000	7.700000
Mean	7.433333	6.507042	8.1		6.575	6.815789	6.233696	6.208829	7.133333	6.553027	7.015385

Just like with genre's, the rating can't determine a film's overall success. Films with a G rating have an unfair advantage to films with an R rating because they can be viewed by a much larger audience. Of course families with children are much more likely to go see a G rated film over an R rated, and therefore make more ticket sales and profit more on average.

Films with "Not Rated" or "Unrated" also have a disadvantage, because they are released after the original theater cut and draw in a subset of that audience. Therefore, these films tend to profit much less than their original released forerunners.

Color

Perhaps the largest evolution in film history was the ability to film in color. Originally films were filmed in black and white, unless they were animated, because cameras from that era just couldn't capture color. The movie *The Wizard of Oz* utilized a Technicolor's 3-strip process by filming through colored filters to produce a colored movie and created a magical feel for its audience. However, the process with expensive and many movies after were still shot in black and white until color capture was cheap and standard. Today, filming in

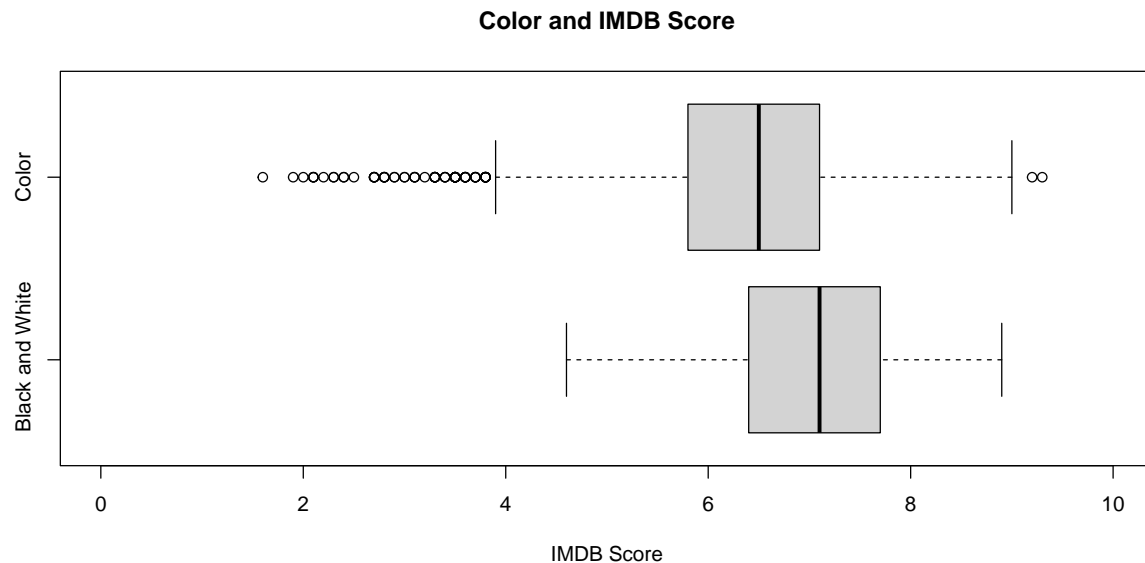
black and white a design choice, for example to capture a certain era or perhaps to show a lack of something within the film. The colors that are put on screen should capture an experience that is being presented to the audience.

Due to the technology change, it's preposterous to claim whether movies filmed in color are more successful than black and white or vice versa. It was really just more of a getting-with-the-times change, although you could argue that people would rather want to go see a colored film over a black and white film.

Budget, Gross, and Profit

Color	Avg. Budget	Avg. Gross	Avg. Profit
Color	40207608	57637396	17429789
Black and White	25847509	39676139	13828630

IMDB Score



	Black and White	Color
Min	4.600000	3.900000
First Quartile	6.450000	5.800000
Median	7.100000	6.500000
Third Quartile	7.700000	7.100000
Maximum	8.900000	9.000000
Mean	7.035635	6.36521

Conclusion

If we do follow the trends then:

The most profitable movie could be an Animated genre, G rated, and colored film directed by Steven Spielberg or Tim Miller.

The highest IMDB scoring movie could be a History genre, PG-13 rated, black and white film directed by Tony Kaye. * Directors with higher average IMDB scores Irvin Kershner and Charles Chaplin have passed away.