

# And the Oscar Goes To...

An Investigation into the Secret behind Critically Successful Movies

Vivian Fan and Amy Huang

## Introduction

What exactly does the phrase “Oscar bait” mean? Throughout the years, moviegoers have noticed patterns in a movie’s awards prospects. Do award voters favor a specific type of film? Our data analysis aims to find the perfect formula. We defined critical acclaim as a movie’s success at the following four awards: the Oscars, the Golden Globes, the BAFTA, and the SAG Awards.

We looked into several factors that could influence a movie’s critical acclaim:

- **Budget** - Does the amount of money invested in the movie affect its critical acclaim?
- **Domestic gross** - Does the amount of domestic box office revenue affect critical acclaim?
- **International gross** - Does the amount of the worldwide box office revenue affect critical acclaim?
- **Bechdel test\*** - Does the movie’s female representation affect critical acclaim?
- **Genres** - Do some genres have a better chance of critical acclaim than others?

\*The Bechdel test determines whether women are well-represented in a movie, which must meet all three criteria to pass the test:

1. At least 2 women are present in the movie, and
2. they must talk to each other
3. about something other than a man.

We used two existing datasets: the Bechdel dataset and the IMDB dataset. Additionally, we scrapped the IMDB website to acquire award nominations and wins. We conducted various analyses on the information we combined to determine which factors were relevant to a movie’s critical acclaim.

Our results show that:

- *Higher budget does not equate higher critical acclaim.*

However, a movie with a moderate budget between \$5 million and \$20 million *may* see its critical acclaim success rise slightly with higher budget investments.

- *Movies with higher domestic or international box office revenues do not necessarily receive higher critical acclaim.*

However, for movies that earn between \$6 million and \$60 million at the domestic box office and between \$10 million and \$200 million at the international box office, higher revenue is occasionally associated with slightly higher critical acclaim.

- *Movies that pass the Bechdel test do not necessarily receive higher critical acclaim.*
- *Films that belong to the following genres receive greater critical success: biography, drama, history, musical, and war.*
- *Films that belong to the following genres receive lower critical success: action, adventure, animation, comedy, science-fiction.*

## Hypothesis

We hypothesize a positive correlation between movie budget and critical acclaim up to a certain budget. Similarly, we hypothesize that both domestic and international gross have positive correlations with critical acclaim up to a certain point.

We hypothesize that films that pass the Bechdel test will have greater critical acclaim than films that do not. We also hypothesize that genres such as drama, biography, history, and war have greater critical success than genres such as comedy, horror, and mystery.

## Combining and Cleaning Data

We started out with two datasets that we scoured from the web: the [IMDB dataset](#) (titles.csv) and the [Bechdel dataset](#) (movies.csv), the latter of which is the intersection of two datasets taken from BechdelTest.com and The-Numbers.com. The Bechdel dataset contains important information, such as movie budget, domestic-, and international box office gross, adjusted for inflation, from a sample of movies released from 1970 to 2013. The IMDB dataset contains movie titles and genres. To measure critical acclaim, we applied web scraping techniques to gather award nominations and wins into a new file (awards.csv), which we then used to calculate the quantitative “score” for each movie’s critical acclaim (**Figure 1**).

	bafta_nom	bafta_win	gg_nom	gg_win	movie	oscar_nom	oscar_win	sag_nom	sag_win
0	0	0	0	0	tt1711425	0	0	0	0
1	0	0	0	0	tt1343727	0	0	0	0
2	6	2	6	1	tt2024544	6	3	3	1
3	0	0	0	0	tt1272878	0	0	0	0
4	0	0	0	0	tt0453562	0	0	0	0
5	0	0	0	0	tt1335975	0	0	0	0
6	0	0	0	0	tt1606378	0	0	0	0
7	0	0	0	0	tt2194499	0	0	0	0
8	0	0	0	0	tt1814621	0	0	0	0
9	0	0	0	0	tt1815862	0	0	0	0

**Figure 1:** Awards Table

We combined the IMDB dataset, the Bechdel dataset, and the awards information into our combined movie dataset (**Figure 2**). We eliminated columns that were irrelevant to our analysis and kept these:

- IMDB identifier
- Bechdel test result (pass/fail)
- Budget (adjusted for inflation)
- Domestic gross (adjusted for inflation)
- International gross (adjusted for inflation)
- Movie title
- Genres
- Total award nominations
- Total award wins
- Critical acclaim (weighted score)

For total award nominations, we added all the nominations that movie received from four prestigious film awards institutions: the Oscars, the BAFTAs, the Golden Globes, and the SAG Awards. For total award wins, we added all the wins for that movie.

imdb	binary	budget_2013\$	domgross_2013\$	intgross_2013\$	primaryTitle	runtimeMinutes	genres	total_wins	total_noms	weighted_score
tt1711425	FAIL	13000000	25682380.0	42195766.0	21 & Over	93	Comedy	0	0	0
tt1343727	PASS	45658735	13611086.0	41467257.0	Dredd	95	Action,Crime,Sci-Fi	0	0	0
tt2024544	FAIL	20000000	53107035.0	158607035.0	12 Years a Slave	134	Biography,Drama,History	6	15	44
tt1272878	FAIL	61000000	75612460.0	132493015.0	2 Guns	109	Action,Comedy,Crime	0	0	0
tt0453562	FAIL	40000000	95020213.0	95020213.0	42	128	Biography,Drama,Sport	0	0	0

**Figure 2.** Combined movie dataset. We added three additional columns for *total wins*, *total nominations*, and *weighted score*.

## Calculating Critical Acclaim

We define critical acclaim to be the weighted amount of prestigious award nominations and wins received by a movie. Our method to calculate the critical acclaim of a movie is as follows: each nod (nomination or win) a movie receives adds the following number of points to the movie's *weighted score*.

Oscar Nomination	+ 2
Oscar Win	+ 3
BAFTA / Golden Globe / SAG Award Nomination	+ 1 (each)
BAFTA / Golden Globe / SAG Award Win	+ 2 (each)

An Oscar nod is worth more than other awards simply because it is the most prestigious motion picture award.

All numbers are then added together to get the weighted score.

For example, if a movie has:

- 3 Oscar nominations
- 1 Oscar win
- 10 BAFTA / Golden Globe / SAG Award nominations
- 4 BAFTA / Golden Globe / SAG Award wins

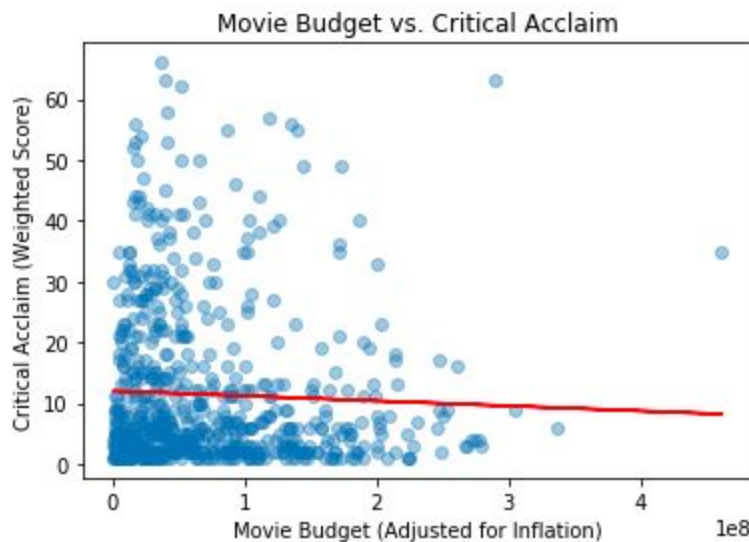
The movie would have  $(3 * 2) + (1 * 3) + (10 * 1) + (4 * 2) = \text{weighted score of 27}$ .

## Data Analysis

Our goal is to analyze how each factor influences a movie's critical success. First, we filtered out movies that made no money domestically and/or internationally, since they would never even have been released in theaters. We also filtered out movies that have a zero weighted score in critical acclaim, since most movies do not receive any accolades and will distort our analysis. We will only be looking at films that have a weighted score of at least 1.

Q1: Does *movie budget* influence critical acclaim?

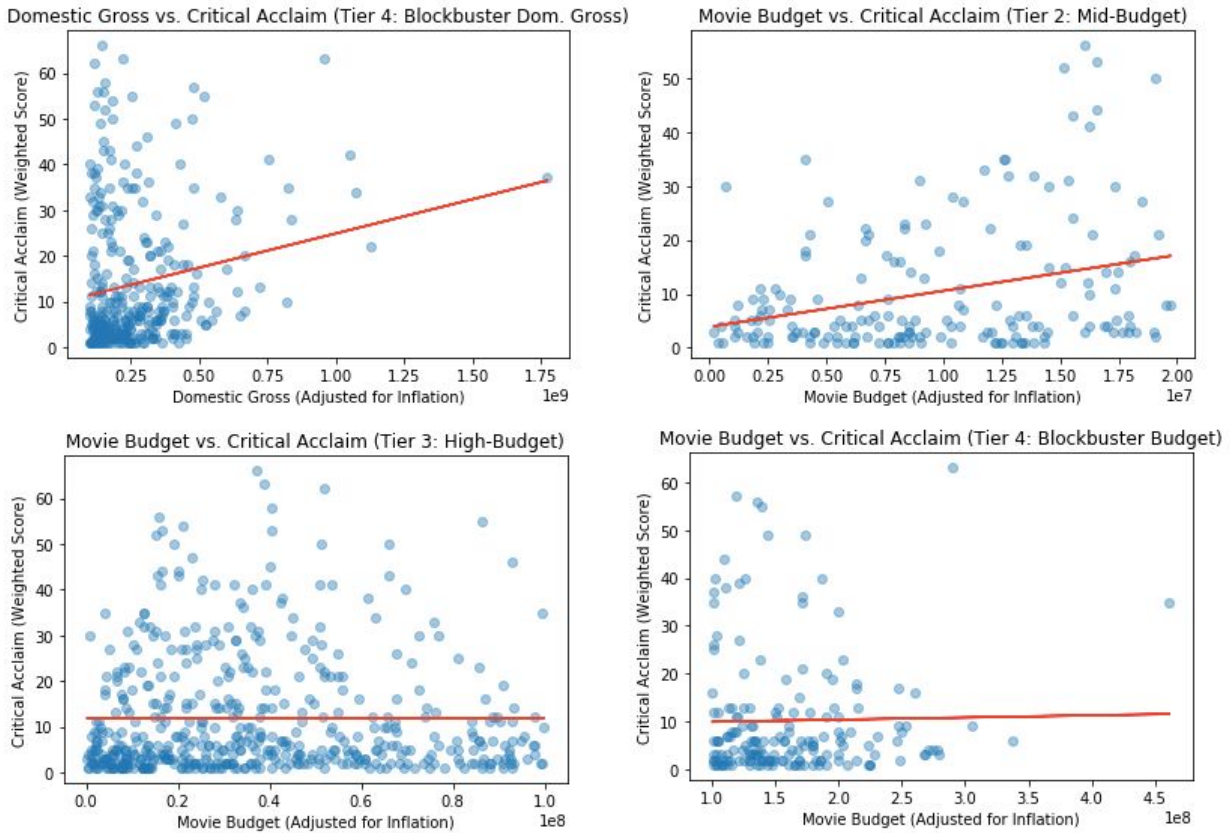
We compared all remaining movie budgets with their weighted score for critical acclaim and found that the coefficient of determination to be  $r^2 = 0.0017$  and slope to be  $-8.2603 \times 10^{-9}$  (**Figure 3**).



**Figure 3.**

About 0.17% of the variance in critical acclaim is dependent on the movie budget. Each additional million dollars in movie budget decreases critical acclaim score by around 0.00826. Therefore, the negative correlation between movie budget and critical acclaim is *very low*. We decided to divide the movie budgets into four tiers to better analyze the relationship:

- **Tier 1: Budget < \$5,000,000**
  - Includes low-budget, independent films.
- **Tier 2: \$5,000,000 < Budget < \$20,000,000**
  - Includes mid-budget films.
- **Tier 3: \$20,000,000 < Budget < \$100,000,000**
  - Includes higher budget films.
- **Tier 4: Budget > \$100,000,000**
  - Includes extremely high-budget, blockbuster films.



**Figure 4.**

For Tier 1 films,  $r^2 = 0.0256$ , and slope =  $8.8428 \times 10^{-7}$  (top left).

For Tier 2 films,  $r^2 = 0.0930$ , and slope =  $6.6517 \times 10^{-7}$  (top right).

For Tier 3 films,  $r^2 = 1.3113 \times 10^{-6}$ , and slope =  $5.5003 \times 10^{-10}$  (bottom left).

For Tier 4 films,  $r^2 = 0.0004$ , and slope =  $4.5132 \times 10^{-9}$  (bottom right).

Only Tier 2 films demonstrate a moderately weak, positive correlation between budget and critical acclaim. Each additional million dollars in movie budget increases critical acclaim score by around 0.665. Around 9.30% of the variance in critical acclaim is dependent on the movie budget.

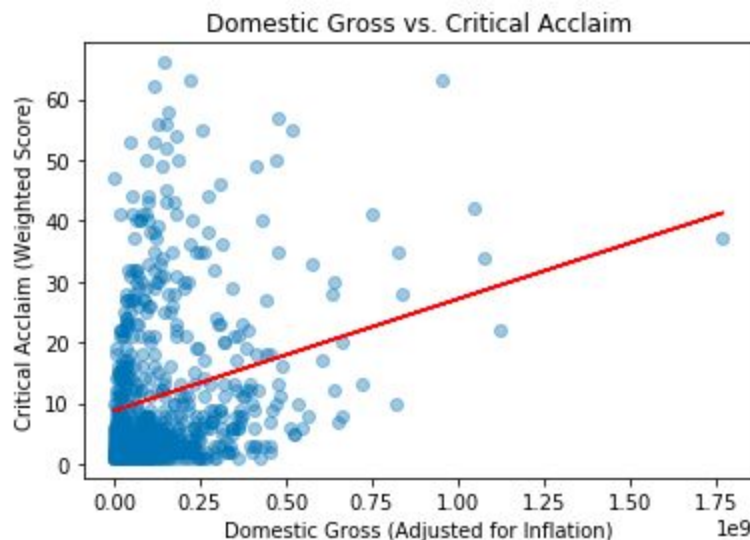
For Tier 1, 3, and 4, films, there is close to no correlation between budget and critical acclaim. For Tier 1 films, each additional million dollars in movie budget increases critical acclaim score by around 0.884. Around 2.56% of the variance in critical acclaim is dependent on the movie budget. For Tier 3 films, each additional million dollars in movie budget increases critical acclaim score by around 0.00055. Around 0.0001% of the variance in critical acclaim is

dependent on the movie budget. For Tier 4 films, each additional million dollars in movie budget increases critical acclaim score by around 0.00451. Around 0.04% of the variance in critical acclaim is dependent on the movie budget.

Even after splitting movies into separate budget tiers, critical acclaim relies *very little* on budget. When the budget is very low or very high (below \$5 million and above \$20 million), budget has nearly no effect on the amount critical acclaim received. For mid-budget movies, for which the budget is between \$5 million and \$20 million, budget has more of an influence on critical acclaim. Even so, in each additional million dollars in budget, the critical acclaim score increases by less than one.

Q2: Does *domestic gross* influence critical acclaim?

We compared a movie's domestic gross with its critical acclaim score. We calculated a coefficient of determination of  $r^2 = 0.0593$  and a slope of  $1.8372 \times 10^{-8}$  (**Figure 5**).

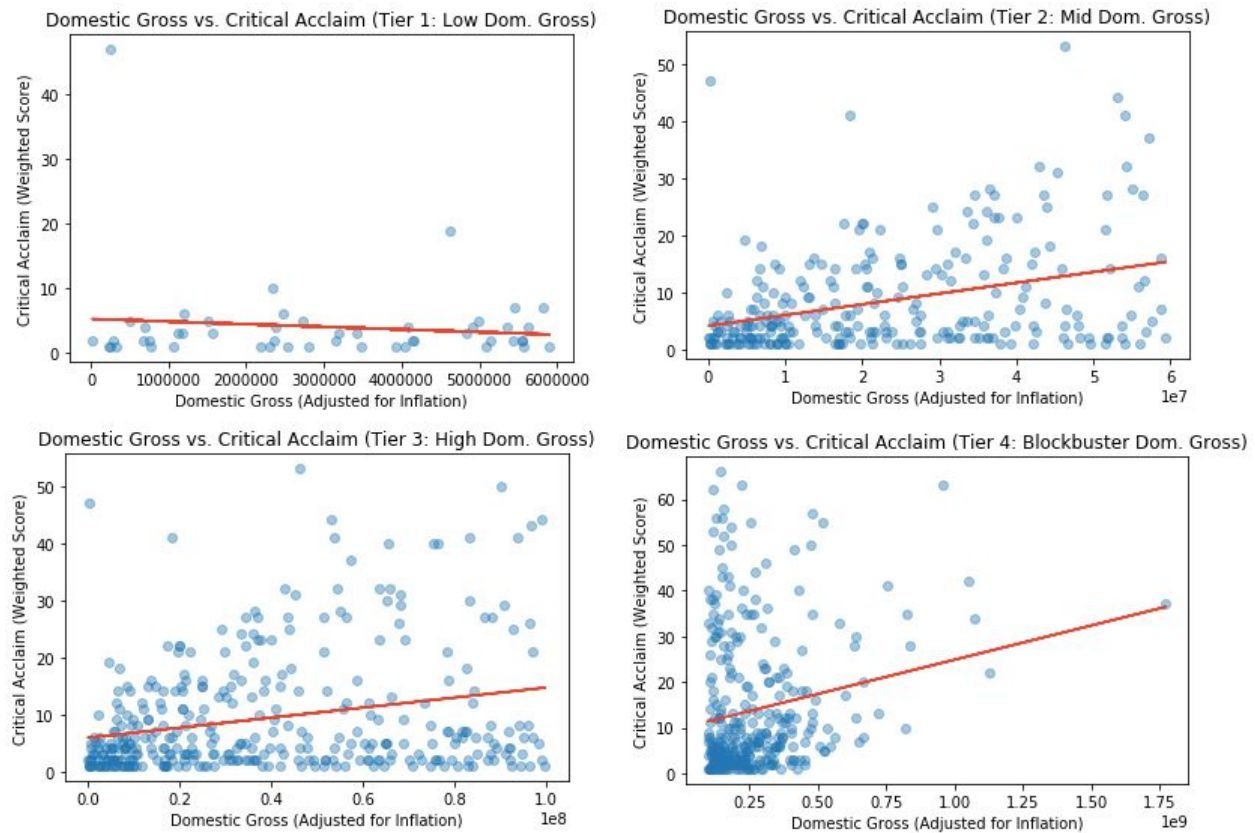


**Figure 5.**

Around 5.93% of the variance in critical acclaim is dependent on the movie's domestic box office revenue. Each additional million dollars in movie domestic gross increases critical acclaim score by around 0.0184. Therefore, domestic gross has a very small influence on critical acclaim. Once again, we split domestic gross into four tiers:

- **Tier 1: Domestic Gross < \$6,000,000**
  - Includes low-domestic gross films, mostly from low-budget, independent films.
- **Tier 2: \$6,000,000 < Domestic Gross < \$60,000,000**
  - Includes mid-domestic gross films.
- **Tier 3: \$60,000,000 < Domestic Gross < \$100,000,000**
  - Includes higher-domestic gross films.

- **Tier 4: Domestic Gross > \$100,000,000**
  - Includes high-domestic gross, blockbuster films.



**Figure 6.**

For Tier 1 films,  $r^2 = 0.0131$ , and slope =  $-4.0817 \times 10^{-7}$  (top left).

For Tier 2 films,  $r^2 = 0.1195$ , and slope =  $1.8838 \times 10^{-7}$  (top right).

For Tier 3 films,  $r^2 = 0.0630$ , and slope =  $8.7826 \times 10^{-8}$  (bottom left).

For Tier 4 films,  $r^2 = 0.0344$ , and slope =  $1.5000 \times 10^{-8}$  (bottom right).

For Tier 2 films, there is a moderately weak, positive correlation between domestic gross and critical acclaim. Each additional million dollars in domestic gross increases critical acclaim score by around 0.188. Around 11.95% of the variance in critical acclaim is dependent on the domestic gross.

For Tier 1, 3, and 4 films, there is close to no correlation between domestic gross and critical acclaim. For Tier 1 films, each additional million dollars in domestic gross decreases critical acclaim score by around 0.408. Around 1.31% of the variance in critical acclaim is dependent on

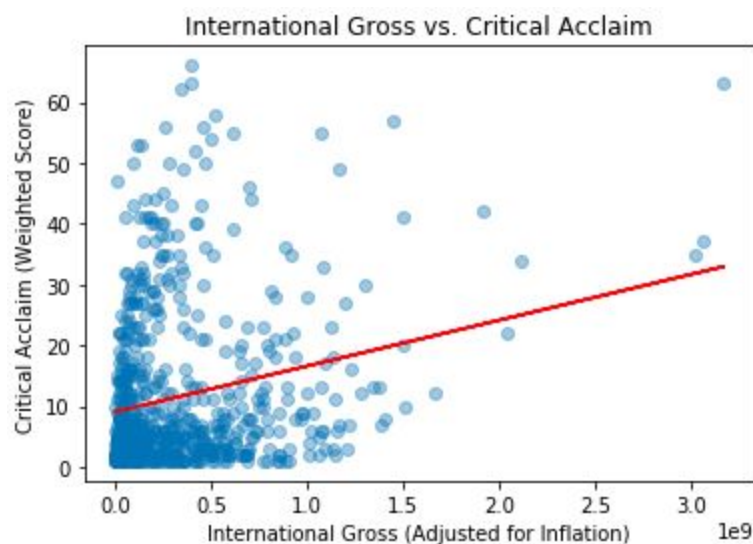


the domestic gross. For Tier 3 films, each additional million dollars in domestic gross increases critical acclaim score by around 0.0878. Around 6.30% of the variance in critical acclaim is dependent on the domestic gross. For Tier 4 films, each additional million dollars in domestic gross increases critical acclaim score by around 0.015. Around 3.44% of the variance in critical acclaim is dependent on the domestic gross.

Critical acclaim relies *very little* on domestic gross, as with budget. When the domestic gross is very low or very high (below \$6 million and above \$60 million), domestic gross has nearly no effect on the amount critical acclaim received. For mid-domestic gross movies, for which the domestic gross is between \$6 million and \$60 million, domestic gross has more of an influence on the amount of critical acclaim. Even so, in each additional million dollars in domestic gross, the critical acclaim score increases only by 0.188.

Q3: Does *international gross* influence critical acclaim?

We calculated a coefficient of determination of  $r^2 = 0.0501$ , and a slope of  $7.5382 \times 10^{-9}$  for the relationship between international gross and critical acclaim (**Figure 7**).

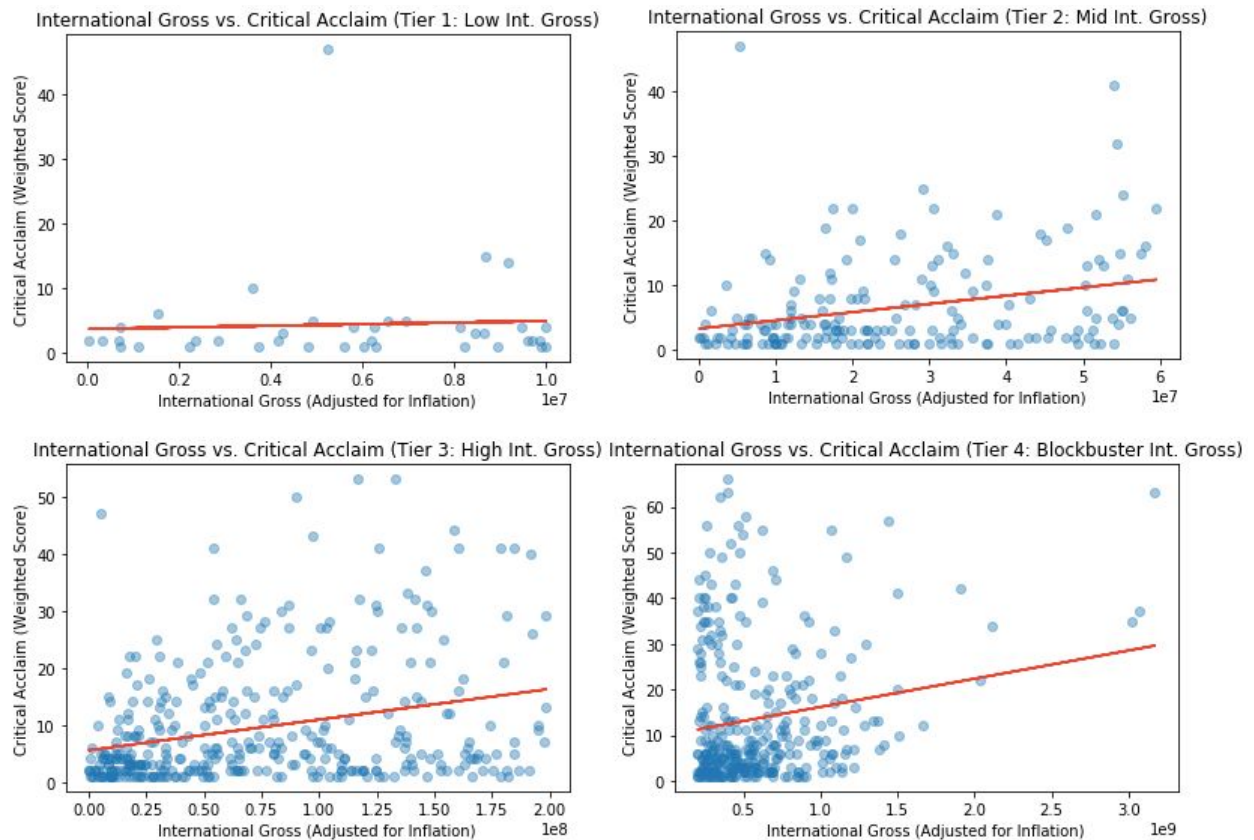


**Figure 7.**

Around 5.01% of the variance in critical acclaim is dependent on the movie's international box office revenue. Each additional million dollars in movie international gross increases critical acclaim score by around 0.00754. Therefore, international gross has a fairly small influence on critical acclaim. Just as for domestic gross, we split international gross into four similar tiers:

- **Tier 1: International Gross < \$10,000,000**
  - Includes low-international gross films, mostly from low-budget, independent films.

- **Tier 2: \$10,000,000 < International Gross < \$60,000,000**
  - Includes mid-international gross films.
- **Tier 3: \$60,000,000 < International Gross < \$200,000,000**
  - Includes higher-international gross films.
- **Tier 4: International Gross > \$200,000,000**
  - Includes high-international gross, blockbuster films.



**Figure 8.**

For Tier 1 films,  $r^2 = 0.0027$ , and slope =  $1.2400 \times 10^{-7}$  (top left).

For Tier 2 films,  $r^2 = 0.0828$ , and slope =  $1.2747 \times 10^{-7}$  (top right).

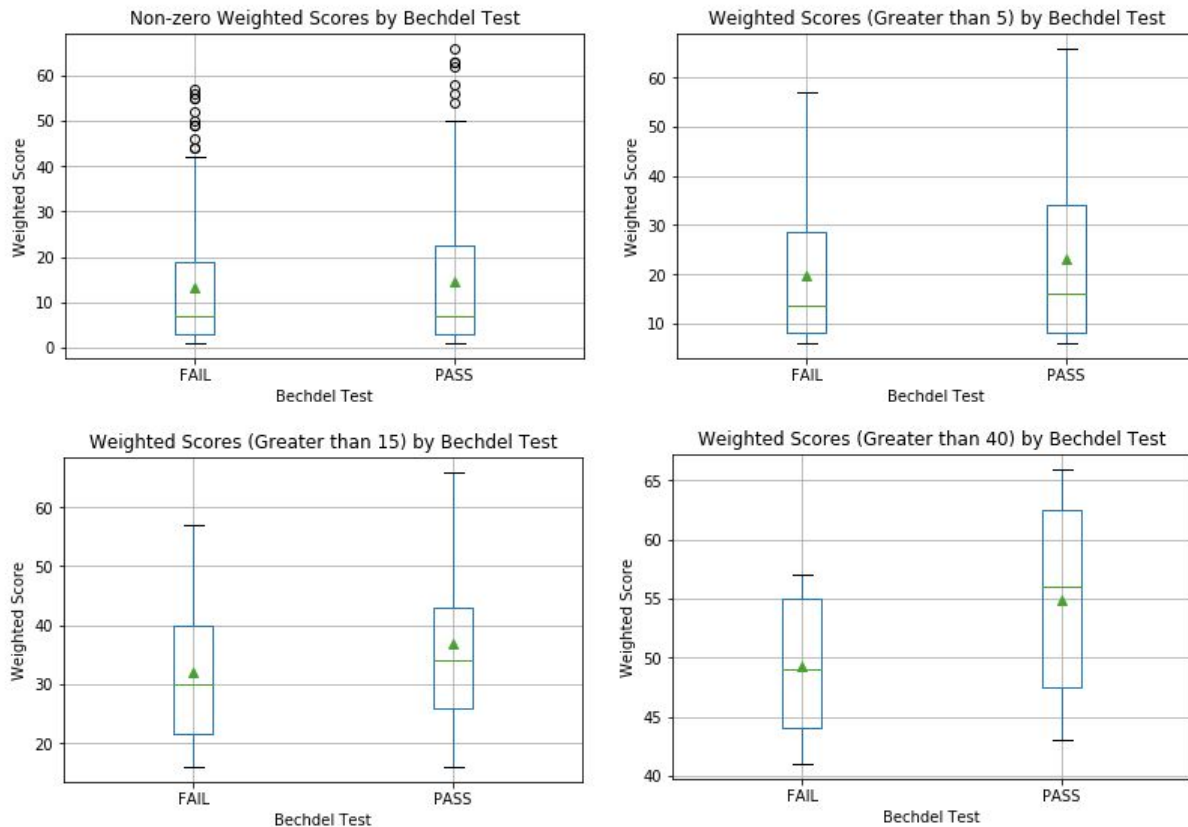
For Tier 3 films,  $r^2 = 0.0799$ , and slope =  $5.3684 \times 10^{-8}$  (bottom left).

For Tier 4 films,  $r^2 = 0.0290$ , and slope =  $6.1754 \times 10^{-9}$  (bottom right).

The pattern for international gross is very similar to that of domestic gross. Overall, critical acclaim depends *very little* on international gross. When the international gross is very low or very high (below \$10 million and above \$200 million), international gross has nearly no effect on the amount critical acclaim received. For mid and high-international gross movies, for which

the budget is between \$6 million and \$200 million, international gross has more of an influence on the amount of critical acclaim. Even so, in each additional million dollars in budget, the critical acclaim score increases by an extremely small amount, between 0.0537 and 0.127.

Q4: Does passing the *bechdel test* influence critical acclaim?



**Figure 9.**

For films with non-zero weighted scores, p-value = 0.9784 (**Figure 9**, top left).

For films with weighted scores greater than 5, p-value = 0.2532 (**Figure 9**, top right).

For films with weighted scores greater than 15, p-value = 0.1189 (**Figure 9**, bottom left).

For films with weighted scores greater than 40, p-value = 0.0869 (**Figure 9**, bottom right).

We analyzed whether a film's passing of the Bechdel test has an effect on its critical acclaim by generating boxplots and performing the Kruskal-Wallis test to determine significance. When considering all films with at least 1 nomination, Bechdel test result has little effect on the weighted score as the p-value is 0.9784. Even when considering only the most critically successful movies (weighted scores greater than 40), the p-value of 0.0869 still does not support

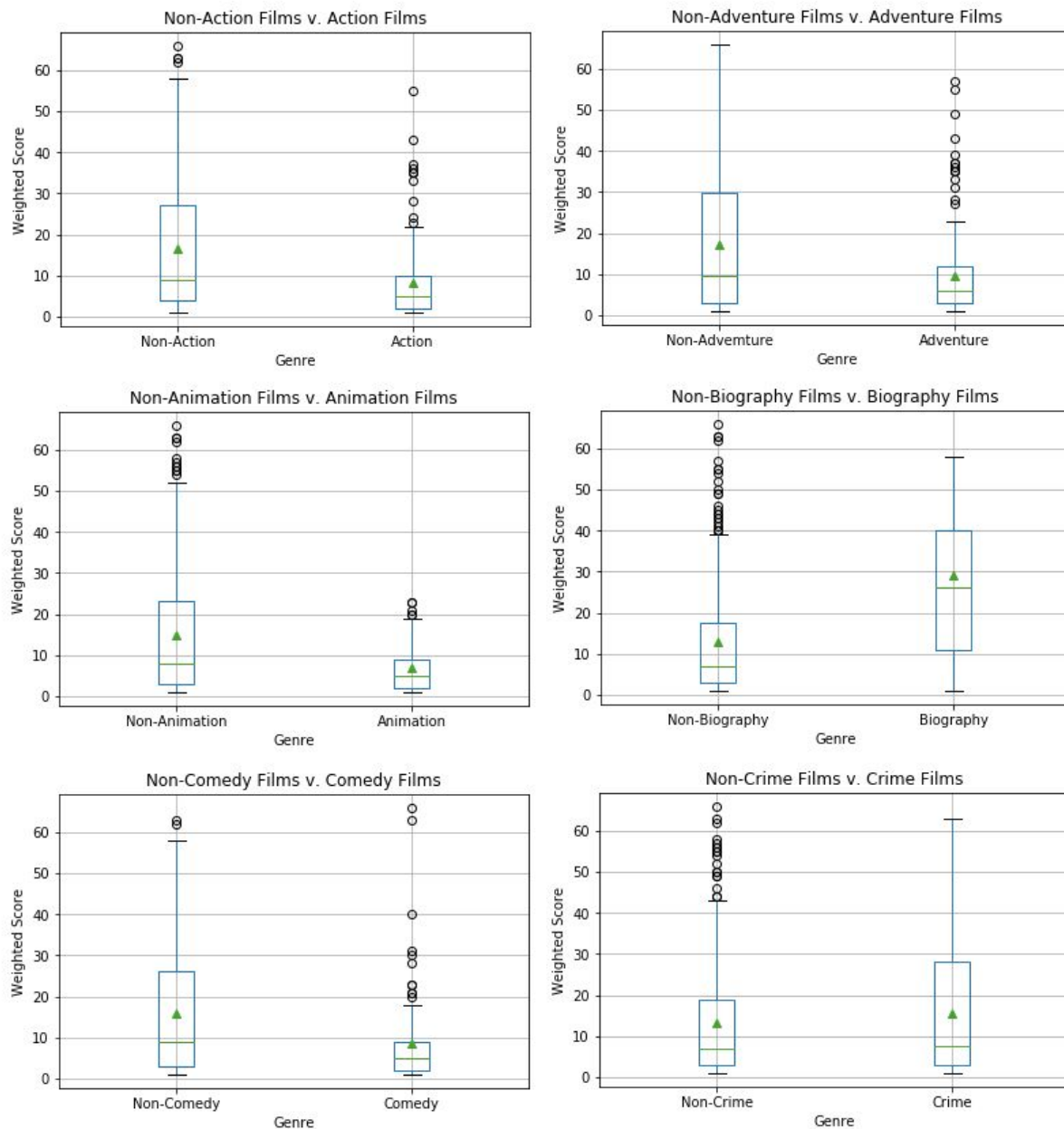
a statistically significant relationship between passing the Bechdel test and critical acclaim. However, the continuous decrease in p-value as we considered greater acclaimed films suggests that there could be a relationship if we increased our sample size to include newer films. The boxplot for films with weighted scores greater than 40 (**Figure 9**, bottom right) shows a greater median score for films that passed the Bechdel test.

Q5: Does the *genre* of a film influence critical acclaim?

It is worth noting that a film can have up to 3 genres. A film may belong to more genres, but only the 3 most important genres determined by IMDB are considered. The Kruskal Wallis statistical analysis of 18 genres can be summarized by the table and from the graphs below (**Figure 10, 11, and 12**).

N/A: Western (could not perform the Kruskal Wallis test because there are only 4 data points)

	Genre in which the median weighted score is <b>greater</b> than the median weighted score of films not in that genre	Genre in which the median weighted score is <b>less</b> than the median weighted score of films not in that genre
Statistically significant (p-value < 0.05)	Biography (median: 26.0) Drama (median: 20.0) History (median: 37.5) Musical (median: 31.5) War (median: 32.0)	Action (median: 5.0) Adventure (median: 6.0) Animation (median: 5.0) Comedy (median: 5.0) Sci-Fi (median: 5.0)
Not statistically significant (p-value > 0.05)	Crime (median: 7.5) Fantasy (median: 7.5) Romance (median: 10.0)	Family (median: 6.0) Horror (median: 6.0) Mystery (median: 4.5) Thriller (median: 6.0)



**Figure 10.**

Non-action films vs. action films,  $p\text{-value} = 1.336 \times 10^{-6}$  (top left)

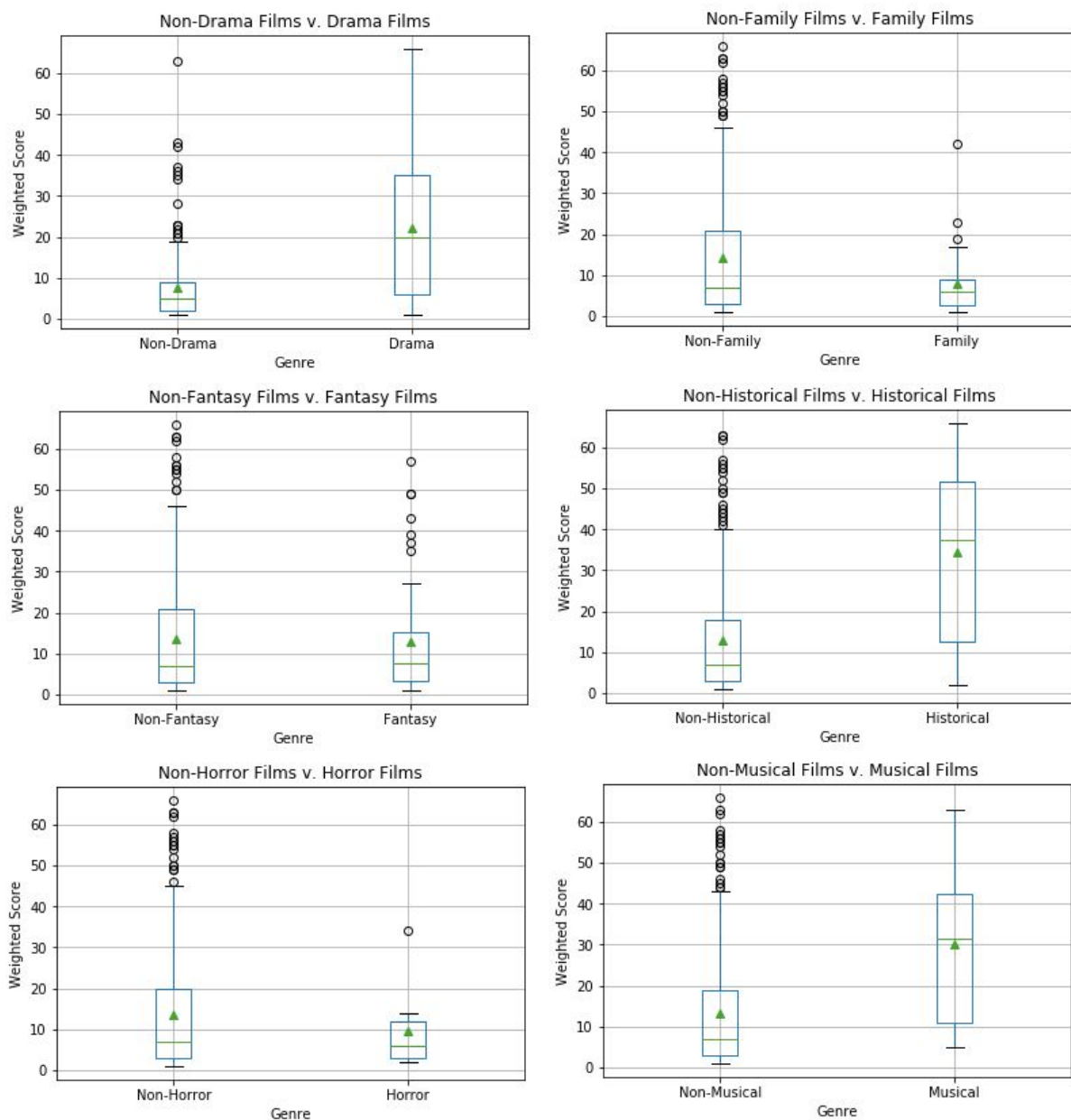
Non-adventure films vs. adventure films,  $p\text{-value} = 0.0004$  (top right)

Non-animation films vs. animation films,  $p\text{-value} = 0.0022$  (middle left)

Non-biography films vs. biography films,  $p\text{-value} = 0.0003$  (middle right)

Non-comedy films vs. comedy films,  $p\text{-value} = 8.8375 \times 10^{-2}$  (bottom left)

Non-crime films vs. crime films,  $p\text{-value} = 0.5899$  (bottom right)



**Figure 11.**

Non-drama films vs. drama films,  $p\text{-value} = 1.5390 \times 10^{-14}$  (top left)

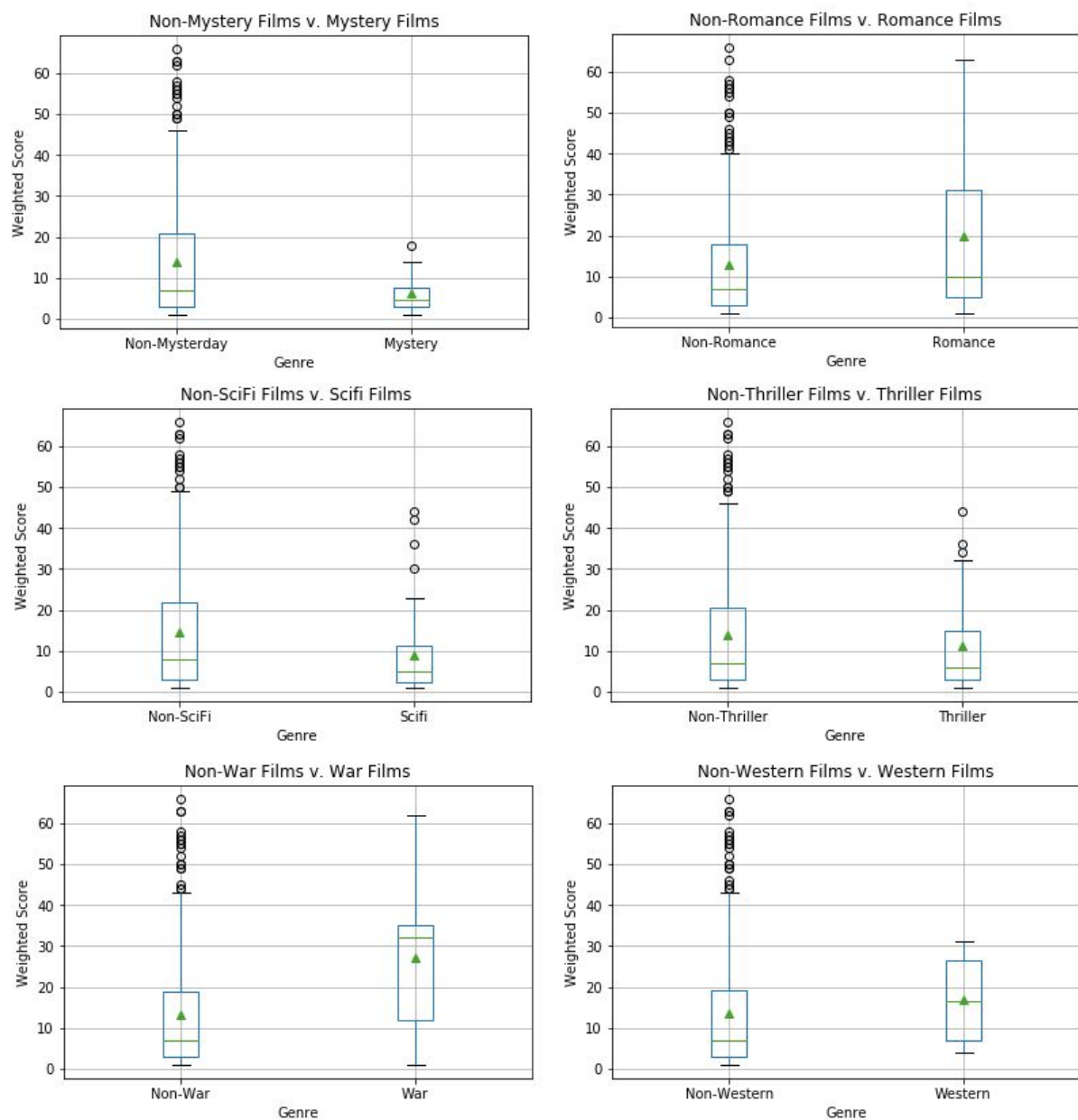
Non-family films vs. family films,  $p\text{-value} = 0.0878$  (top right)

Non-fantasy films vs. fantasy films,  $p\text{-value} = 0.9272$  (middle left)

Non-historical films vs. historical films,  $p\text{-value} = 0.0004$  (middle right)

Non-horror films vs. horror films,  $p\text{-value} = 0.6592$  (bottom left)

Non-musical films vs. musical films,  $p\text{-value} = 0.0004$  (bottom right)



**Figure 12.**

Non-mystery films vs. mystery films, p-value = 0.0972 (top left)

Non-romance films vs. romance films, p-value = 0.0513 (top right)

Non-SciFi films vs. SciFi films, p-value = 0.0193 (middle left)

Non-thriller films vs. thriller films, p-value = 0.7178 (middle right)

Non-war films vs. war films, p-value = 0.0170 (bottom left)

Non-western films vs. western films, no p-value as the sample size is too small (bottom right)



## Discussion

Our hypothesis regarding the budget, domestic-, and international gross was slightly off. We expected a much stronger correlation between each of those variables and critical acclaim for all films not in the highest tiers. As expected, the highest tiers (Tier 4) for budget, domestic-, and international gross demonstrate no significant correlation between each of those variables and critical acclaim. Unexpectedly, across all three variables, only the middle tiers (Tier 2 for budget and domestic gross, Tier 2 and 3 for international gross) demonstrate a slightly noticeable positive correlation ( $r^2$  between 0.08 and 0.12). Low-budget and grossing films also have no budget-critical acclaim or gross-critical acclaim correlations.

We can conclude that for the most part, how much budget is invested into a movie and how well it performs at the box office are not indicators of whether the film be critically acclaimed. However, if the film has a moderate budget between \$5 million and \$20 million, a moderate domestic gross between \$6 million and \$60 million, and/or a moderate-to-high international gross between \$10 million to \$200 million, higher numbers in those categories *may* occasionally indicate higher critical acclaim.

The data does not support our hypothesis regarding Bechdel test and critical acclaim. The boxplots for each acclaimed tier (weighted scores of  $> 1$ ,  $> 5$ ,  $> 15$ , and  $> 40$ ) suggests at first glance that Bechdel-passing films have higher median weighted scores than films that failed the test. Yet the differences in medians are not statistically significant. However, as we filtered the dataset to include only the more critical acclaimed films, we discovered that the p-value lowers. For films with weighted scores  $> 40$ , the p-value for films that pass the Bechdel test and films that do not is 0.0869. The decrease in p-value suggests the possibility that Bechdel-passing films are more critically acclaimed when only considering acclaimed movies. In other words, films that pass the test and films that fail the test can both achieve critical success, but films that pass the test has the potential for higher acclaim.

The idea of “Oscar bait” is most obvious when looking at genres. Films labeled as biography, drama, history, musical, and war have median weighted scores between 20.0 and 37.5 and receive significantly higher critical acclaim, while films labeled as action, adventure, animation, comedy, and science-fiction have median weighted scores of 5.0 or 6.0 and receive statistically lower critical acclaim. The results can be explained by the nature of the genres; biopics, dramas, historical movies, musicals, and war films provide more opportunities for acting nominations and are held to higher esteem culturally. With a p-value of  $1.5390 \times 10^{-14}$ , films in the drama genre are significantly more critically successful than films that are not in that genre. Some outliers worth noticing include *Shakespeare In Love* and *Chicago* in the comedy genre, scoring 66 and 63 respectively while the comedy median is 5.



In conclusion, films in the biography, drama, history, musical, and war genres that pass the Bechdel test with a moderate budget that does moderately well at the box office have the greatest potential for awards success.

## Future Direction

The research can be improved upon and expanded in several ways. First, we could add newer films to the dataset. Voters for these awards are often reactive, and the country's political climate has changed drastically these past years. An expanded dataset could paint a more complete and applicable picture of films and critical success. Second, we could group the films by decades to analyze how voters' taste in movies change over time and determine its implication for future predictions. Lastly, we could use other measures of "critically acclaim" such as MetaCritic scores to analyze how the variables discussed affect a movie's success.