Box Office Analysis

Graham Odell

6/23/2020

Project Overview

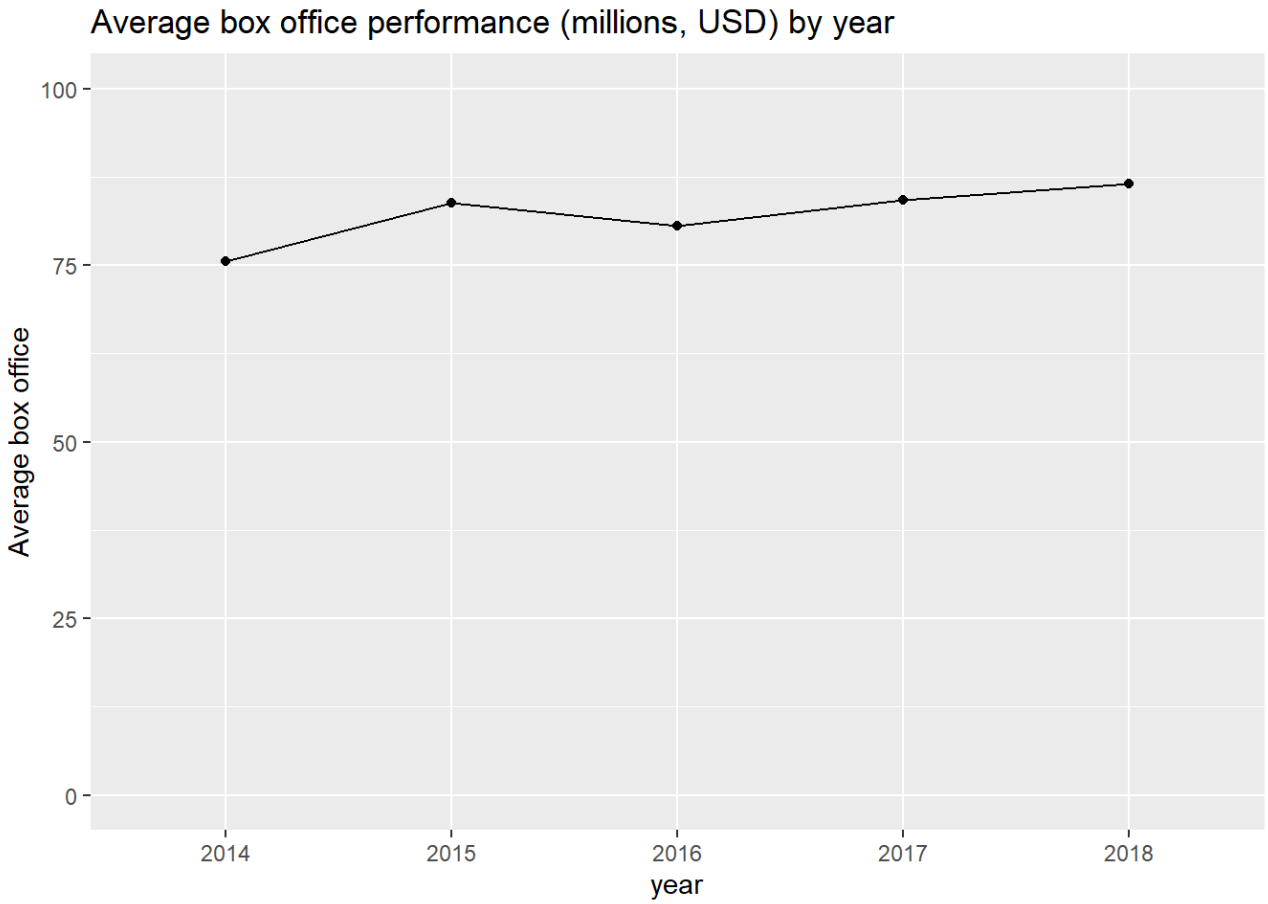
This project utilizes a custom dataset to explore patterns in box office performance of films released between 2014 and 2019. The dataset includes variables measuring, among other factors, box office performance (dollars earned in the North American market), prevalence of releases (number of theaters at opening), audience and critic reactions (Rotten Tomatoes scores), studio (name and size), genre and time of release (month and year). In this project, I focus on patterns relating to studio, month of release and Rotten Tomatoes score. This focus is driven by recent film industry commentary on the remarkable success of certain studios (e.g., Disney), the changing calendar of film releases (e.g., big-budget blockbusters released in February and March) and the outsized influence of review aggregators such as Rotten Tomatoes.

I try to illuminate patterns in the data and connections between the variables through standard statistical techniques, such as correlation matrices. linear regression models and machine learning algorithms, and extensive data visualization. More specifically, I use an OLS approach to estimate the effects of various predictors on the outcome variable of logged total box office receipts in the domestic (i.e., North American) market and then compare these models with one generated through the XGBoost algorithm.

In brief, my major findings are that the number of theaters at opening is the strongest predictor of total box office performance. Rotten Tomatoes scores are also substantial predictors, but perhaps not as much as recent film industry commentary suggests. The time of year of a release matters to some extent, especially in months such as December. When including other predictors, genre does not appear to have a major impact on the box office success of a film, whereas the size of the studio (measured in number of movies released) does. A number of these findings are not surprising, but it is nonetheless useful to have quantitative evidence that confirms one’s subjective impressions.

Dataset Exploration

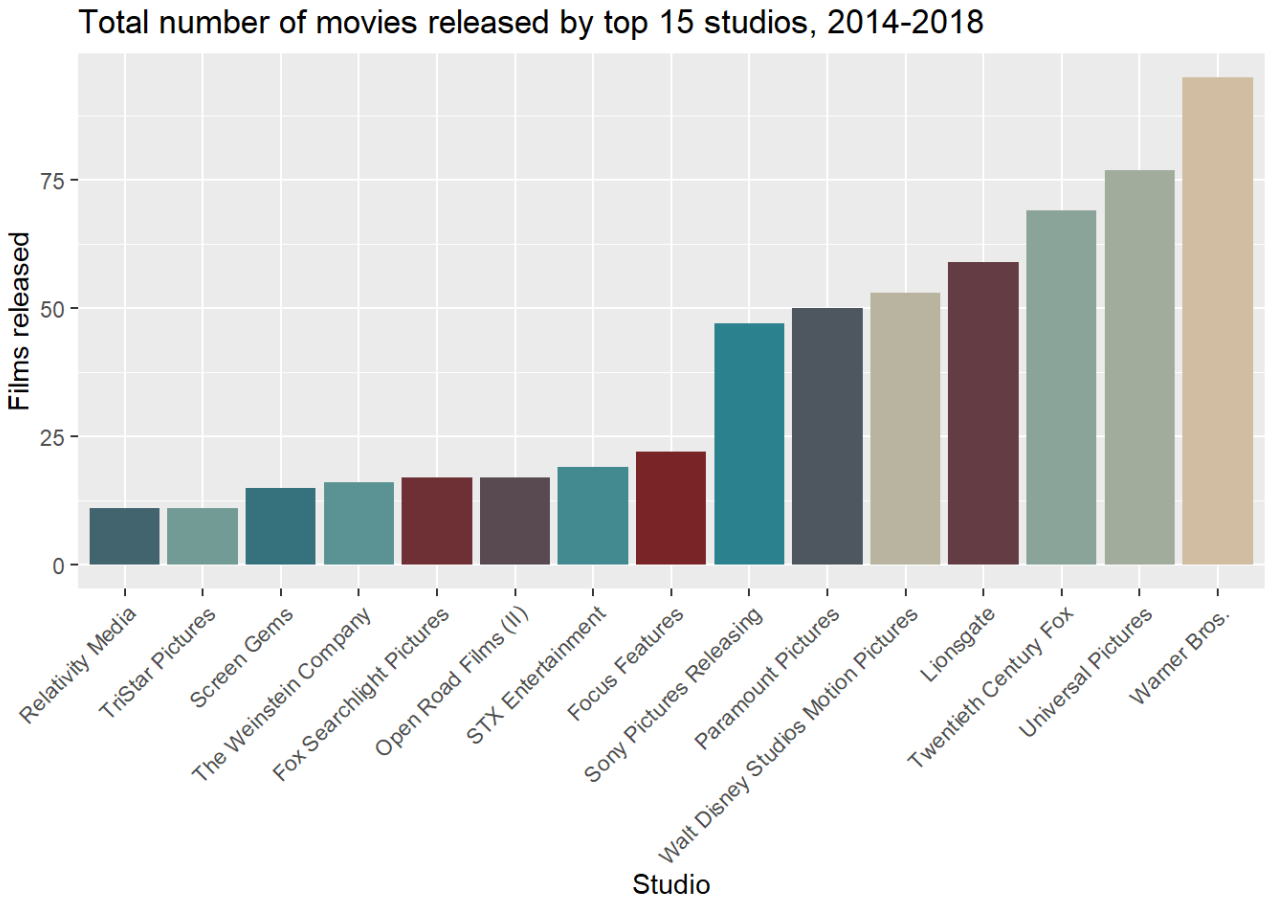
As can be seen in the plot below, there was a general increase in average box office performance per film between 2014 and 2018. This trend makes sense given the long-term increase in ticket prices, though this probably does not explain all of the increase observed. 2016 appears to have been a down year, though only slightly.



Studio-level patterns

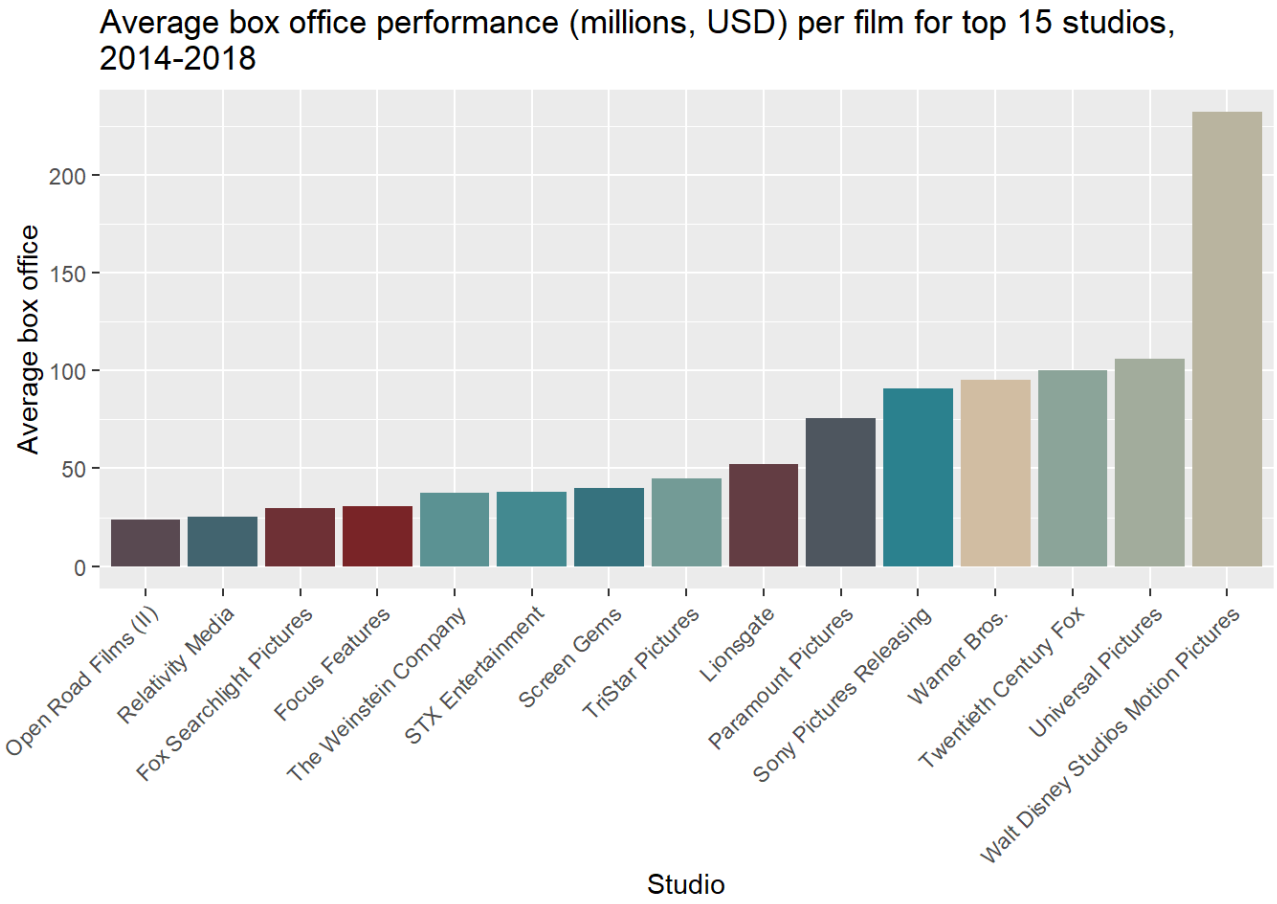
The following set of plots summarizes some key measures of film studios' performance and industry influence. Note that for the sake of visual clarity, I have excluded studios that released ten movies or less in the five year period covered by the dataset. There are a total of 41 studios represented in the dataset, of which 15 are displayed in the following plots.

The plot below displays the total number of movies released by each studio (more than ten total movies released). The "Big Seven" - Warner Bros., Universal, Fox, Lionsgate, Disney, Paramount and Sony - not surprisngly dominated the film schedule between 2014 and 2018. Indeed, we see a significant jump in the number of movies released between Focus Features and Sony Pictures. I use this cutoff to distinguish between midrange studios (between 10 and 45 movies released) and majors (more than 45 studios released). This cutoff is less arbitrary than the one used to distinguish between midrange studios and minor ones (ten movies or less), and, as we will see in the quantitative analyses below, the distinction between major and mid is more significant than the one between mid and minor when predicting box office performance.

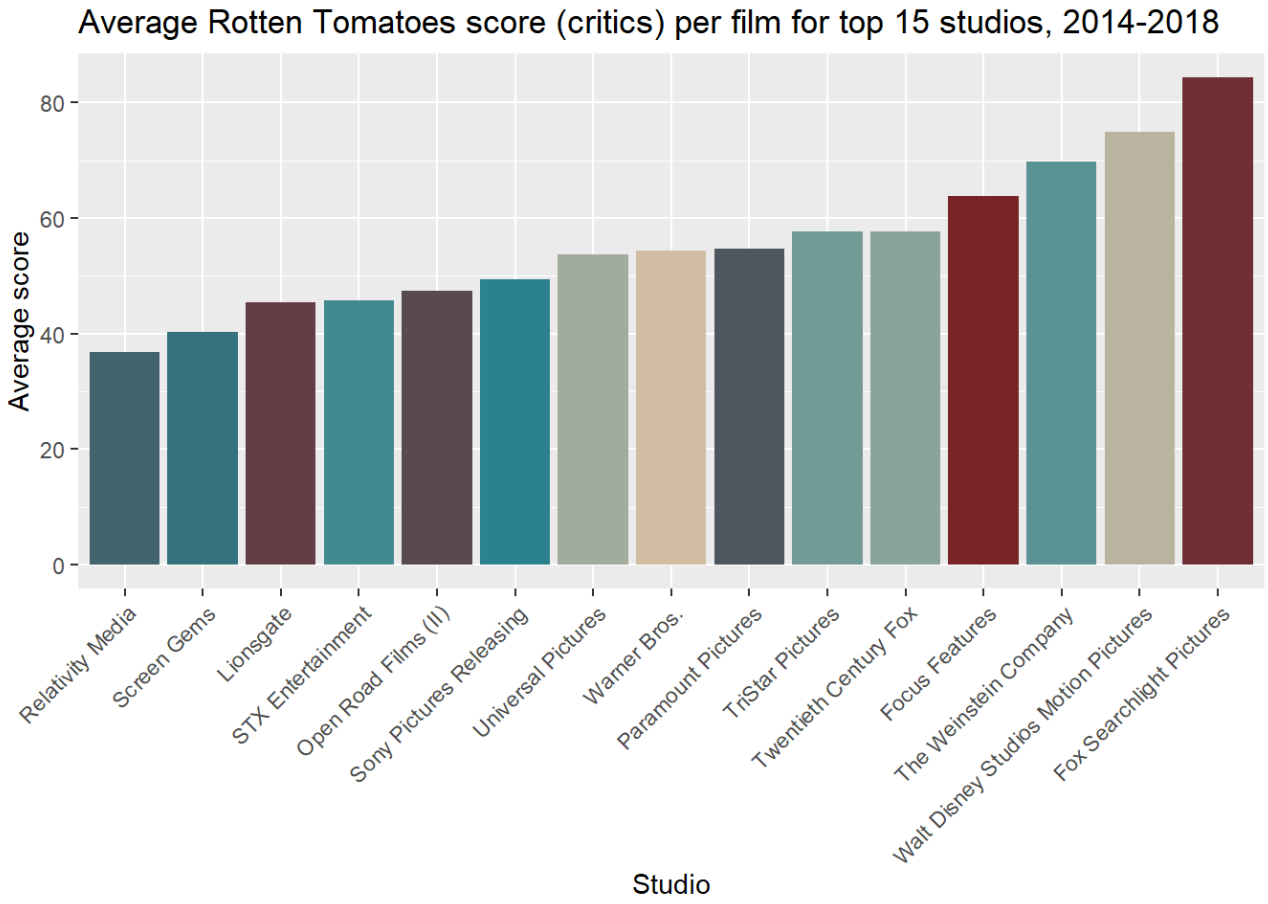


The second plot displays how well the typical movie released by each studio did at the North American box office. Studios are ordered from lowest average performance to highest. Disney is the clear industry leader, with its releases on average earning about 225 million dollars. Universal is a distant second with about 110 million dollars earned per movie. Notice that excluding Disney, there are no truly significant drops in average performance from studio to studio. The most significant of these other drops is between Paramount and Lionsgate, of about 25 million dollars.

In retrospect, it might make sense to include an additional variable that distinguishes Disney from non-Disney films. However, this approach would encourage overfitting since we would be using an estimate of box office performance to predict box office performance. Instead, by focusing on the number of movies released by a given studio, I am relying on a predictor that, at least in principle, is conceptually independent from the outcome I am trying to predict. The argument could be made, though, that past box office performance determines how much cash a studio has on hand to make more films in the future. This is certainly true, but given that the purpose of the project is to predict individual film performance, characteristics of a film's studio should be considered in any analysis. Rather than measure these characteristics directly (which would require some form of multilevel modeling), I instead collapse any potentially significant studio-level factors into the studio variable and any other variables derived from it. Future additions to this project can take a more sophisticated multilevel approach to further tease out the studio characteristics that explain why some studios release better performing movies than others.

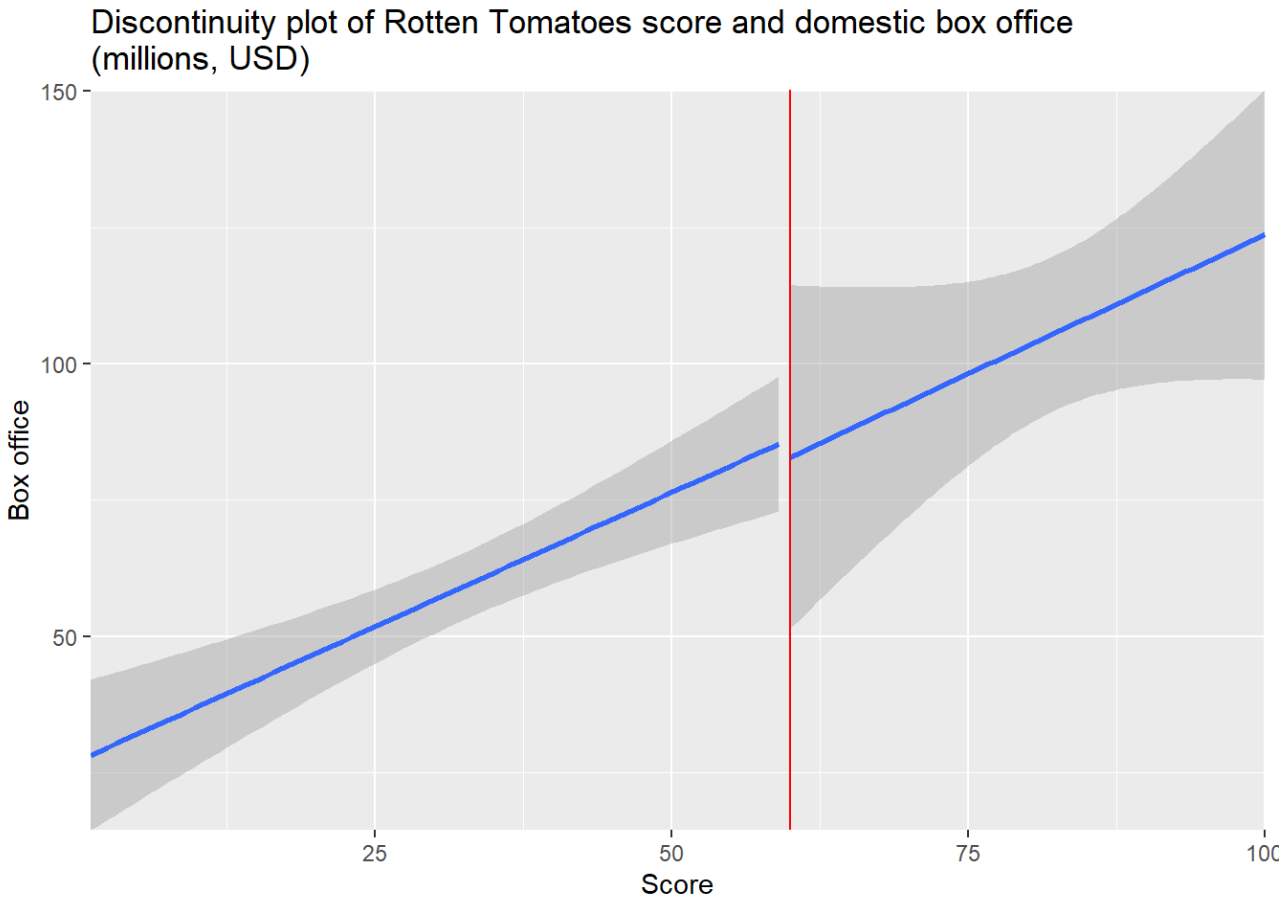


Moving on, we can visually inspect the relationship between studio and Rotten Tomatoes score. We again see a different ranking of studios, with Fox Searchlight boasting the best average score and Relativity Media at the bottom. The “Big Seven” are roughly evenly distributed across the rankings, with Sony having the sixth lowest average score and Disney having the second highest. Films from boutique studios such as Fox Searchlight, The Weinstein Company and Focus Features, not surprisingly, do very well with the critics. The high average score of Disney is notable, since the typical major studio on average puts out rather middling films, judged by critic reception (see Sony, Universal, Warner Bros. and Paramount - all around the 50% mark). Indeed, Disney is the only major studio that releases films with average critic scores of greater than 70%. Given the relatively high number of movies Disney puts out (more than ten a year), this consistency is adjudged quality is rather impressive.

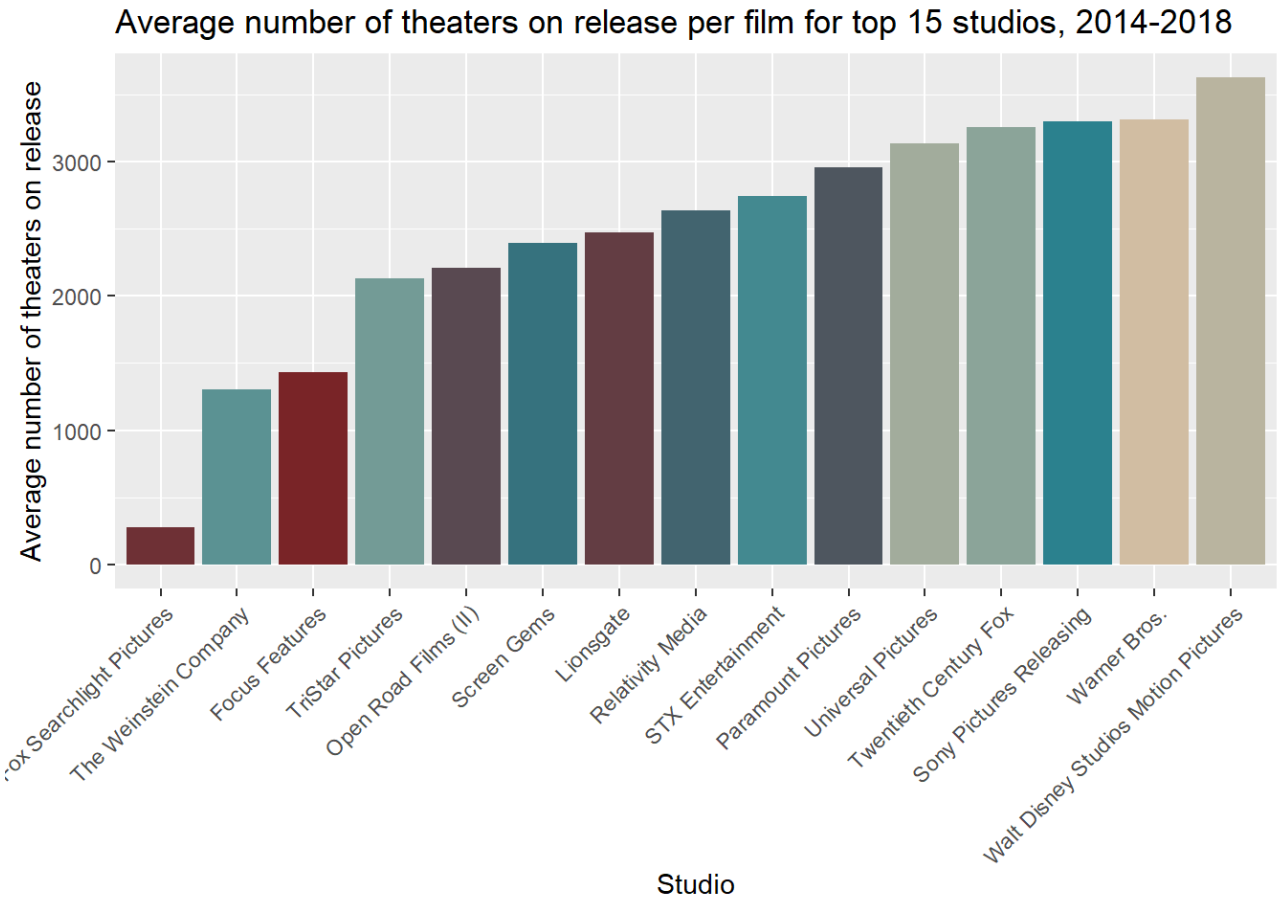


It's worth noting here how Rotten Tomatoes scores work. They represent the percentage of aggregated reviews that are "positive", a classification based in part on Rotten Tomatoes' editors' judgments about the scales that reviewers use. If a movie's aggregate percentage of positive reviews is 60% or greater, then the movie is classified as "Fresh" and receives an appealing red tomato icon next to its name on the Rotten Tomatoes site. If the percentage is less than 60%, then the film is classified as "Rotten" and is branded with a rude green splat icon.

One may wonder if the "Fresh" rating has an effect on the box office performance of a movie, independent of the percentage score. The following plot (which computes the relationship between the two variables separately for RT scores above and below the 60% threshold) suggests there is no special advantage a film receives for having a "Fresh" rating. A film with a 59% rating is expected to have a box office performance about as good as one with a 60% rating. Note that the red line indicates the threshold for obtaining a "Fresh" rating.



The final plot that looks at studio performance visualizes the average number of theaters on a film's release. Not surprisingly, boutique studios like Fox Searchlight and Focus Features tend to not open their movies in wide releases. Six of the "Big Seven" occupy the top six ranks - again, not a surprising finding. Interestingly, despite Lionsgate being considered one of the "Big Seven" based on the number of films released per year, the studio has a middling average number of theaters on release, around 2500. Disney holds the top spot, with more than 3500 theaters showing one of its movies on opening day, on average.



Before moving to the next stage of the analysis, let's examine the correlations between these averaged characteristics of the top 15 studios. The average box office performance of a studio's films is moderately correlated with the total numbers of movies it releases and the average number of theaters its movies are shown in on release. There is almost no correlation between a studio's average Rotten Tomatoes score and the number of films it releases (my proxy for studio size). The only negative correlation is between the average Rotten Tomatoes score and the average number of theaters on release. This finding suggests that critics tend to prefer "smaller" movies (i.e., films released less widely). This interpretation is supported when we look at the correlation between critics score and number of theaters on release, for movies rather than studios: -0.24. Though weaker than the correlation when our unit of analysis is studio, the correlation between the two variables on a film-unit basis is still negative. Interestingly, the correlation between the Rotten Tomatoes *audience* score and number of theaters on release is virtually the same: -0.23. One complaint that often appears on the internet is that critics and audience members have divergent attitudes on what makes a good movie. These findings suggest this concern may be misplaced. I will have more to say on this issue later in the analysis.

Correlations, studio-level values

	Average number of theaters	Total movies	Average box office performance	Average Rotten Tomatoes score (critics)
Average number of theaters		0.64	0.65	-0.45
Total movies	0.64		0.6	0.01
Average box office performance	0.65	0.6		0.32
Average Rotten	-0.45	0.01	0.32	

Tomatoes score (critics)

It is clear from the above visualizations and analysis that there are substantial differences across studios when it comes to the kinds of movies that are made and how well they perform at the box office. These findings justify including a predictor based on a movie's studio in any models that seek to predict a film's box office performance. However, including a variable that has a label for every studio would be overly complicated. So instead I create a new variable with just three categories based on the total number of movies a studio released between 2014 and 2018. If a studio released ten movies or less, I classify it as "minor". If the studio released more than ten but less than 46, it is classified as "mid". All other studios are classified as "major". These major studios correspond with the "Big Seven" I listed above.

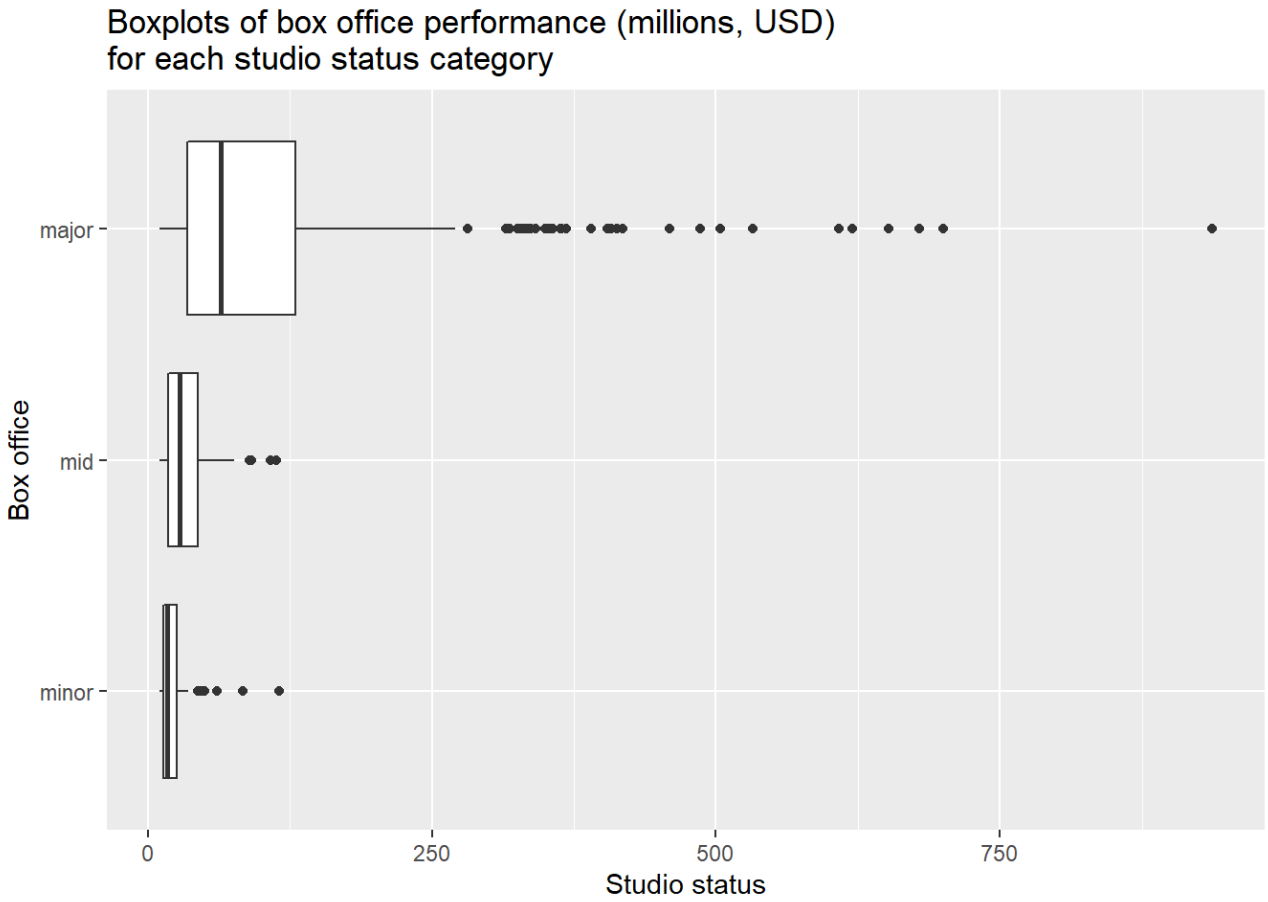
Observed patterns by studio status

This section explores the data after collapsing the studio variable into three categories or statuses - minor, mid and major - as described above. As the table below shows, movies released by major studios on average earn about 105m USD, dramatically more than those of mid-sized or minor studios. Despite only releasing 0.58 films per year, the typical minor film studio can expect to earn not that much less per film than a mid-sized studio, which on average releases 3.2 per year.

Studio status	Average domestic box office (millions)
minor	22.96
mid	33.47
major	105.71

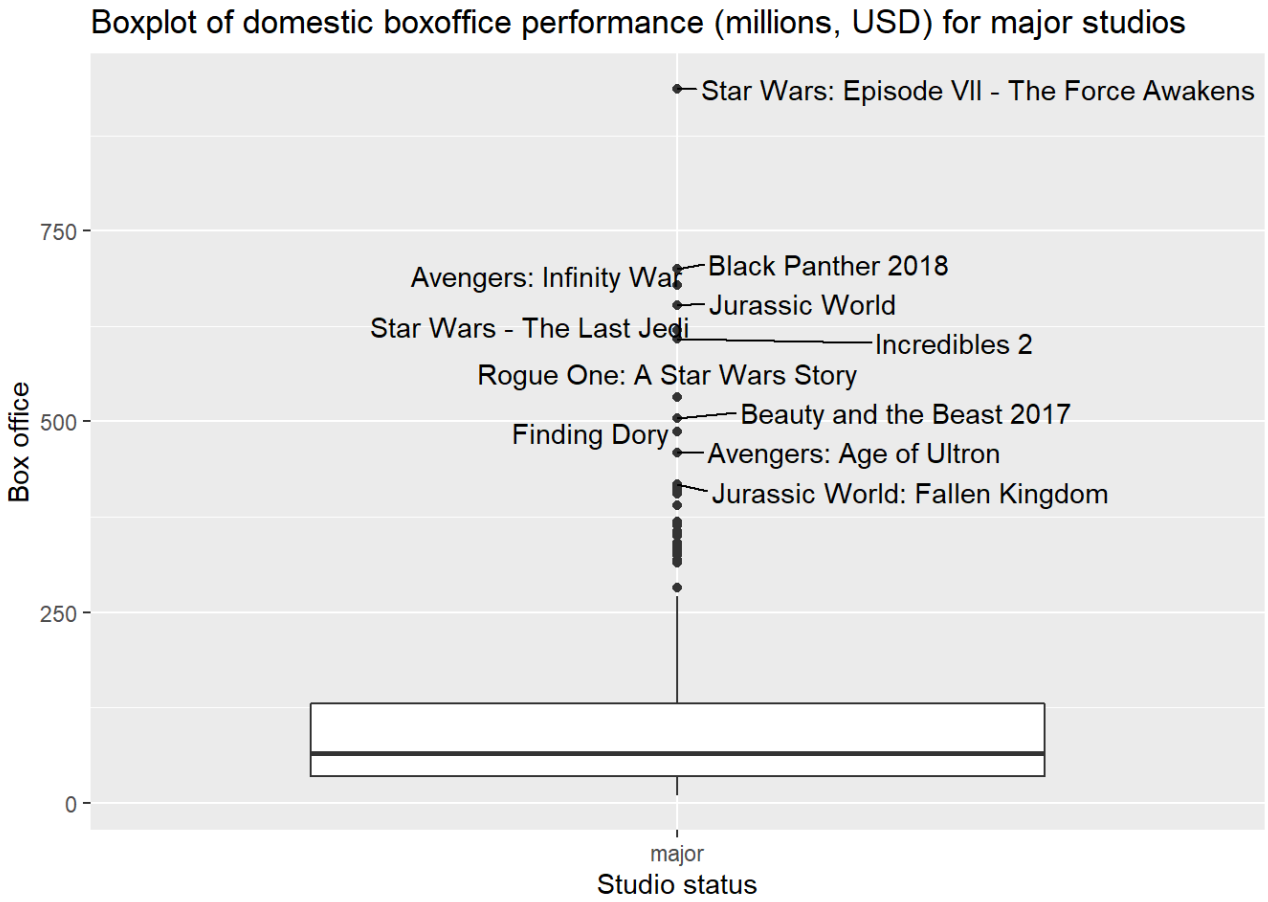
Boxplots of domestic box office results by studio status

The first plot shows the distributions for all three studio categories together. The median values (visualized by the black lines inside the boxes) are not too different from each other. However, the major studio category has far more outliers (defined as at least 1.5 times greater/smaller than the 75th/25th percentile of that distribution) and a much wider range. A movie appears much more likely to be a box office smash if it is released by a major studio than by a mid-size or minor one.

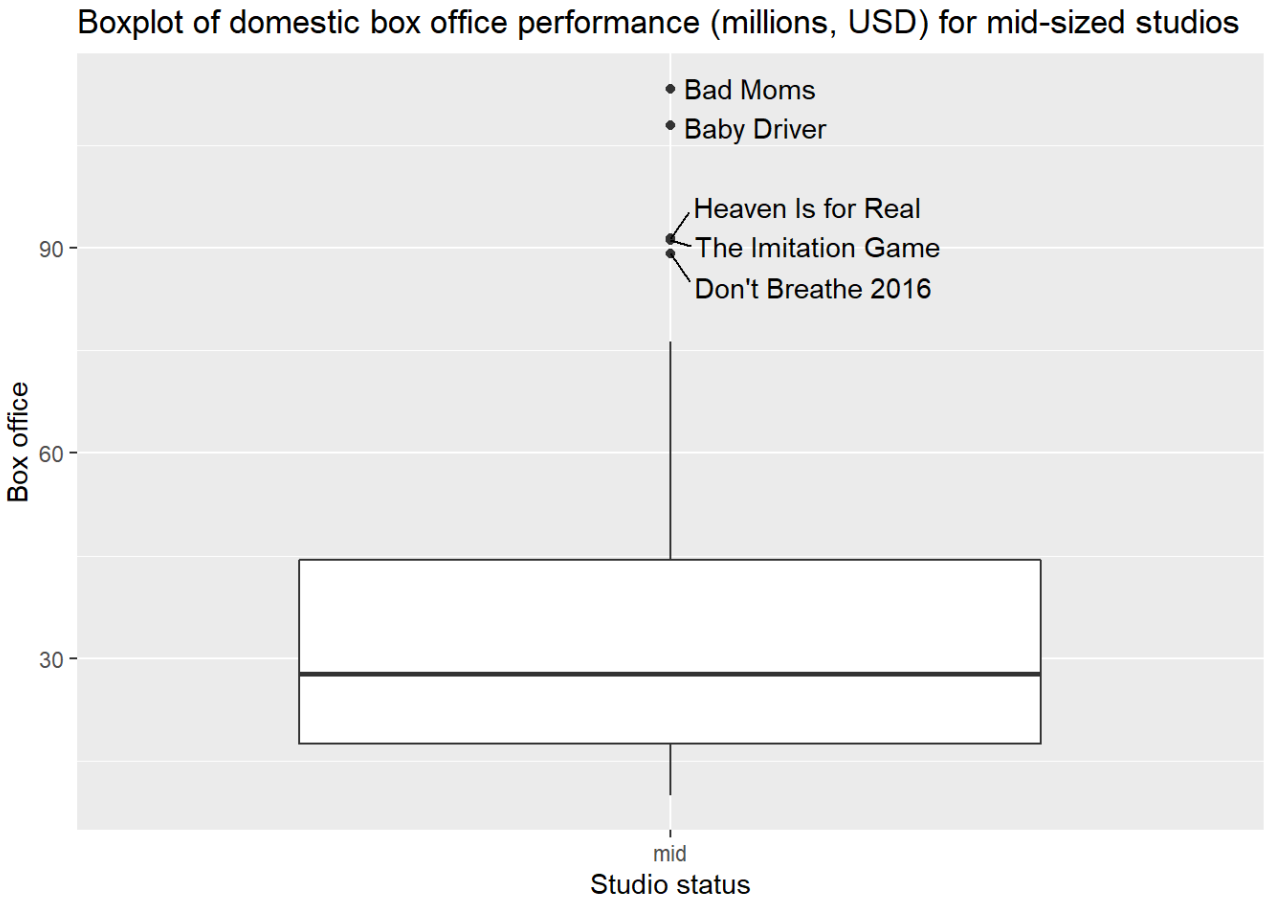


The next three plots explore the distribution of domestic box office performance (in millions of USD) for each studio status category - minor, mid-sized and major. Each plot displays the standard boxplot values (e.g., median, lower quartile, upper quartile) and names the outliers for each category. For the major-studio plot, I only label the “extreme” outliers ((defined as at least 3 times greater/smaller than the 75th/25th percentile of that distribution) so that the text labels are legible.

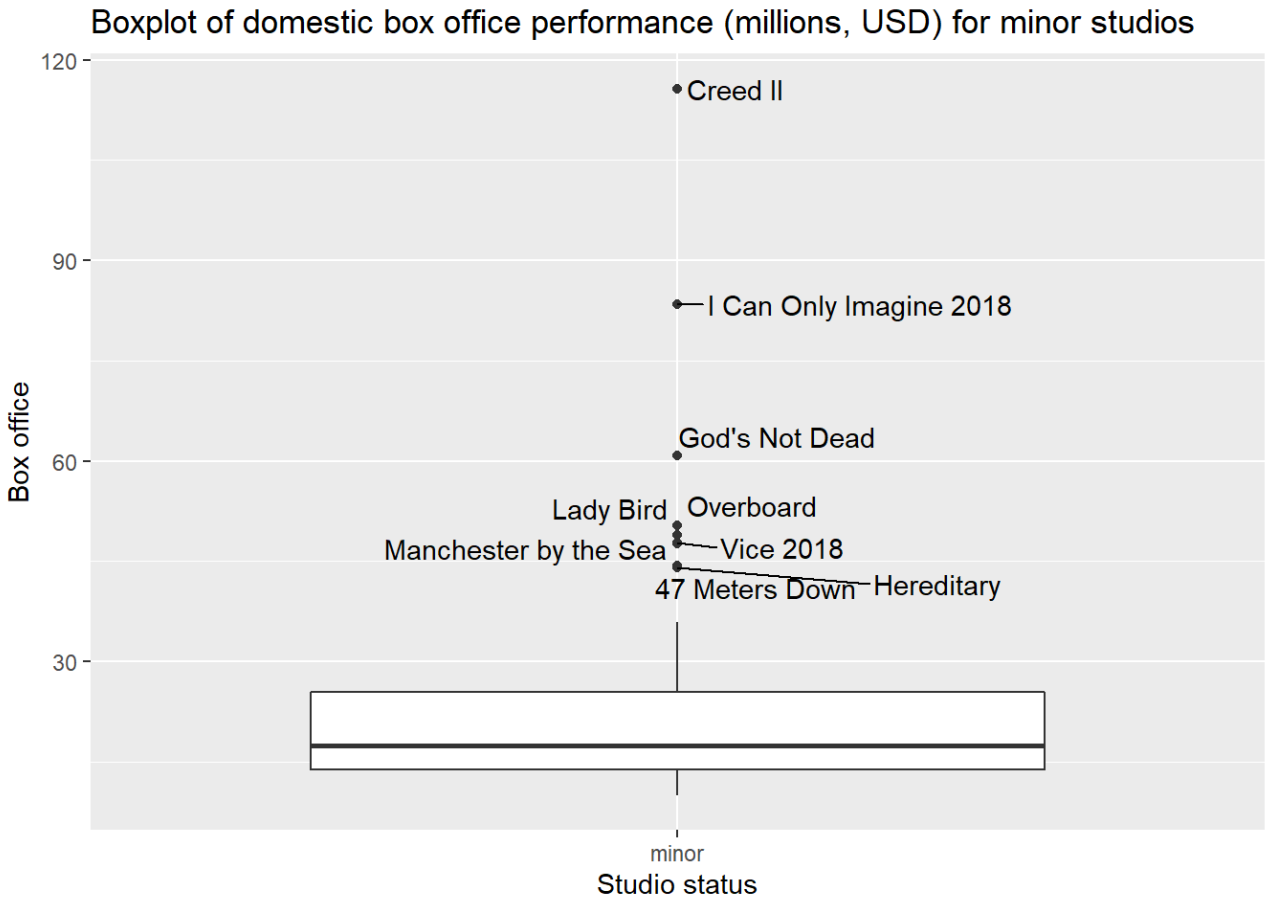
The first plot shows that fifty percent of major-studio films earn between 16.535 and 112.045 million USD. However, there are many outliers that earn well above 250 million USD at the box office. Star Wars and Marvel films dominate the group of outliers, with The Force Awakens earning the most money with over 875 million USD. In fact, virtually all of these outlier films were released by Disney.



The second boxplot (displayed below) visualizes the distribution of domestic box office performance for mid-sized studios. Fifty percent of these studios' films earned between 14.33 and 41.11 million USD. There are some outliers, but not nearly as many as for the major studios. Bad Moms was the highest performing mid-sized studio film between 2014 and 2018, earning more than 105 million USD.



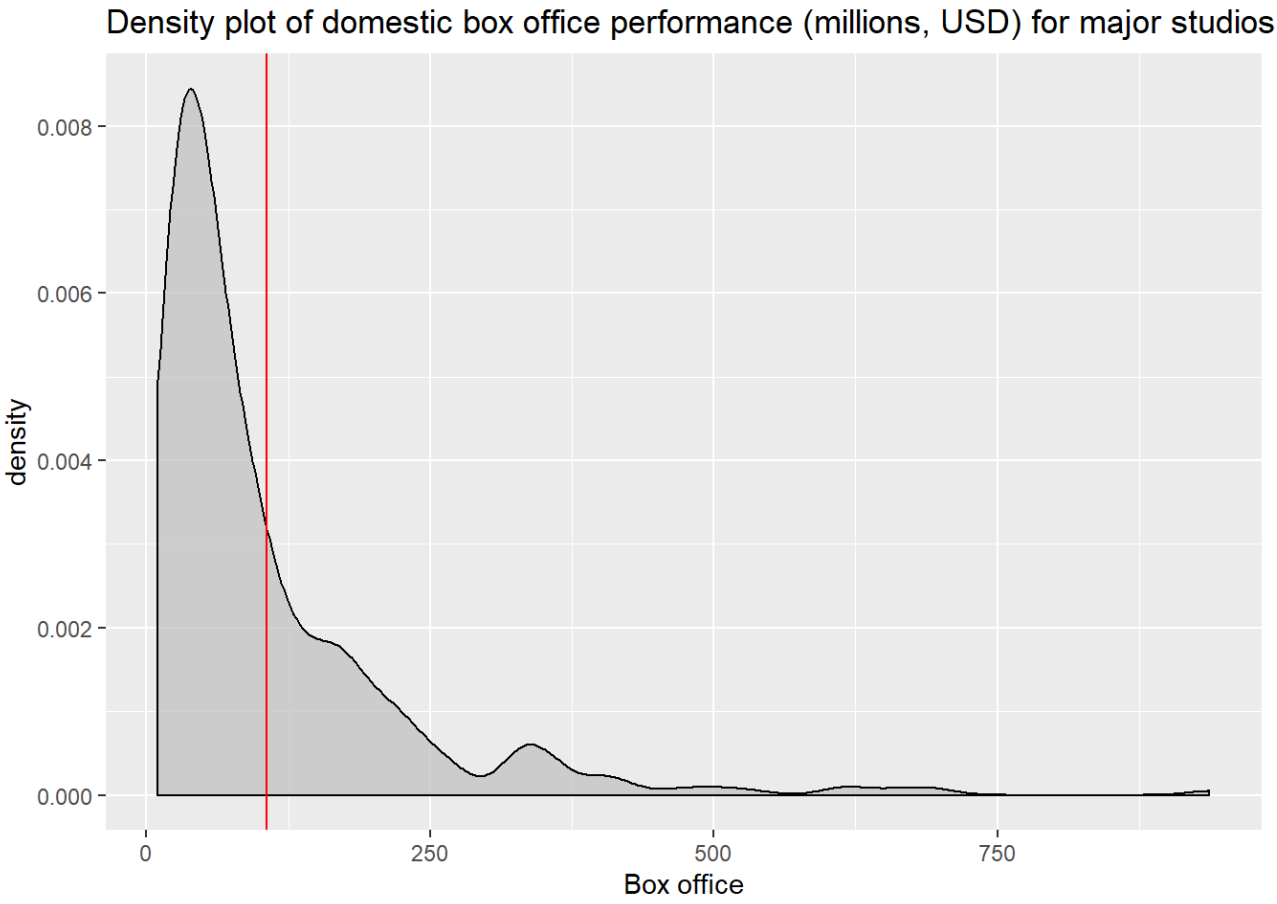
The third and final boxplot (displayed below) visualizes the distribution of domestic box office performance for minor studios. Fifty percent of these films earned between 11.635 and 23.165 million USD, a fairly narrow range. Very few of these movies made more than 45 million USD, with Creed II earning the most at almost 120 million.

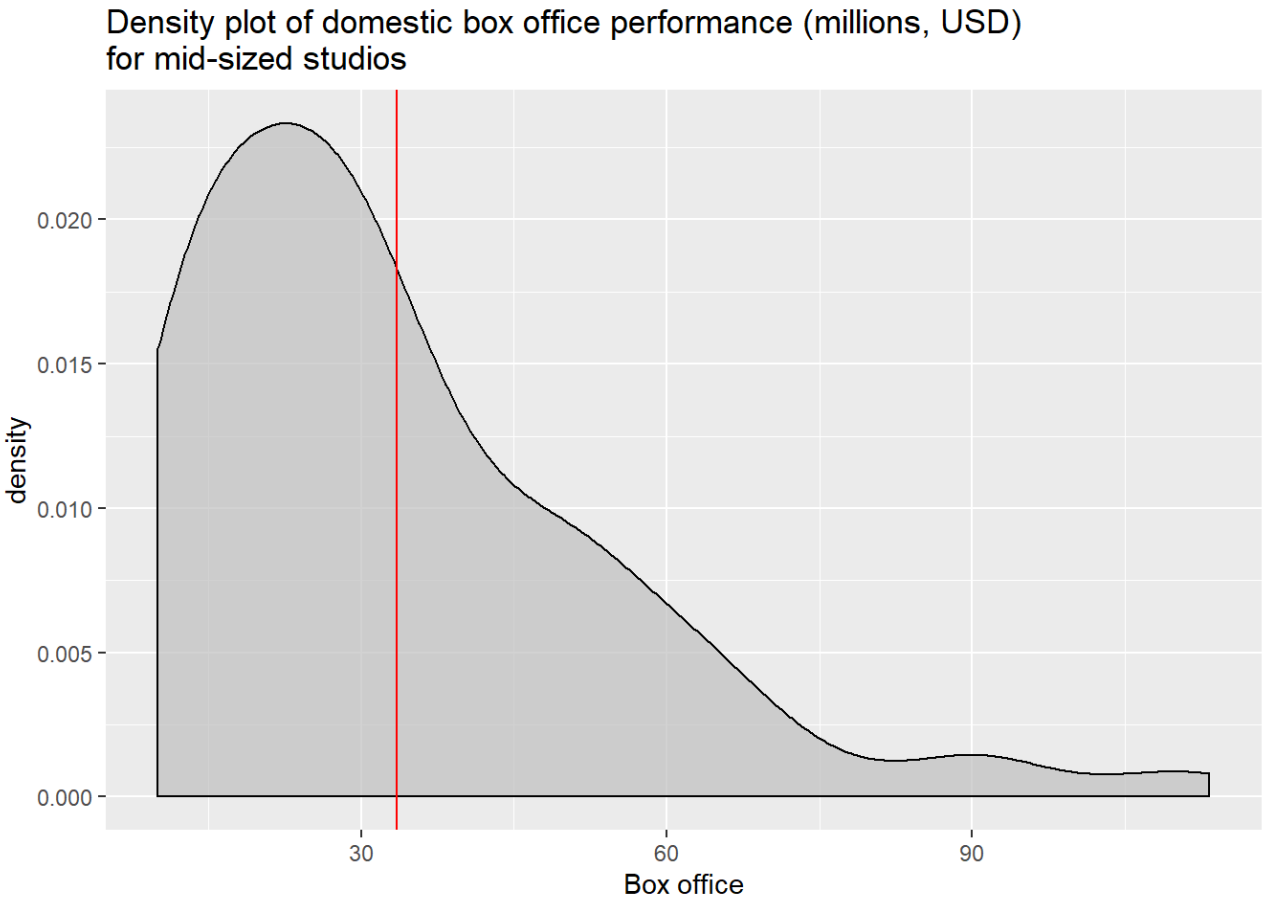


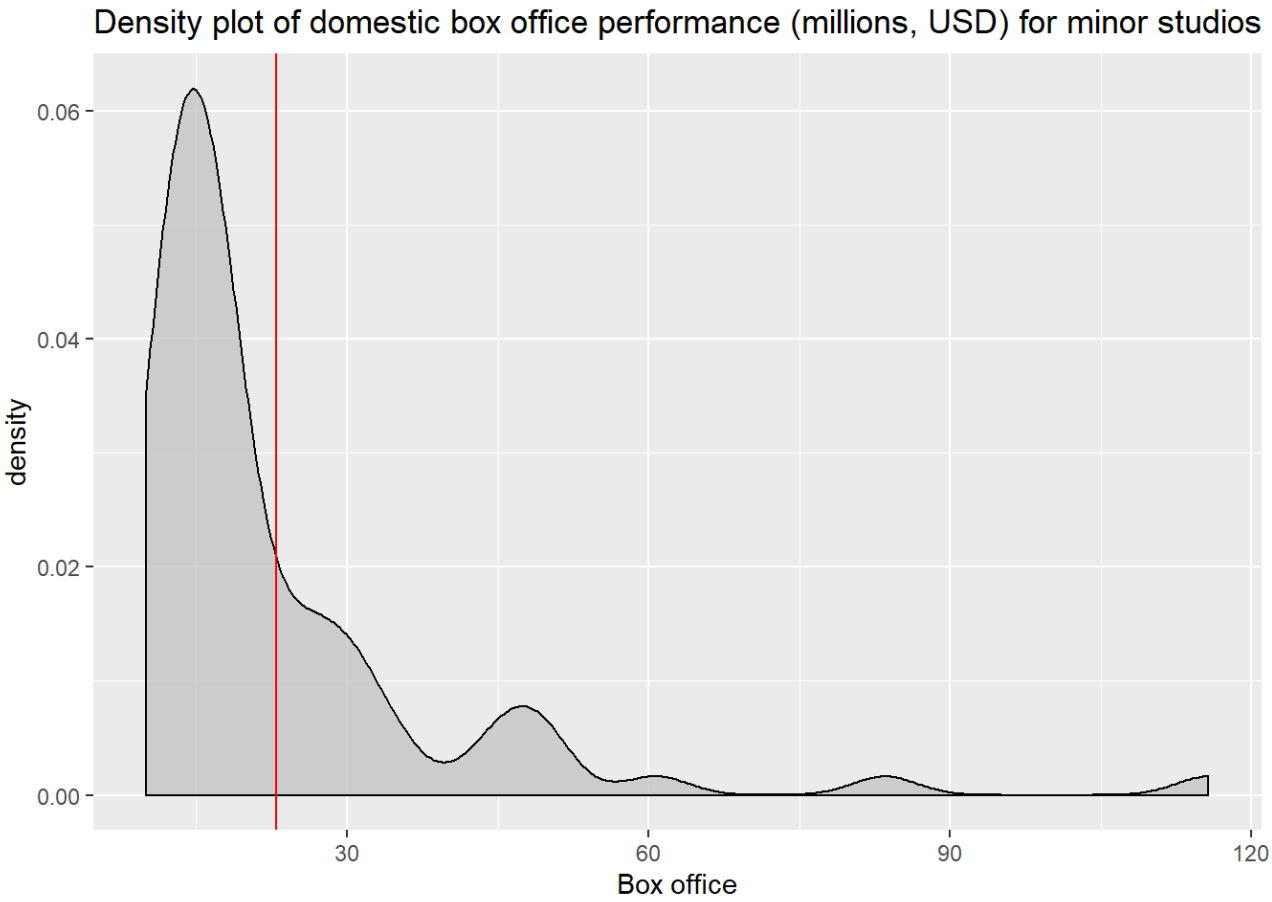
Based on the interquartile ranges (i.e., the range of the middle 50% of values, centered around the median value) for these three categories, it's clear that movies from all three categories are capable of making relatively little money (i.e., less than 20 million USD). However, the three categories have substantially different upper ends of their interquartile ranges, with the difference between that of the major studios and the minor studios the most dramatic. It appears, then, that movies released by major studios are much more likely to make a lot of money than those released by mid-size or minor ones. At the same time, though, major studio films frequently make as much as films from smaller studios do.

The next three (density) plots provide an alternative way of looking at the distribution of box office performance by studio category. Density is a measure of proportion for a given value along the x-axis - in this case, domestic box office performance. The higher the density at a given point, the more movies earned that much money. The red lines indicate the mean domestic box office performance for that studio category.

The first plot shows that the distribution for major studios is highly skewed upward, with most movies earning less than the mean but a number of very large outliers pulling the mean to the right. A similar (though less extreme) pattern is apparent for minor studios, whereas mid-sized studios have a less skewed distribution.


Density plot of domestic box office performance (millions, USD) for major studios
density
0.008
0.006
0.004
0.002
0.000
0
250
500
750
Box office


Density plot of domestic box office performance (millions, USD)
for mid-sized studios
density
0.020
0.015
0.010
0.005
0.000
30
60
90
Box office



Indeed, the skewness scores of these three distributions are 3, 1.38 and 2.79 for minor, mid-sized and major studios, respectively. A non-skewed distribution should have a score close to zero, and, as a rule of thumb, no more than 1 or less than -1. All three of these studio categories exhibit substantial skewness. Since we are using domestic box office performance (in millions of USD) as the outcome variable to predict, we should address this skewness since it can interfere with our analysis. Specifically, films with very high box office returns can excessively influence the estimation of the coefficients we use to develop predictive models. To mitigate this skewness, I log this outcome variable when using it in any models. For visualization purposes, however, I continue to use the original version of the domestic box office performance variable since it is much more easily interpretable.

Numeric predictors (summary)

The following table displays correlations between all numeric predictors in the dataset. The names of these variables should be self-explanatory, though I should note that "(adjusted)" means that the variable takes into account only those movies that came out in the years prior to the one in which the film was released. This adjustment seems justified since the correlation between the non-adjusted and adjusted average director RT critics scores is 0.856154 and the correlation between the two versions of the lead actor RT score is 0.7768476. These correlations are high, but not perfect, suggesting some degree of divergence between adjusted and non-adjusted RT scores. In addition, considering the reception of only those movies that were released prior to the film in question makes theoretical sense, as we would not expect the reception of future movies to influence how well a movie did in the past.

Correlations, all numeric variables

	Total boxoffice (millions)	Rotten Tomatoes score (critics)	Rotten Tomatoes score (audience)	Average director RT score (adjusted)	Average lead actor RT score (adjusted)	Number of theaters on release	Maximum number of theaters	Opening as a percentage of total boxoffice	Opening boxoffice (millions)	year

Total boxoffice (millions)		0.26	0.3	0.1	0.08	0.47	0.58	0.02	0.94	0.03
Rotten Tomatoes score (critics)	0.26		0.75	0.42	0.32	-0.24	-0.11	-0.49	0.17	0.09
Rotten Tomatoes score (audience)	0.3	0.75		0.23	0.17	-0.23	-0.09	-0.48	0.19	0.01
Average director RT score (adjusted)	0.1	0.42	0.23		0.39	-0.16	-0.1	-0.28	0.04	0.06
Average lead actor RT score (adjusted)	0.08	0.32	0.17	0.39		-0.12	-0.06	-0.19	0.06	0.04
Number of theaters on release	0.47	-0.24	-0.23	-0.16	-0.12		0.89	0.64	0.56	0.04
Maximum number of theaters	0.58	-0.11	-0.09	-0.1	-0.06	0.89		0.4	0.59	0.06
Opening as a percentage of total boxoffice	0.02	-0.49	-0.48	-0.28	-0.19	0.64	0.4		0.23	-0.03
Opening boxoffice (millions)	0.94	0.17	0.19	0.04	0.06	0.56	0.59	0.23		0.03
year	0.03	0.09	0.01	0.06	0.04	0.04	0.06	-0.03	0.03	

The correlation table reveals some interesting patterns. Total box office is moderately correlated (i.e., absolute value is greater than 0.30) with the RT audience score, the number of theaters on release, and maximum number of theaters. Opening box office is strongly correlated with total box office (i.e., absolute value is greater than 0.7). When selecting numeric variables to include in my predictive models, I only want to include those variables that have at least a moderate correlation with my outcome variable: total box office. However, I also want to only include variables that can be useful predictors, that is, prior to the film's release. Of course, the RT audience score does not fit this criterion, but if we relax this rule just a bit and say that any variables whose values are determined prior to the first evening of a film's release can be useful predictors. To see if final RT scores (which the two RT variables measure) are good proxies for the reception a movie receives (by both critics and audiences) before the movie's first evening showings, I took a random sample of 21 films from my dataset and used the Internet Wayback Machine to obtain RT scores the morning of the film's release (i.e., prior to the its first evening showings) and compared those to the final

RT scores obtained in February 2020. The following two tables display correlations between the main RT variables and their respective alternatives, i.e., films' scores at different snapshots in time - the day of release, 3 days after release, 14 days after release and 28 days after release. The extremely high correlations suggest that final RT scores are on average good measures of a film's pre-release or release day reception.

(NOTE: the days after release markers are approximate. Most movies did not have archived pages for every relevant day. I picked the closest day to each desired interval when possible, though some movies did not have archived pages anywhere close to the desired interval. These intervals I have left blank.)

Finally, I relax my criteria to also allow for the inclusion of the RT critics score in predictive models, even though it has a correlation of just 0.26 with total box office. Since one of this project's motivating questions is how much impact Rotten Tomatoes scores have on a film's success, I think including both variables in the models would be substantively interesting.

Correlations, Rotten Tomatoes scores (critics), day of release and times after

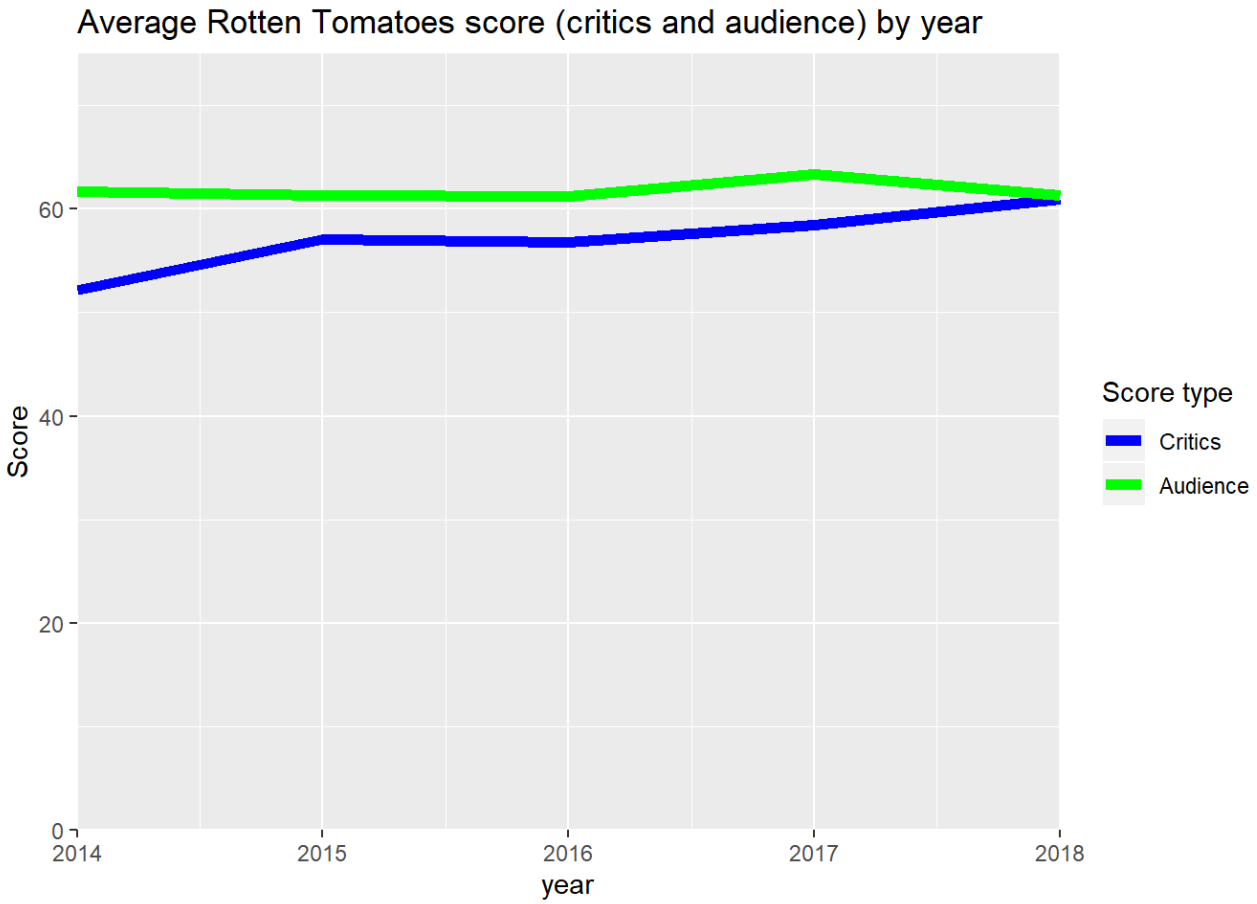
	rt.critics	rt.critics.day1	rt.critics.day3	rt.critics.day14	rt.critics.day28
rt.critics		0.9986	0.9932	0.9961	0.998
rt.critics.day1	0.9986		0.9978	0.998	0.9982
rt.critics.day3	0.9932	0.9978		0.9993	0.9969
rt.critics.day14	0.9961	0.998	0.9993		0.9992
rt.critics.day28	0.998	0.9982	0.9969	0.9992	

Correlations, Rotten Tomatoes scores (audience), day of release and times after

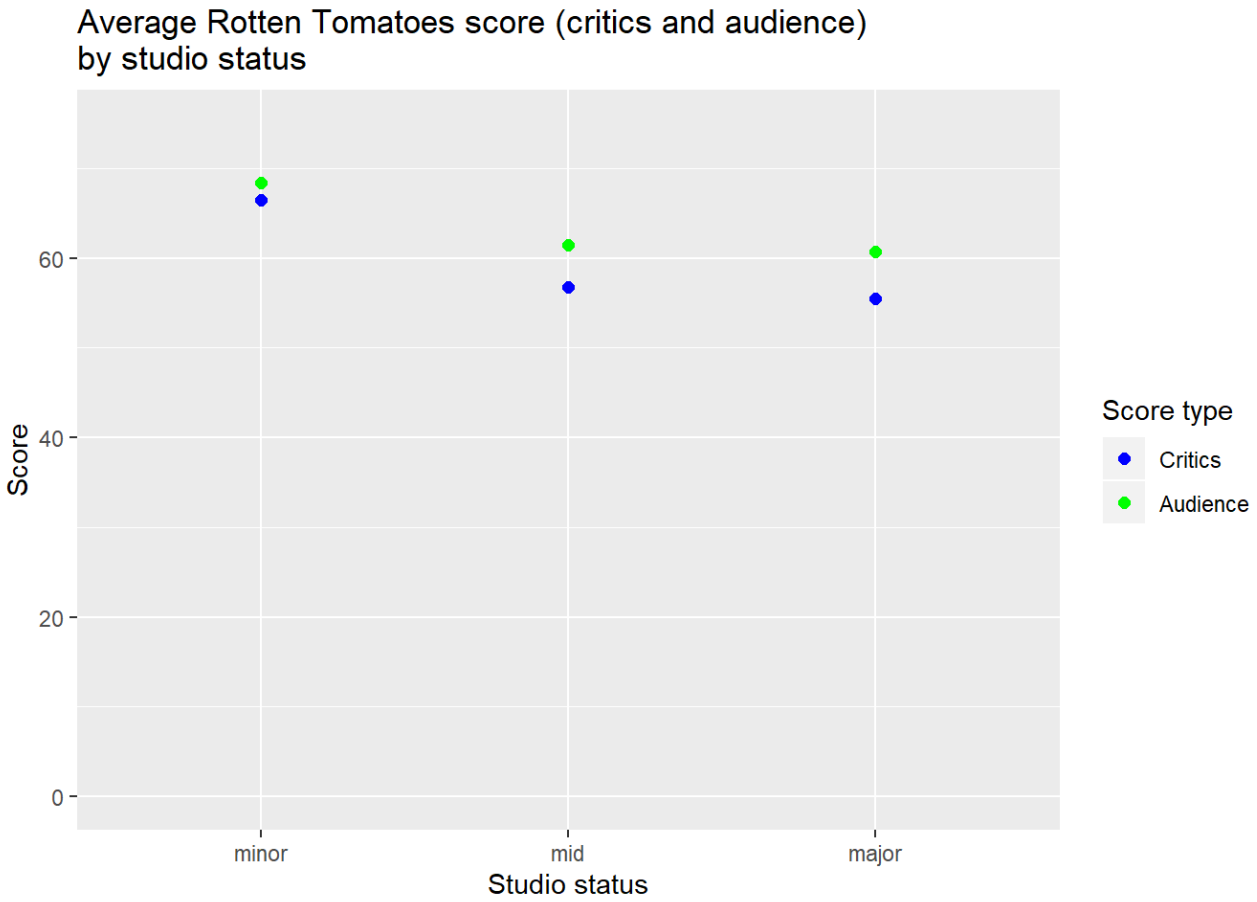
	rt.audience	rt.audience.day1	rt.audience.day3	rt.audience.day14	rt.audience.day28
rt.audience		0.9704	0.9554	0.9807	0.9869
rt.audience.day1	0.9704		0.9879	0.9864	0.9834
rt.audience.day3	0.9554	0.9879		0.9902	0.9857
rt.audience.day14	0.9807	0.9864	0.9902		0.9918
rt.audience.day28	0.9869	0.9834	0.9857	0.9918	

Rotten Tomatoes scores

This section further explores patterns relating to the two variables that measure Rotten Tomatoes scores. The first plot below displays how RT scores have varied over time within my dataset's scope, i.e., the years 2014 to 2018. The average audience score has stayed largely unchanged at slightly higher than 60%, meaning that the average movie has received a "Fresh" rating from audiences. The average critics score, on the other hand, has slightly but steadily increased from a little over 50% in 2014 to just over 60% in 2018. Across all years, the average critics score is lower than the average audience score.

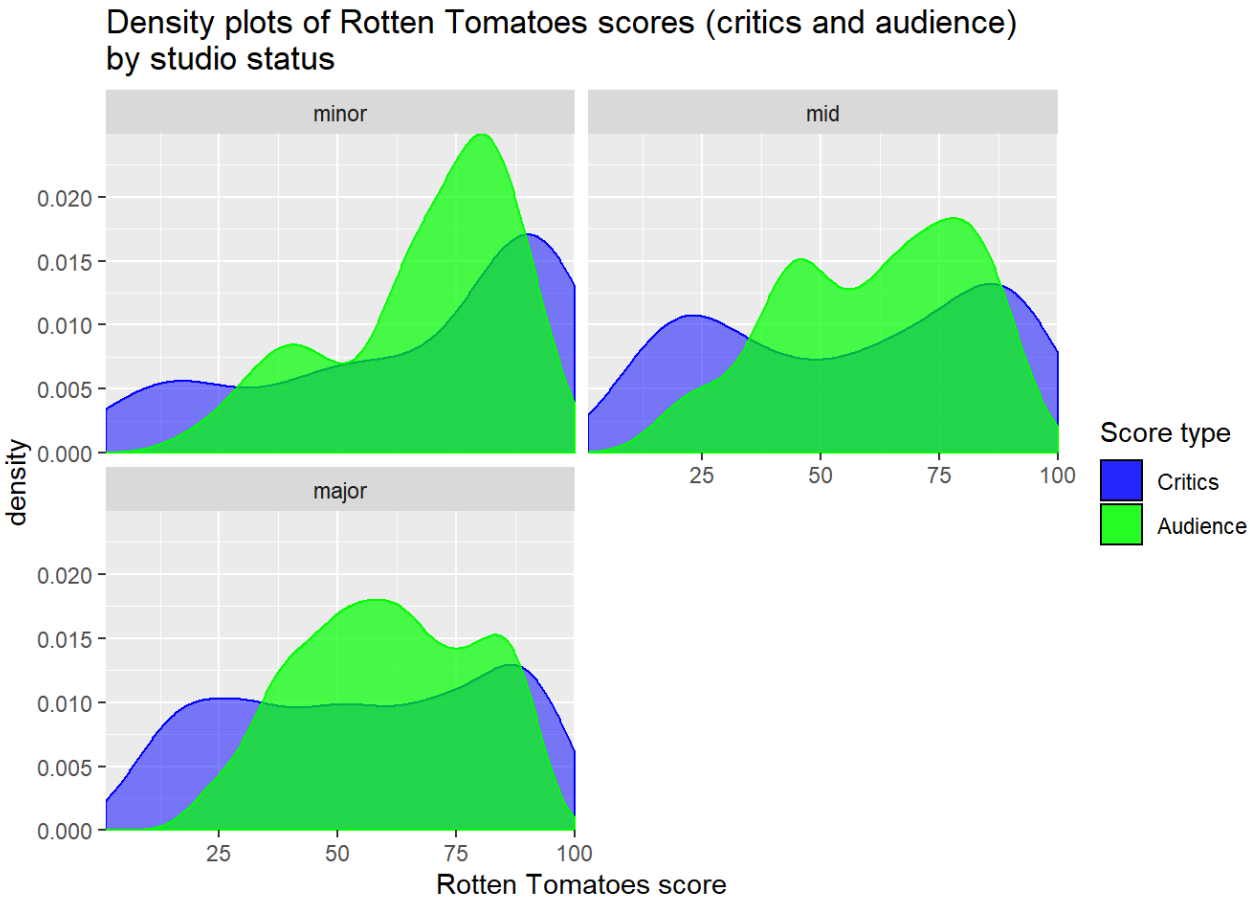


The second plot displays average RT scores for each studio status category: minor, mid-sized and major. Once again, the average audience score is consistently higher than the average critics score. Notably, though, films from minor studios tend to be better received than their mid-sized and major studio counterparts and the gap between audience and critics scores is narrower for this group than the other two. There seems to be little difference in average reception between mid-sized and major studio films.



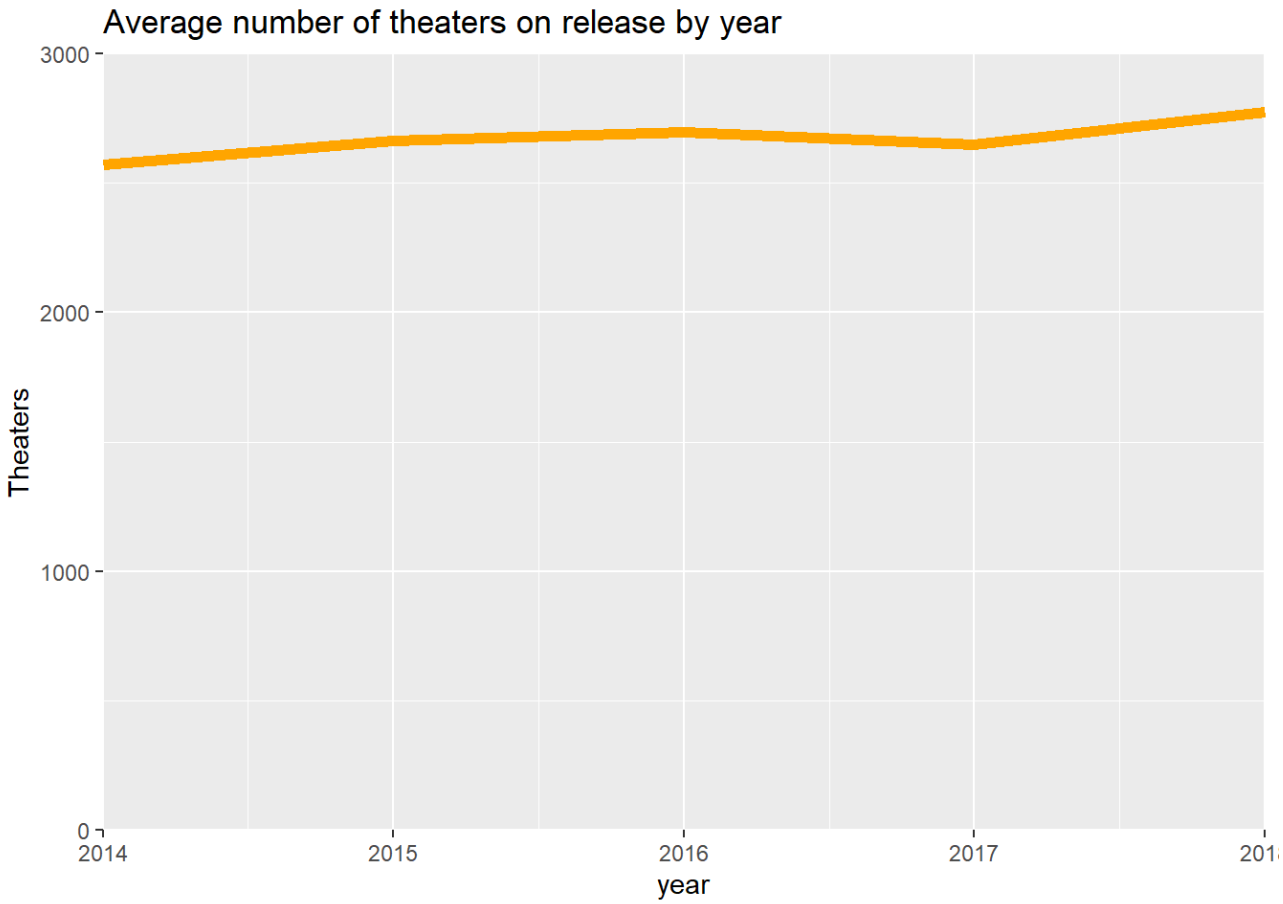
The third plot looks at the interaction between RT score and studio status in greater detail. This set of density plots shows a couple of interesting patterns. First, RT audience scores seem to be more densely distributed than critics scores, with the density curves higher in the middle of the distributions for audience scores and higher at the tails for critics scores. This suggests audiences tend to be not as strongly denigrating or acclamatory as critics are. Second, all three plots are to some extent double peaked. For instance, audiences seem more likely to rate a movie a 45% or 75% score than a 55% for mid-sized studio releases (bear in mind that these percentages represent the portion of audiences/critics who have favorably reviewed a movie). This double-peakedness is more of an issue for the critics score variable because the peaks are further apart. The relative similarity in density height between the peaks for the mid-sized and major studio categories is notable as well (for audience scores, only the mid-sized category has peaks with similar heights). The spread and roughly equal height of these peaks for the critics score variable suggests that it may be a weak predictor for other variables, a possibility that is supported by the weaker correlation between the critics score and box office variables. The double-peakedness is also present for the audience score variable, but appears less severe.

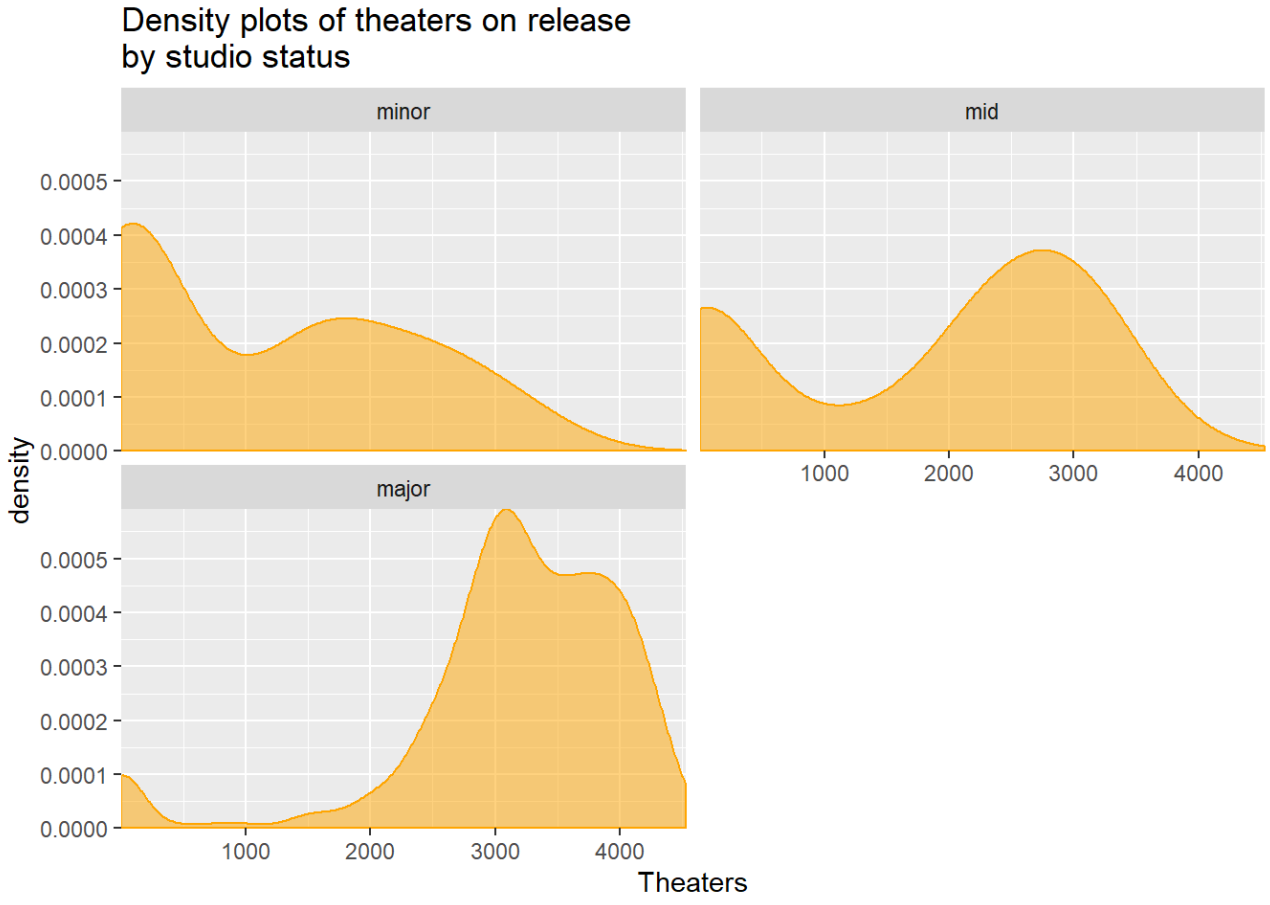
In addition, the notable variations in RT score distribution across the three studio categories suggests that interacting these numeric variables with the categorical studio one may be sensible when building a predictive linear model.



Number of theaters on release

The next two plots provide visual information on the third numeric variable I use when developing predictive models for box office success: the number of theaters a film is shown in on release. The first plot shows the average number of theaters on release for each year between 2014 and 2018. There is a modest, but steady increase across these years. The second plot displays the density distribution of the theaters on release variable for each studio status category. The density plots for minor and mid-sized studio films are similarly shaped, though the highest peak for the mid-sized category is further to the left on the x-axis, indicating that mid-sized studio films tend to be released more widely than minor studio ones. There are still quite a few mid-sized studio films that are released in very few theaters, however. Films in the major studio category seem to be overwhelmingly released in a large number of theaters. These patterns suggest we should test the interaction between the theaters on release variable and the studio status variable when constructing predictive models.



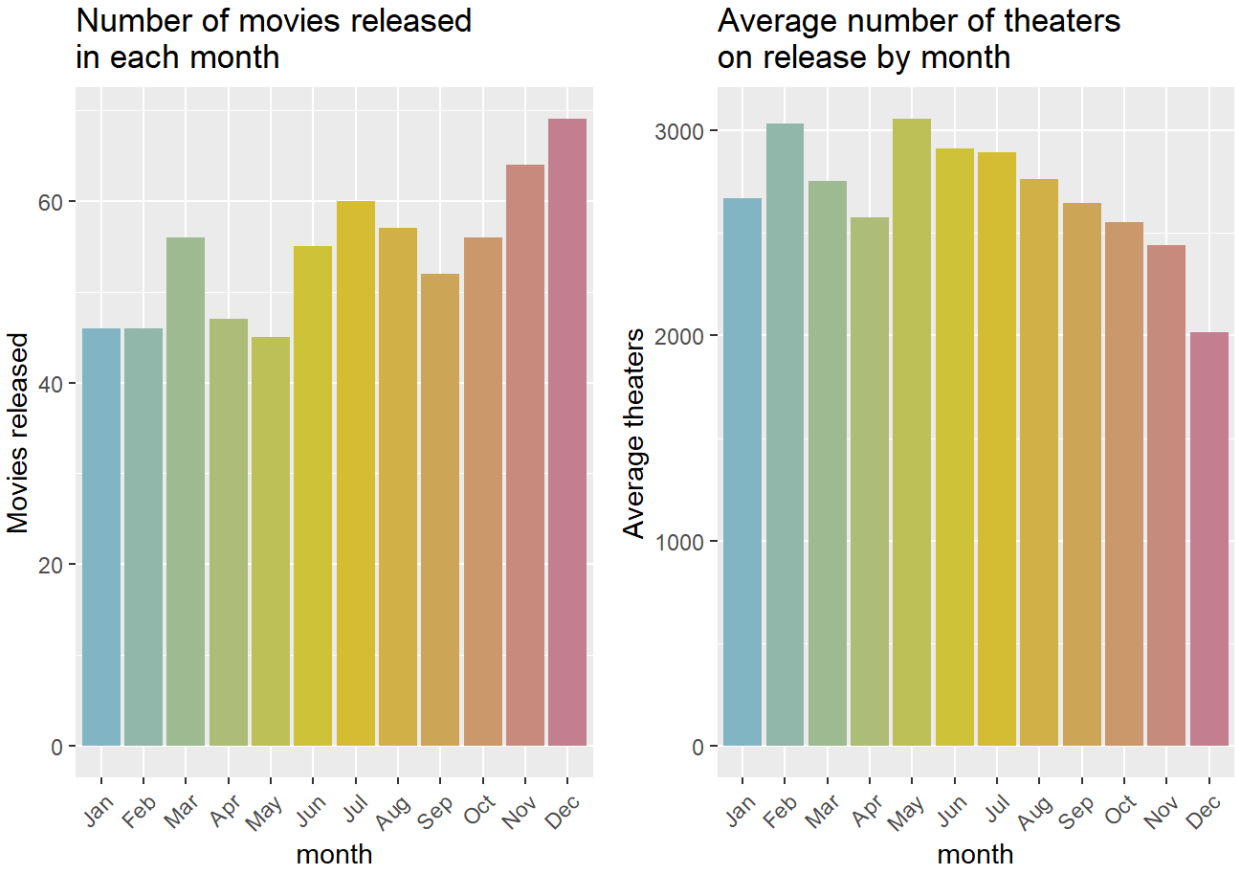


Month of release

This subsection explores patterns relating to which month a film is released in. The first five plots visualize summary statistics for each of the twelve months.

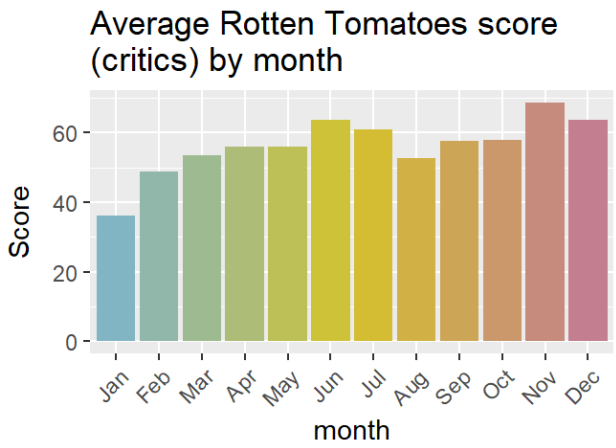
The plot below on the left shows the number of movies released in each month. We can see that the top three months are July, November and December. March, June, August and October are also popular months for movie releases. These results make intuitive sense as the summer and holiday seasons are considered to be the most popular times for people to go to the movies. October is known as a month for smaller movies with potential widespread appeal to be released and March is increasingly becoming a time for major studios to release tentpole movies.

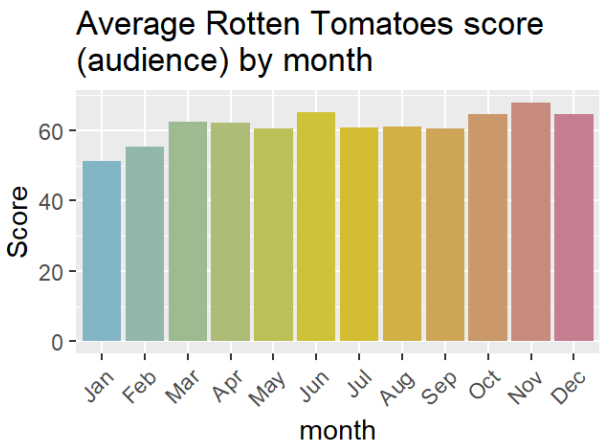
The plot on the right shows the average number of theaters on release for each month. The widest average releases occur in February, May, June and July. The latter three months make intuitive sense because that is when major blockbusters are usually released. February is rather surprising, though, but can be explained by the large number of movies that released in very large numbers of theaters (i.e., more than 3710, or the 80th percentile for the theaters on release variable). Such films include The Lego Movie and its sequel The Lego Batman Movie, Black Panther, and the Fifty Shades movies. Of the six movies that released in February in more than 3700 theaters, five were released after 2016. A disproportionate number of movies that are released in February have wide releases, especially in recent years. Conversely, a disproportionate number of films that are released in December have very small releases. This finding likely captures the tendency for studios to release Oscar-bait films very late in the year, so as to release them as close to the awards ceremonies as possible.

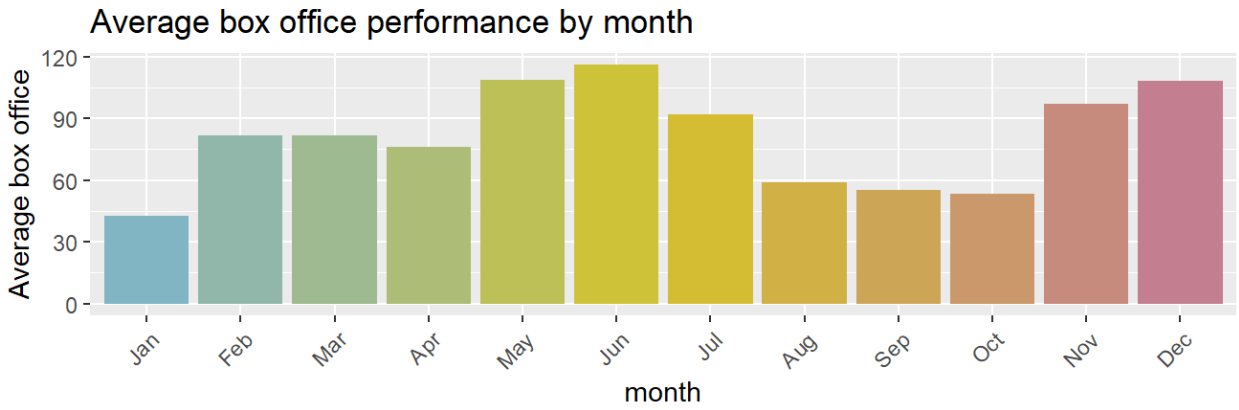


The next three plots visualize the relationship between the two Rotten Tomatoes score variables and box office performance and month of release. The two Rotten Tomatoes plots are roughly similar, with the most highly rated movies on average releasing in June, November and December. Since the latter two months often have a lot of Oscar-ambitious films released, it is not surprising they would have two of the highest average scores across all months. June is an interesting result, since this is right in the middle of blockbuster season, when many tentpole movies that may do well at the box office but not so well with critics are released. It is possible that studios tend to release the “best” films from their annual stables of potential blockbusters earlier in the summer season. Regardless of the reasons for why these variations across months are observed, these five plots suggest that it may be beneficial to interact the month variable with at least some of the numeric variables when building

predictive models.



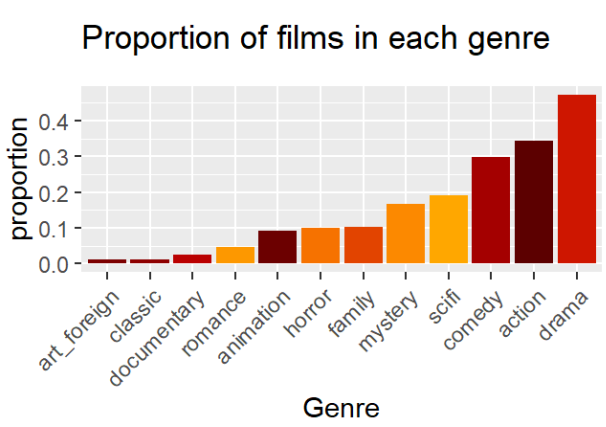


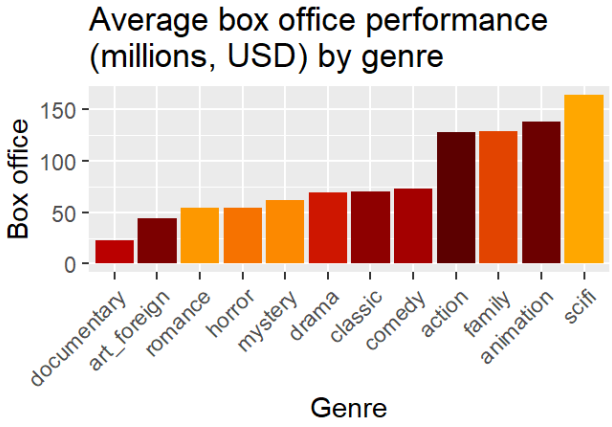


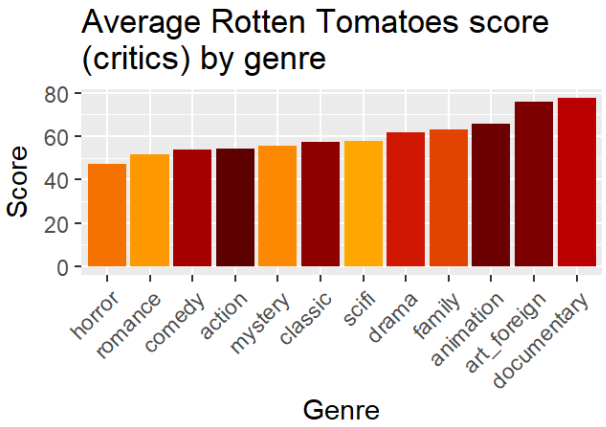
Genre

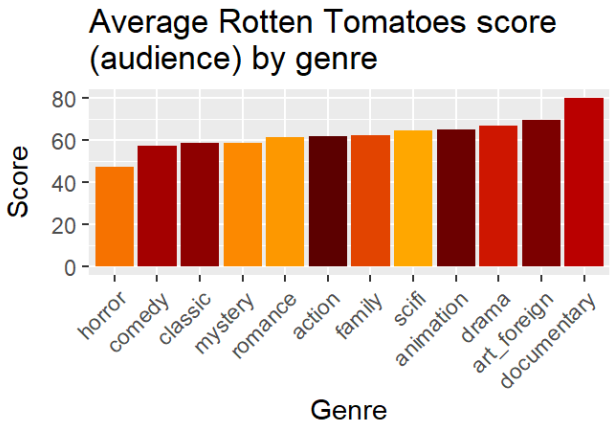
The last set of variables I want to explore before beginning the process of building predictive models is the one that measures genres. Unlike the categorical variables studio status and month of release, genre is not mutually exclusive. Therefore, I created a series of binary genre variables, corresponding to the genres used by Rotten Tomatoes. The genre classifications for each film were also gleaned from the RT website.

Four plots are displayed below that visualize the relationships between genre and number of films, box office performance, RT critics score and RT audience score. Drama is the most frequent genre category, followed by action and comedy. However, scifi, animation and family films perform the best at the box office, on average. Documentaries and art/foreign films tend to be the best reviewed, by both critics and audiences. The variation in average box office performance by genre suggests that considering these genre variables when building predictive models could be advantageous. While there are differences across genre in terms of RT score, these differences seem fairly minimal, with the possible exceptions of documentaries and horror films. There does not seem to be much support for spending time on determining which interactions would be valuable for model-building.









Predicting domestic box office performance

Given the above data exploration, I have decided to focus on the following predictors when constructing models for predicting box office performance: Rotten Tomatoes critics score, Rotten Tomatoes audience score, number of theaters on release, year, studio status, month, and genre. Before working towards a final predictive model, however, I explore the linear relationship between RT scores and domestic box office. Because of the extreme skew of the domestic box office variable, I first log it before including it in any models.

Bivariate models: Rotten Tomatoes scores

The following tables show the results of the OLS regression analysis of the relationship between the RT score variables and the outcome variable (i.e., logged domestic box office performance in millions of USD). The RT critics model indicates that there is a positive linear relationship between the predictor and the outcome variable. For what it's worth, both the intercept and coefficient estimate are statistically significant by conventional standards. We obtain similar results for the model estimating the relationship between the RT audience predictor and the outcome. The coefficient for the RT audience variable is almost double in magnitude compared to that of the RT critics variable. Given that the two variables are measured on the same scale (0-100%), this indicates a stronger relationship between the audience score and the outcome than between the critics score and the outcome.

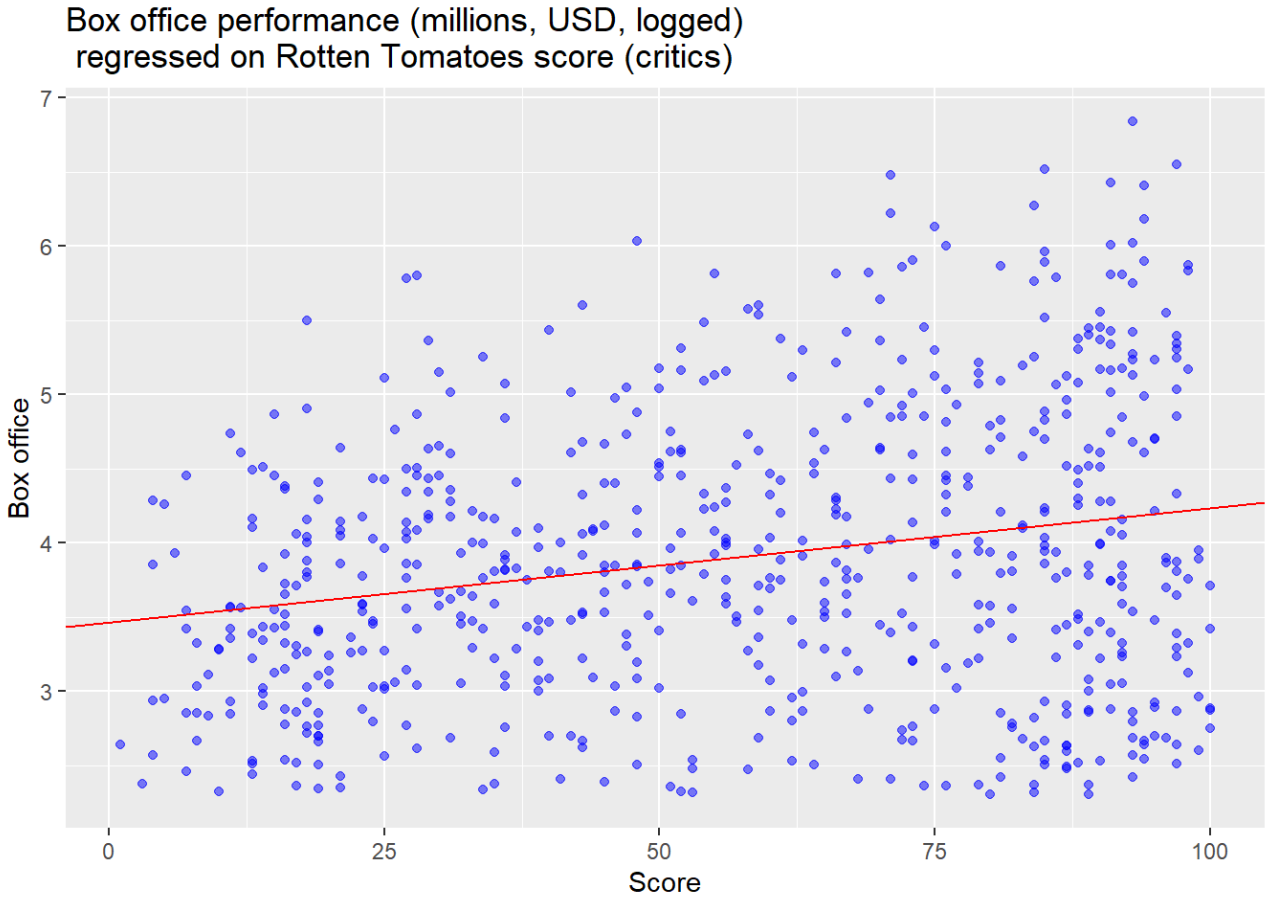
Results for RT critics model

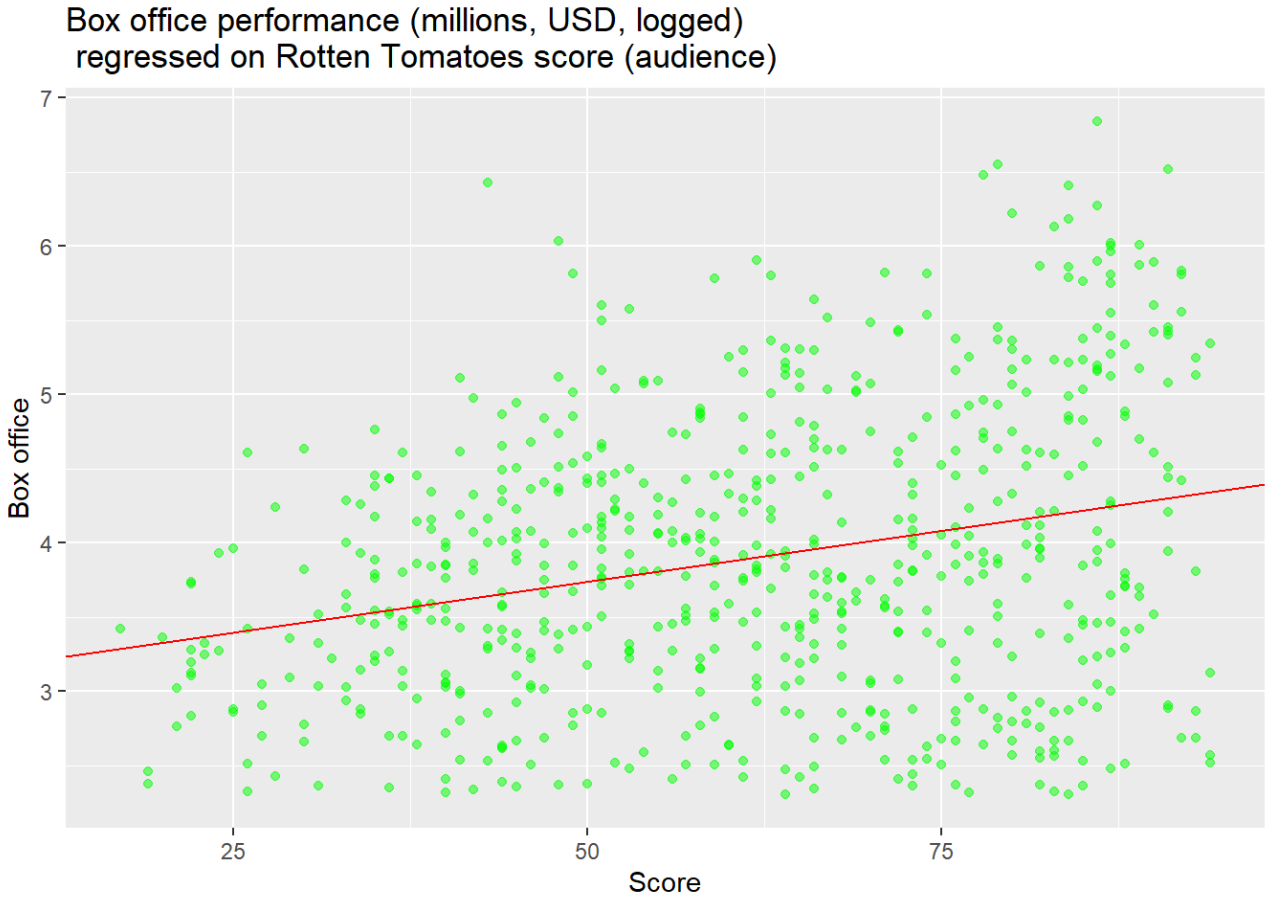
term	estimate	std.error	statistic	p.value
(Intercept)	3.4631	0.0821	42.1862	0
rt.critics	0.0077	0.0013	5.9916	0

Results for RT audience model

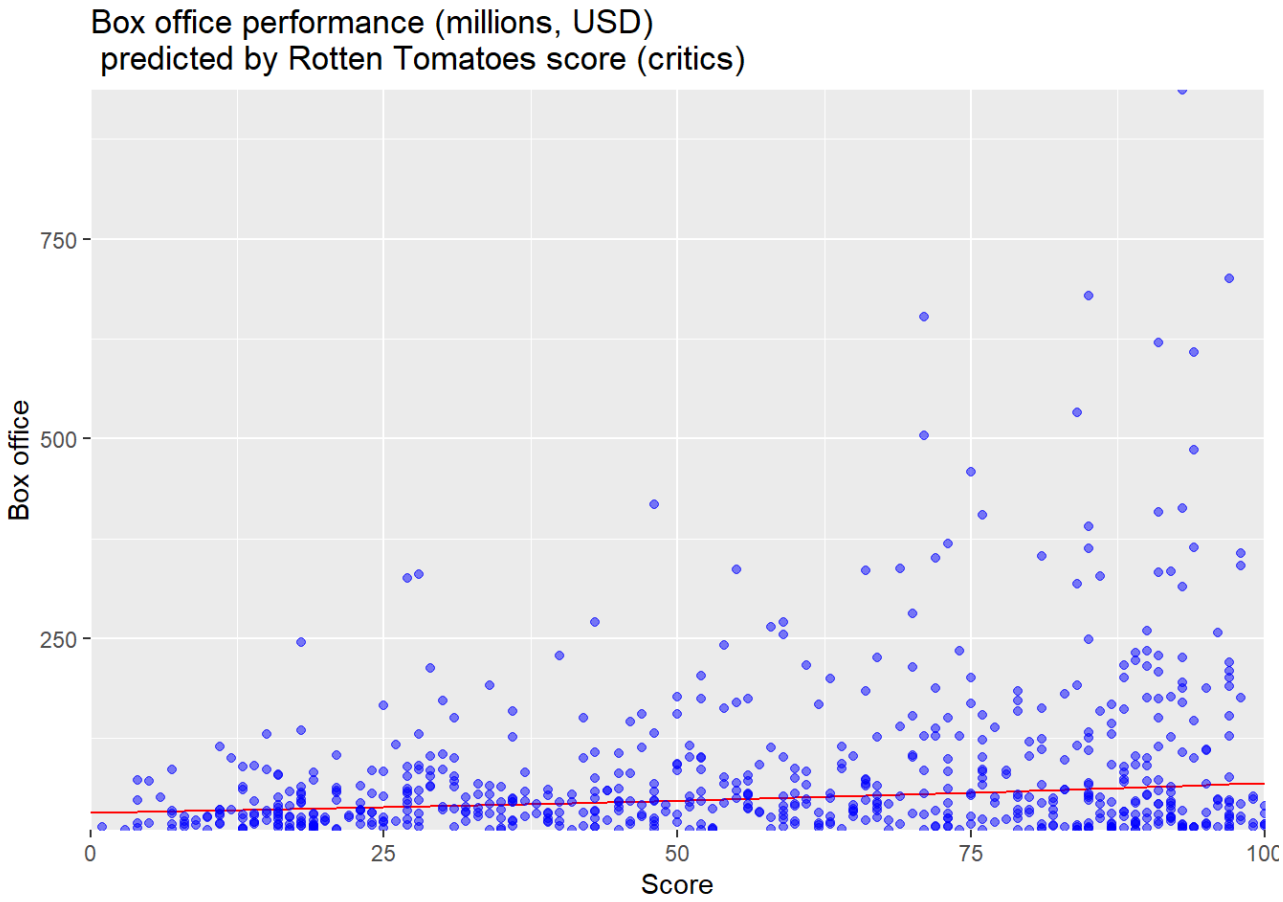
term	estimate	std.error	statistic	p.value
(Intercept)	3.0565	0.1239	24.6604	0
rt.audience	0.0137	0.0019	7.1435	0

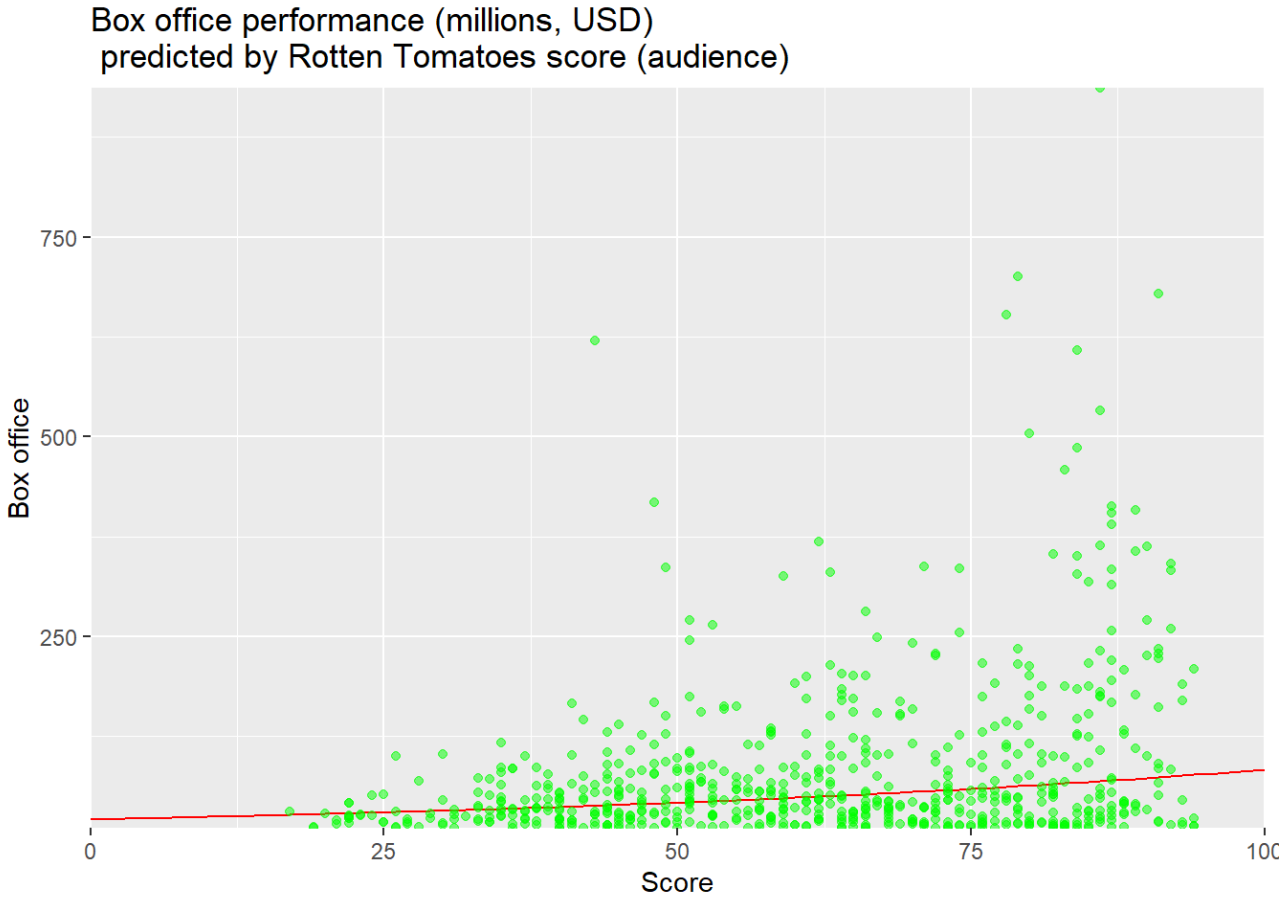
The following scatter plots show, however, that neither variable has an especially strong relationship with box office performance. There is quite a lot of variation around the regression line in both plots, indicating that neither is a good predictor of box office performance.





The next two plots use the appropriate bivariate regression model to compute predicted values for the outcome variable (i.e., box office performance) and then transform those predictions into the original outcome scale (unlogged USD in millions). In both plots, we can see that the RT score variables underestimate the performance of dozens of films. Once again, we see that, at least on their own, these variables do not do a good job of predicting box office performance.





Multivariate models

Rather than make a bivariate model for each predictor, I instead present below a single model that includes all predictors I have selected. A second model includes the same predictors but with interaction terms. I tried a number of different interactions between numeric variables and categorical ones, and determined that three are worth including due to their effects on the model's r.squared and AIC statistics. R.squared represents the variation in the outcome variable that is explained by the set of predictors, whereas AIC (Akaike's Information Criterion) considers both the quality of the model in fitting the data and the number of predictors. The quality of a model is considered in relation to how much information about the data-generating process is lost or gained when adding or subtracting predictors. Generally speaking, a model that has the same level of quality as another but uses fewer predictors will have a lower (and thus better) AIC score.

The three interactions I settled on are: Rotten Tomatoes critics score and studio status, theaters on release and studio status, and theaters on release and month.

The following table presents the two OLS models side-by-side. For each term in the regression, the coefficient (standard error) is provided. Stars indicate the degrees of statistical significance at conventional standards. The coefficients are not standardized and so must be interpreted on the scales of the respective variables. The difference in coefficient size between RT critics and RT audience is even more substantial than it was when we looked at the two predictors' relationships with the outcome variable independently. In model two, the inclusion of interaction terms seems to significantly weaken the relationship between the RT critics score variable and the outcome. Conversely, the coefficient for theaters on release (dom.theaters.open) changes relatively little when including interaction terms, suggesting this predictor has a more robust relationship with the outcome than RT critics does. Interestingly, the coefficient for the month of February changes sign when we include interaction effects, probably because, as noted above, February has a disproportionate number of films with very wide releases (greater than 3700 theaters on opening day). December has a substantial and robust effect in both models, as do the action and scifi genres. The major category of the studio status variable reverses sign (from positive to negative) when including interaction terms, perhaps because what that category is really capturing is how wide of a release a film has.

	Model 1	**Model 2**
(Intercept)	53.365832	64.162473
	(33.969811)	(32.800035)
rt.critics	0.004299***	0.000261
	(0.001285)	(0.002565)
rt.audience	0.017316***	0.015954***
	(0.001980)	(0.001946)
dom.theaters.open	0.000435***	0.000315**
	(0.000028)	(0.000120)
year	-0.026026	-0.031127
	(0.016851)	(0.016271)
monthFeb	0.043242	-1.669422**
	(0.123861)	(0.570538)
monthMar	0.111074	0.040238
	(0.118653)	(0.342136)
monthApr	0.003771	-0.268014
	(0.123928)	(0.349930)
monthMay	0.177825	-0.252475
	(0.127521)	(0.362056)
monthJun	0.180758	-0.287570
	(0.120187)	(0.352827)
monthJul	0.226996	0.041770
	(0.118125)	(0.338604)
monthAug	-0.043641	0.187486
	(0.116932)	(0.375706)
monthSep	-0.055866	-0.028280
	(0.119496)	(0.342692)

monthOct	-0.133023	0.106345
	(0.118685)	(0.343380)
monthNov	0.323190**	0.445377
	(0.117834)	(0.320767)
monthDec	0.561903***	0.802166*
	(0.117559)	(0.313241)
genreAction	0.126879*	0.118112*
	(0.057384)	(0.055210)
genreAnimation	0.012517	-0.063379
	(0.103106)	(0.099833)
genreArtForeign	0.008579	-0.039488
	(0.241705)	(0.234006)
genreClassic	-0.169716	-0.122385
	(0.227245)	(0.218697)
genreComedy	0.061899	0.077606
	(0.063609)	(0.061529)
genreDocumentary	-0.324608	-0.039914
	(0.168413)	(0.170908)
genreDrama	-0.076750	-0.063203
	(0.059047)	(0.057132)
genreHorror	0.161160	0.197687*
	(0.089220)	(0.086336)
genreFamily	0.060771	0.058516
	(0.096075)	(0.092303)
genreMystery	-0.010543	0.004429
	(0.066598)	(0.064636)
genreRomance	-0.103084	-0.070506
	(0.111004)	(0.107687)
genreSciFi	0.169808*	0.137606*
	(0.067535)	(0.066747)
studio.status.threecatmid	0.204525*	0.464838
	(0.089033)	(0.310177)
studio.status.threecatmajor	0.430219***	-0.414620
	(0.088770)	(0.267109)
rt.critics:studio.status.threecatmid		-0.002241
		(0.003341)
rt.critics:studio.status.threecatmajor		0.005520*
		(0.002737)
studio.status.threecatmid:dom.theaters.open		-0.000073
		(0.000088)
studio.status.threecatmajor:dom.theaters.open		0.000223**
		(0.000080)
dom.theaters.open:monthFeb		0.000570**
		(0.000191)
dom.theaters.open:monthMar		0.000022
		(0.000120)
dom.theaters.open:monthApr		0.000096
		(0.000124)
dom.theaters.open:monthMay		0.000122
		(0.000122)

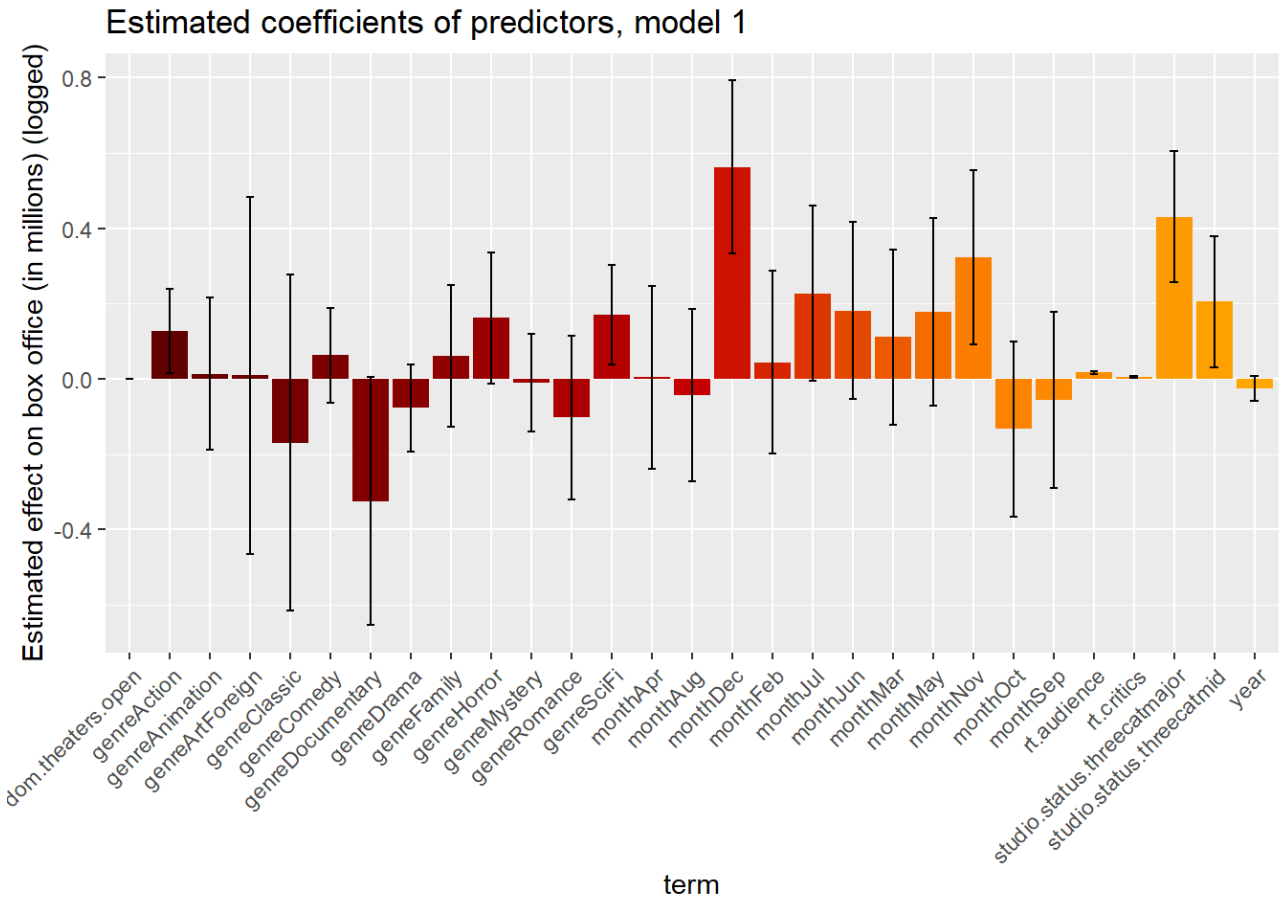
dom.theaters.open:monthJun		0.000165
		(0.000121)
dom.theaters.open:monthJul		0.000057
		(0.000117)
dom.theaters.open:monthAug		-0.000065
		(0.000131)
dom.theaters.open:monthSep		-0.000019
		(0.000121)
dom.theaters.open:monthOct		-0.000088
		(0.000123)
dom.theaters.open:monthNov		-0.000056
		(0.000114)
dom.theaters.open:monthDec		-0.000106
		(0.000113)
R^2	0.653459	0.689782
Adj. R^2	0.637276	0.667258
Num. obs.	651	651
RMSE	0.579288	0.554830

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

When looked at holistically, the two models both do a good job of predicting domestic box office performance, with the first having an r.squared of around 0.65 and the second having one of around 0.69. The root mean squared error (a measure of model accuracy) is also superior for the second model: 0.5353 versus 0.5658 for the first. The AIC for the second model is also superior to that of the first: 1125.83 versus 1167.91.

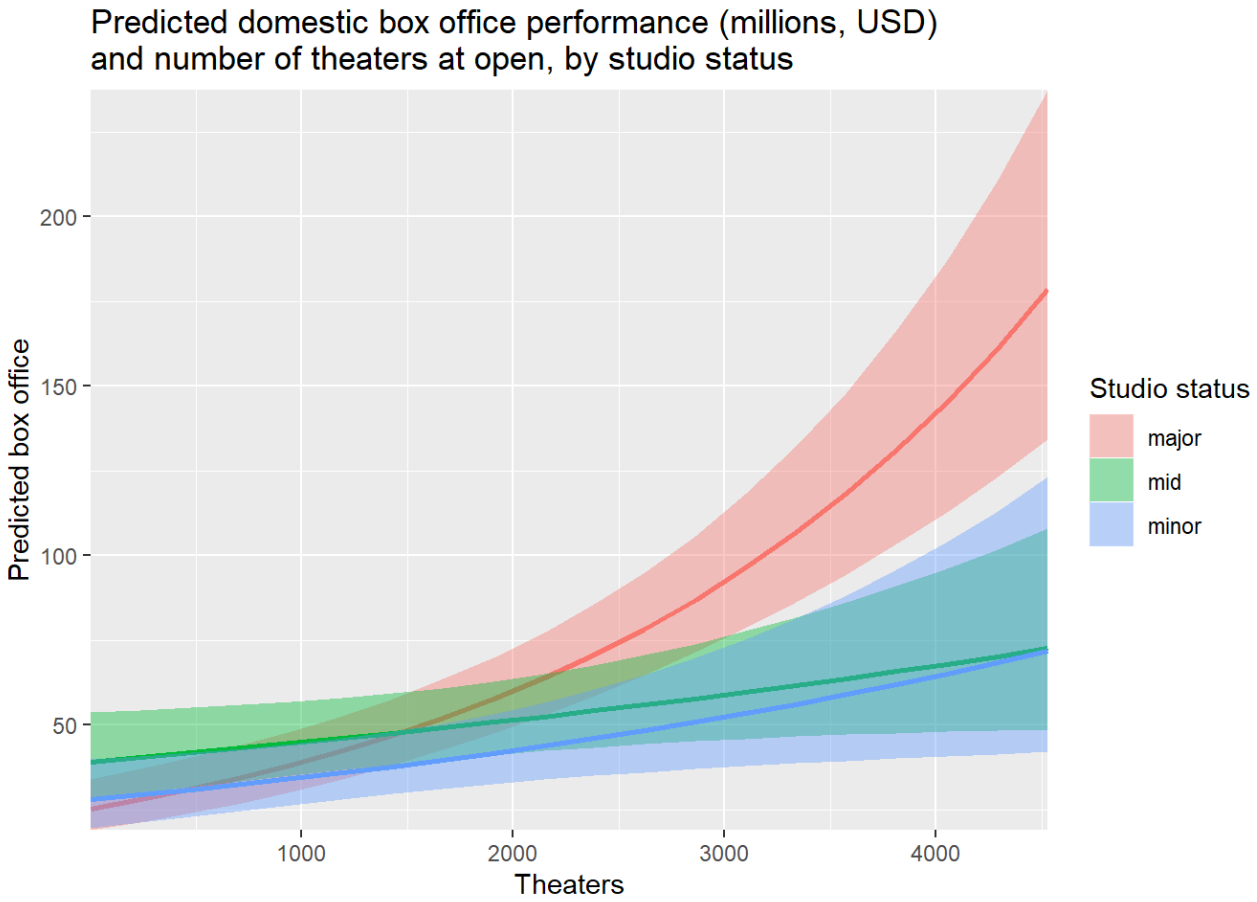
Visualizing the models

An alternative way of displaying the model results is to plot the point estimates of the variable coefficients along with their 95% confidence intervals. I do this just for model one, since the interaction terms in model 2 would make such a visualization overly complicated. Instead, I use other plots further below to display the most important features of model 2. Each bar represents the coefficient point estimate for each predictor included in model 1. The black bars represent 95% confidence intervals for these estimates. If the bar crosses 0, then the point estimate is not statistically significant according to conventional standards. More importantly, though, each black bar gives us the range of possible values that we could reasonably say encompasses the true value of the variable coefficient. Some “statistically significant” coefficients - like that associated with the month of December - have a fairly wide range of plausible values. Others, like the two Rotten Tomatoes score variables, mave much narrower ranges. We can be more confident in our assessment of the true impact of these latter variables on box office performance than for others. Some coefficients, like for the genre Documentary, have a very broad range of plausible values and are only barely “not statistically significant” by conventional standards.



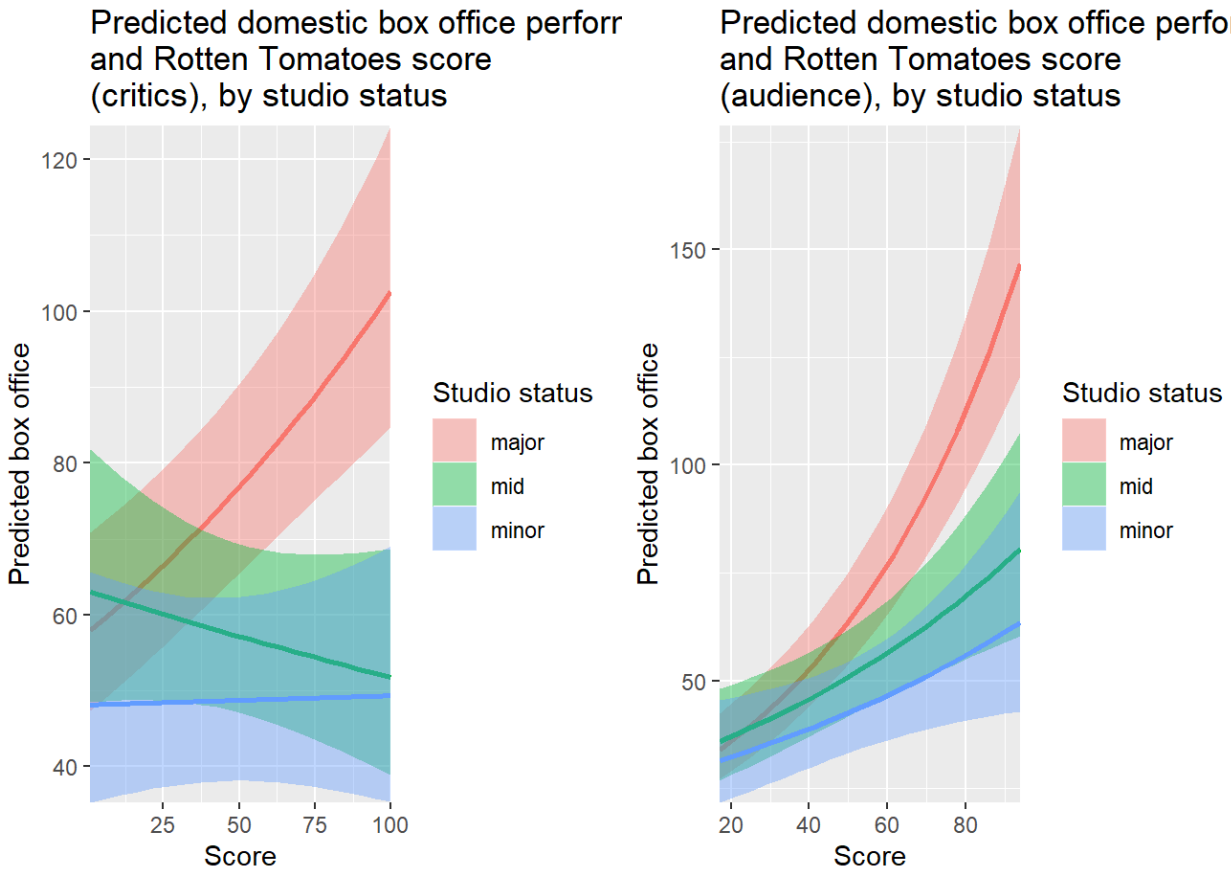
The next few plots take an alternative approach to displaying the uncertainty surrounding Model 2's coefficients, specifically those associated with one or more interaction term. The first plot visualizes the interactive relationship between studio status and number of theaters on release and domestic box office performance. In particular, it shows the changes in the linear relationship between number of theaters and box office when moving from one studio category to another. Each line represents point predictions for box office performance given a certain number of theaters on release and a certain studio category. The bands represent the 95% confidence intervals for these predictions.

It is clear that there isn't much difference in the relationship betrween number of theaters and box office performance when comparing films made by minor studios and those made by mid-sized ones. Major studio films, on the other hand, exhibit a stronger relationship between number of theaters and box office, especially for films released in more than 3000 theaters. Major studio films seem to better take advantage of large-scale releases than do those released by smaller studios. One possible reason may be that major studios have the marketing budgets to encourage large numbers of people to go fill the seats in the theaters their movies are shown in, whereas smaller studios lack such resources.

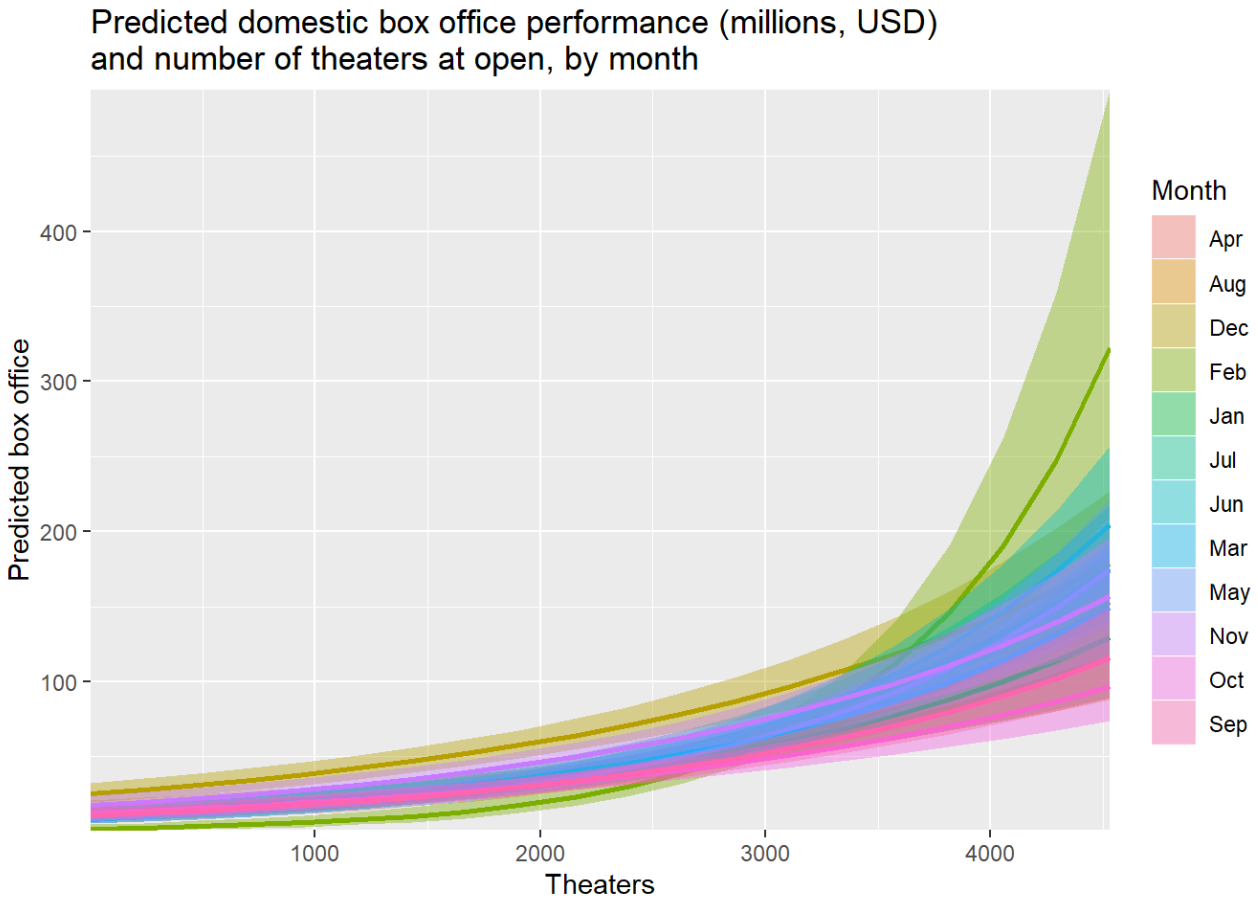


The next two plots visualizes the relationship between the two Rotten Tomatoes score variables and studio status. It is important to note that only the first plot is based on Model 2. The second plot illustrates what happens when we interact the RT audience variable, instead of the RT critics one, on studio status. In other words, the second plot displays predictions based on a different model. I am displaying this alternative specification to demonstrate the unique relationship the RT critics variable has with studio status. Specifically, for mid-sized and minor studio films, the relationship between RT critics scores and box office performance is negligible and even negative for mid-sized studio films. The overlapping confidence bands for these two categories suggest that there is no real difference between them when it comes to predicting the relationship between RT critics scores and box office performance. However, major studio films show a notably different pattern. For these films, there appears to be a very strong, positive relationship between critics score and box office performance. It may be that moviegoers pay more attention to Rotten Tomatoes critics scores for major studio movies when deciding whether to go see a movie than for mid-sized/minor studio films. Given the potentially more niche nature of mid-sized/minor studio films, viewers of these movies may already make up their mind as to whether to go see them and are not as interested in what the critics say. Major studio films, though, are often intended to be tentpoles that appeal to many types of people and so there is no automatic market segment interested in seeing them. If this hypothesis is true, then the box office success of major studio films can be considerable, but may be substantially tempered by a lukewarm reception by the critics.

Interestingly, we do not see the same pattern when substituting an interaction between the RT audience score variable and studio status. The slopes for all three categories and positive, and while the prediction line for the major studio category is significantly steeper than those for the other two categories, the difference is not as remarkable as it is for the RT critics score-studio status interaction. Indeed, the AIC score for the model decreases when we swap this new interaction in or include both interactions. Therefore, the RT critics score-studio status interaction does more predictive heavy-lifting and is worth including in Model 2.



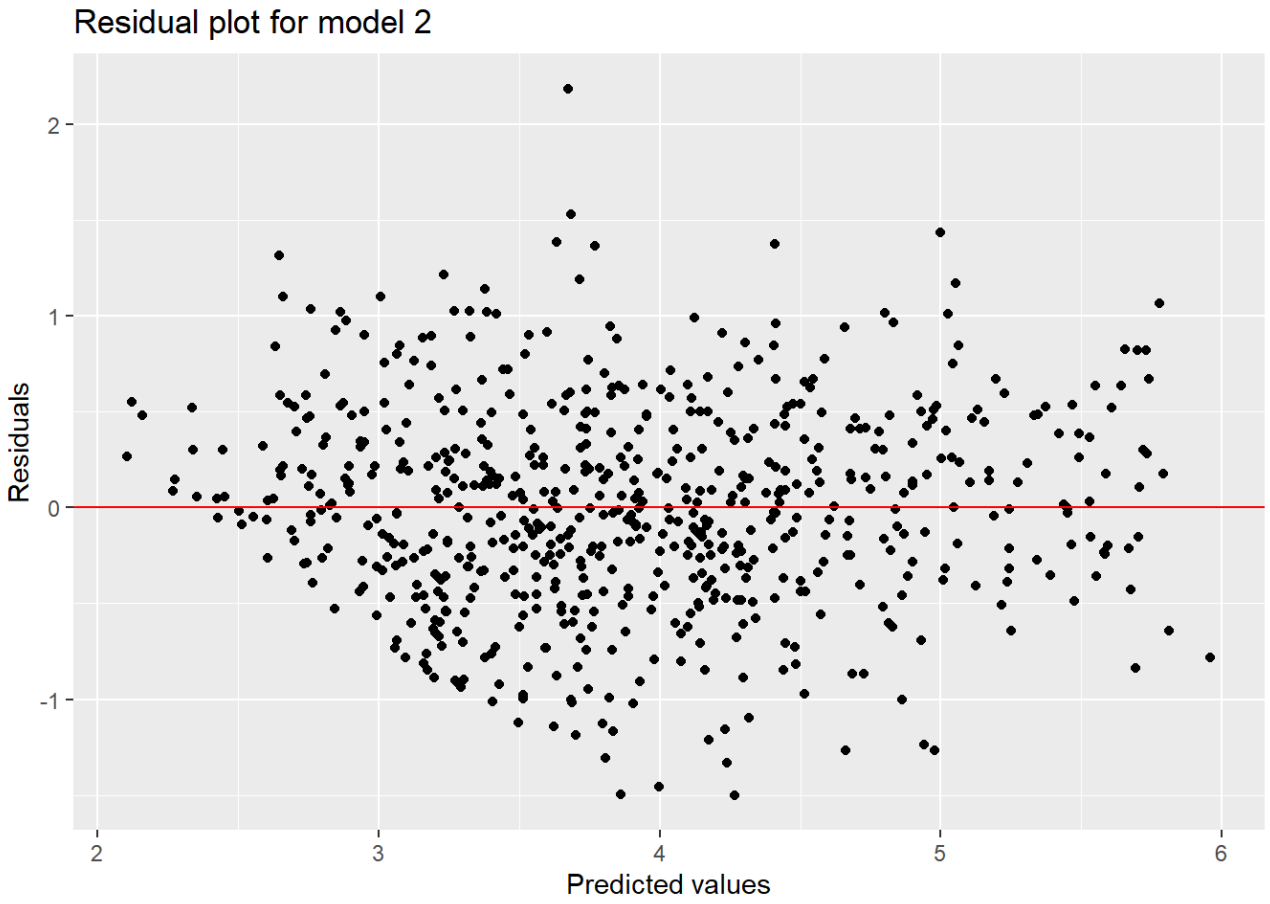
The final plot visualizes the interaction between number of theaters on release and month when predicting box office performance. Most of the monthly prediction lines are bunched fairly close to each other, with the major exception of December. In the final month of the year, films that are released very widely (i.e., more than 3700 theaters) seem to do much better than films with similar releases in other months. Given that this is a time of year when many people have ample holiday/vacation time, and most beaches are closed, it is not surprising that big movie releases do very well during December.

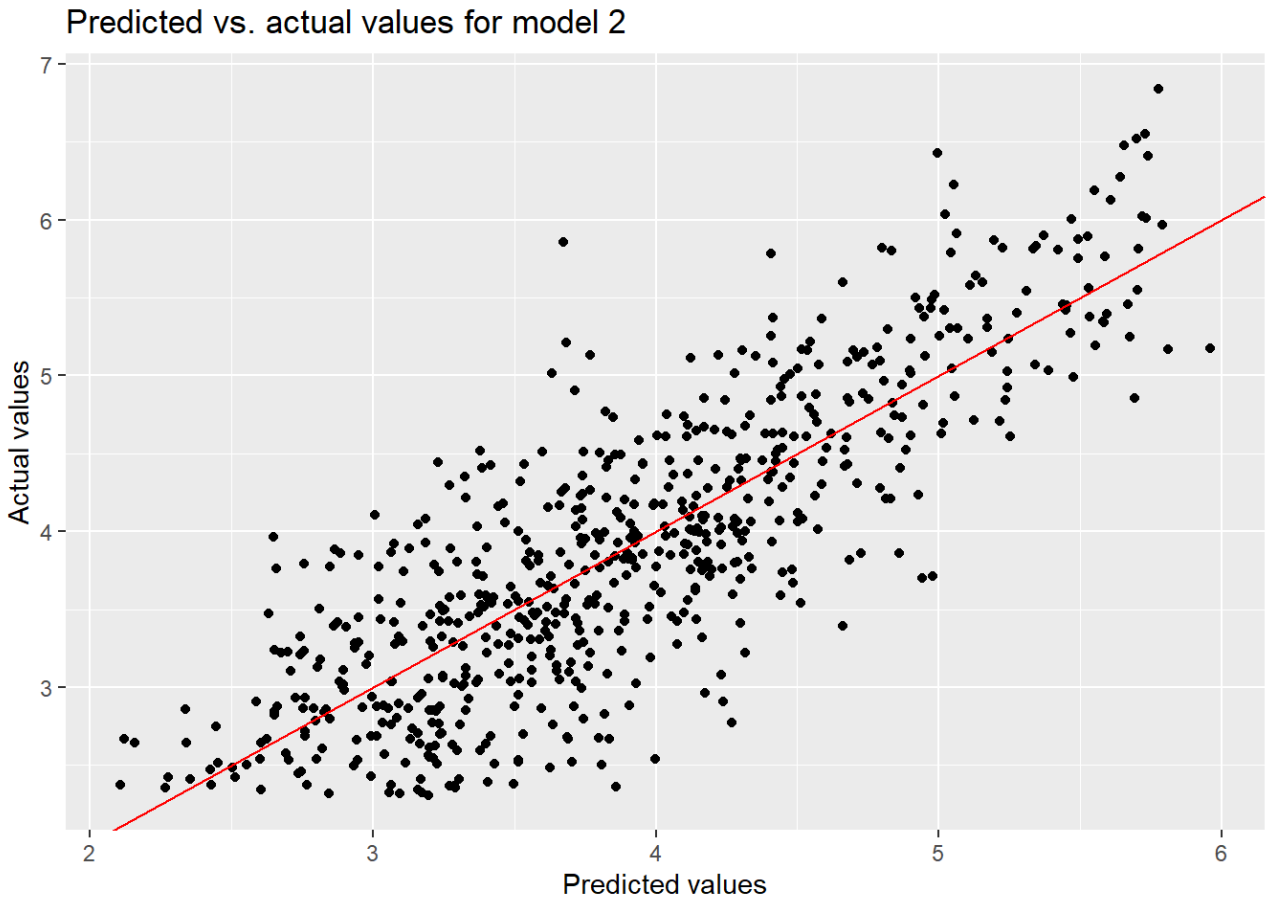


Robustness tests

This section presents the results of a number of checks on the robustness of the models I have so far created. The following two plots display the relationship between the predicted values of the outcomes and their residuals, that is, the difference between the actual outcome for a given observation (i.e, the actual box office performance of a given film) and its predicted outcome. Both plots are for Model 2 only, since this is the one I use for prediction purposes given its higher r.squared and lower AIC. The first plot displays this relationship directly, with residuals plotted on the y-axis and predicted values on the x-axis. Ideally, the residuals should be evenly distributed on either side of the zero line and should not display any noticeable pattern. Both criteria seem to be met by this plot, though there appears to be a noticeable line demarcating the bottom-left quadrant of the plot. This may be because I excluded films that made under ten million USD at the box office since I did not want very small films to bias my analysis.

The second plot displays the same relationship, but rotated such that the y-axis now represents the actual values for the outcome variable. Again, there should be balance on either side of the red line and no clear patterns should be present, both of which we observe.





Next, I perform two robustness tests designed for OLS regression models: the Bonferroni Outlier Test and the Correlation Test for Normality. The outlier test indicates that a single observation is a significant enough outlier to have an outsized influence on

the model: American Sniper. This film was released in just 4 theaters on release but ended up making 350.126372 million USD at the boxoffice. To be technical, this observation has a Bonferroni p-value of 0.0268067. I am not concerned about a single outlier invalidating the model and I would never want to discount any professional duet between clint_eastwood and bradley_cooper.

The correlation test for normality obtains the correlation between the model's residuals and the expected values of the residuals under an ideal normal distribution. Since OLS regression assumes that residuals are normally distributed, this is an appropriate test for determining if this criterion is met. the reported correlation is 0.9988271, indicating the model's residuals are very nearly normally distributed.

Finally, we can measure variance inflation factors (VIFs) to identiy any collinearities between predictors that may bias their coefficients. In general, VIFs should not be above four and certainly not above ten. The following table displays the VIFs for model 2. We can see that the VIFs for many of the variables are very high. However, these are due to the fact that we have interaction terms in the model, which purposefully introduce collinearity into the model. Instead, a better approach would be to examine the VIFs for model 1, which is identical to model 2 except without the interaction terms.

Robustness measures for model 2

Variables	Tolerance	VIF
rt.critics	0.08851250	11.297839
studio.status.threecatmid	0.03111510	32.138740
studio.status.threecatmajor	0.03096905	32.290303
rt.audience	0.34863052	2.868366
dom.theaters.open	0.02035690	49.123389
year	0.89605554	1.116002
monthFeb	0.02212174	45.204396
monthMar	0.05138044	19.462660
monthApr	0.05765072	17.345836
monthMay	0.05606168	17.837497
monthJun	0.04911003	20.362440
monthJul	0.04929227	20.287155
monthAug	0.04193195	23.848163
monthSep	0.05478514	18.253125
monthOct	0.05100885	19.604443
monthNov	0.05184461	19.288410
monthDec	0.05219829	19.157715
genreAction	0.68736079	1.454840
genreAnimation	0.55869357	1.789890
genreArtForeign	0.94566693	1.057455
genreClassic	0.92945991	1.075894

genreComedy	0.59707374	1.674835
genreDocumentary	0.71916222	1.390507
genreDrama	0.58135243	1.720127
genreHorror	0.69634103	1.436078
genreFamily	0.59332211	1.685425
genreMystery	0.81195801	1.231591
genreRomance	0.89913472	1.112180
genreSciFi	0.68414131	1.461686
rt.critics:studio.status.threecatmid	0.06248580	16.003637
rt.critics:studio.status.threecatmajor	0.05331386	18.756849
studio.status.threecatmid:dom.theaters.open	0.07116064	14.052712
studio.status.threecatmajor:dom.theaters.open	0.02744096	36.441872
dom.theaters.open:monthFeb	0.02089129	47.866839
dom.theaters.open:monthMar	0.04497321	22.235458
dom.theaters.open:monthApr	0.05558416	17.990736
dom.theaters.open:monthMay	0.04407832	22.686891
dom.theaters.open:monthJun	0.04068563	24.578702
dom.theaters.open:monthJul	0.04024835	24.845741
dom.theaters.open:monthAug	0.03973159	25.168889
dom.theaters.open:monthSep	0.05060987	19.758990
dom.theaters.open:monthOct	0.05009750	19.961077
dom.theaters.open:monthNov	0.04758043	21.017046
dom.theaters.open:monthDec	0.05869553	17.037074

We can see in this table that the VIFs are all lower than four and some are quite close to one, which would include zero collinearity. We can conclude that collinearity between predictors is not biasing our coefficient estimates.

Robustness measures for model 1

Variables	Tolerance	VIF
rt.critics	0.3844717	2.600972
rt.audience	0.3669523	2.725150
dom.theaters.open	0.4195639	2.383427
year	0.9106682	1.098095
monthFeb	0.5116638	1.954408
monthMar	0.4657048	2.147283

monthApr	0.5010704	1.995727
monthMay	0.4926335	2.029907
monthJun	0.4613670	2.167472
monthJul	0.4415166	2.264920
monthAug	0.4718891	2.119142
monthSep	0.4911686	2.035961
monthOct	0.4654529	2.148445
monthNov	0.4188059	2.387741
monthDec	0.4039888	2.475316
genreAction	0.6936126	1.441727
genreAnimation	0.5709827	1.751367
genreArtForeign	0.9662435	1.034936
genreClassic	0.9384228	1.065618
genreComedy	0.6090060	1.642020
genreDocumentary	0.8073630	1.238600
genreDrama	0.5932950	1.685502
genreHorror	0.7108031	1.406859
genreFamily	0.5969947	1.675057
genreMystery	0.8337300	1.199429
genreRomance	0.9224379	1.084084
genreSciFi	0.7284712	1.372738
studio.status.threecatmid	0.4116733	2.429111
studio.status.threecatmajor	0.3056628	3.271578

Stepwise model

My final robustness check is to run a stepwise forward regression procedure to see which predictors should be included in a model according to standard criteria. Stepwise forward regression adds predictors to a model based on how much added accuracy they provide, given the existing set of predictors. The procedure starts by adding a single predictor that contributes the most to the model's accuracy in predicting the outcome, the adds a second one from the remaining pool of candidate predictors, and so forth until all predictors have been added. I use Mallow's Cp to decide which predictors to include in this version of the model. Mallow's Cp measures the contribution a particular predictor provides to a model, while attempting to avoid overfitting (i.e., adding so many predictors to a model, based on a particular sample of data, such that the model poorly predicts data from outside the sample). Mallow's Cp is thus similar to AIC. In principle, one can use P (the number of predictors included in the model up to that point in the stepwise procedure) as a threshold for inclusion in the final model, such that Mallow's Cp should be higher than P. A more conservative approach would be to use 2 times P as the threshold. Either way, the same set of predictors is recommended by stepwise forward regression and Mallow's Cp.

The table below highlights these predictors in green. Notably, not a single genre predictor is deemed worthy to be included in the model. In addition, the RT audience score variable contributes much more to the model's accuracy than does its critics counterpart. This may be due to the peculiar interaction the RT critics score has with studio status, as discussed above. We also see that the r.squared of the model flattens out after including the RT critics variable, increasing by less than 0.02 after including the remaining displayed variables. Recall that the r.squared of Model 2 is 0.69, only a bit higher than the r.squared for the stepwise model. For prediction purposes, Model 2 is superior. But the stepwise version is more parsimonious and may test better on new data as a result.

Results for forward selection stepwise model

predictors	mallows_cp	rsquare	rmse
dom.theaters.open	466.9648227	0.3783669	0.7589407
rt.audience	140.0510296	0.5619128	0.6376112
month	51.0743946	0.6126811	0.6046830
studio.status.threecat	24.1719378	0.6288098	0.5928906
rt.critics	12.0973046	0.6366639	0.5870469
genreSciFi	4.7852441	0.6418604	0.5832940
genreDrama	2.7304877	0.6441231	0.5819083
genreDocumentary	-0.8662109	0.6472463	0.5798081
genreAction	-2.5185217	0.6492844	0.5785893
genreHorror	-2.4891886	0.6503841	0.5781405
year	-2.4841409	0.6514973	0.5776787
genreComedy	-1.8144188	0.6522397	0.5775231

xgboost model

The fourth and final model I construct is quite different in design than the first three. This model is built using the XGBoost (eXtreme Gradient Boosting) machine learning algorithm. XGBoost uses a series of decision trees to arrive at a model that best predicts the outcome variable (here, domestic box office performance) while minimizing overfitting. In theory, a model that is not overfitted should not only be effective at explaining the data used to create it (the training data) but also good at predicting the outcome given new data (the test data). A variety of hyperparameters are defined when constructing an XGBoost model, which I present in the table below:

XGBoost Hyperparameters

seed	gamma	eta	max_depth	subsample	lambda	nrounds
2020	0	0.1	3	0.8	1000	1000

The seed represents the value utilized by the random number generator whenever any arbitrary choices are made by the machine learning algorithm. Gamma (the tree complexity parameter) sets the threshold that the reduction in error the leaf node of a tree must meet in order to be kept in the tree model. Eta (the learning rate) represents the weight given to each new tree added by the algorithm and can be set lower to reduce the risk of overfitting. The lower eta is, the more cautious the algorithm is but also the longer the algorithm will take to arrive at the optimal model given the training data. Max depth is the maximum number of nodes permitted in each tree. The smaller the max depth, the less likely overfitting becomes. Subsample is the ratio

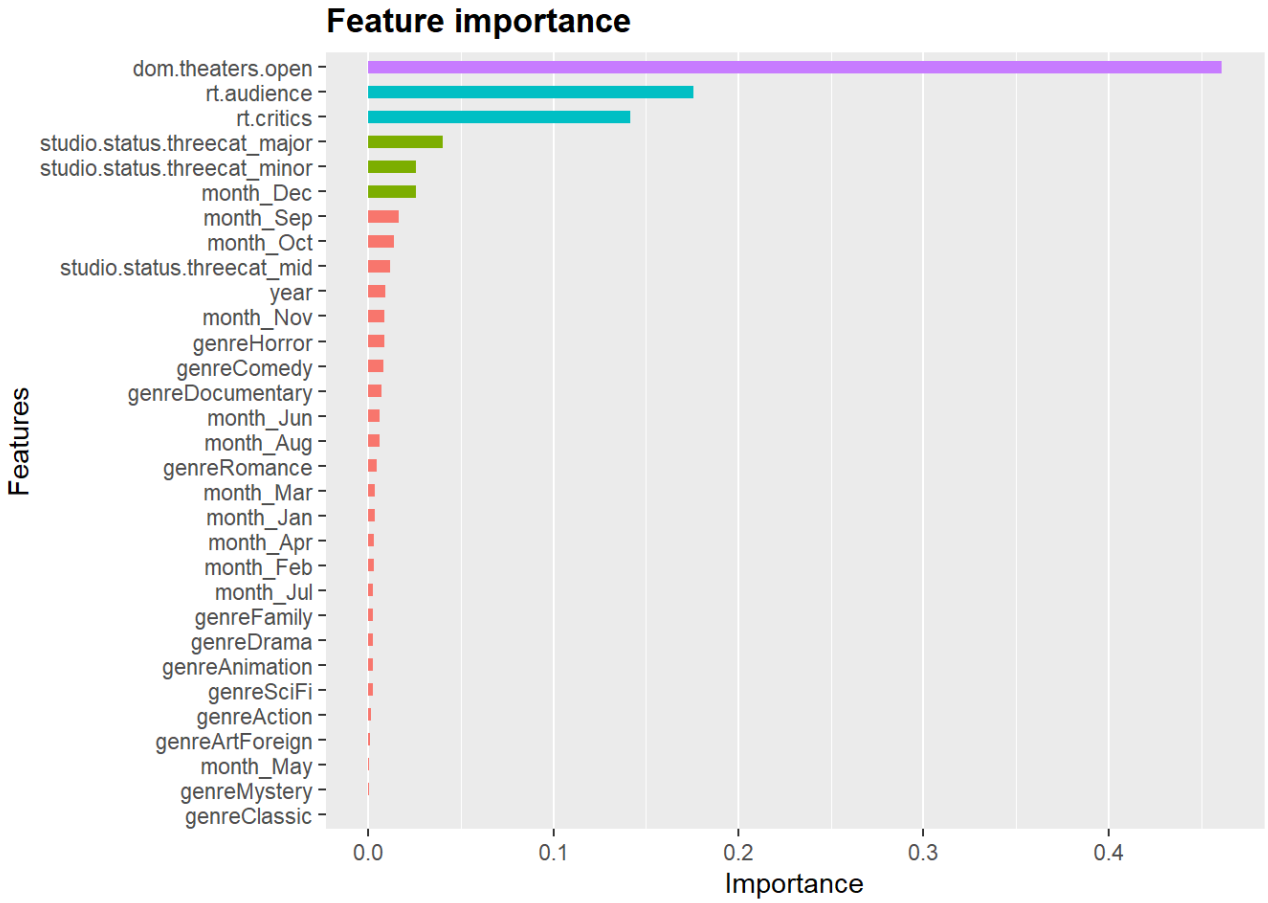
of the training data that is sampled for growing trees. Lambda (a regularization paramter) represents the penalty added to the calculation of an individual leaf node's contribution to the tree's predictive accuracy and is designed to reduce the impact of individual observations on this calculation. In other words, gamma helps ensure that leaf nodes that distinguish fewer observations are penalized more than those that distinguish more observations. Nrounds is the total number of trees that are grown by the algorithm. When eta is low, nrounds should be high to give the algorithm time to converge on a final model.

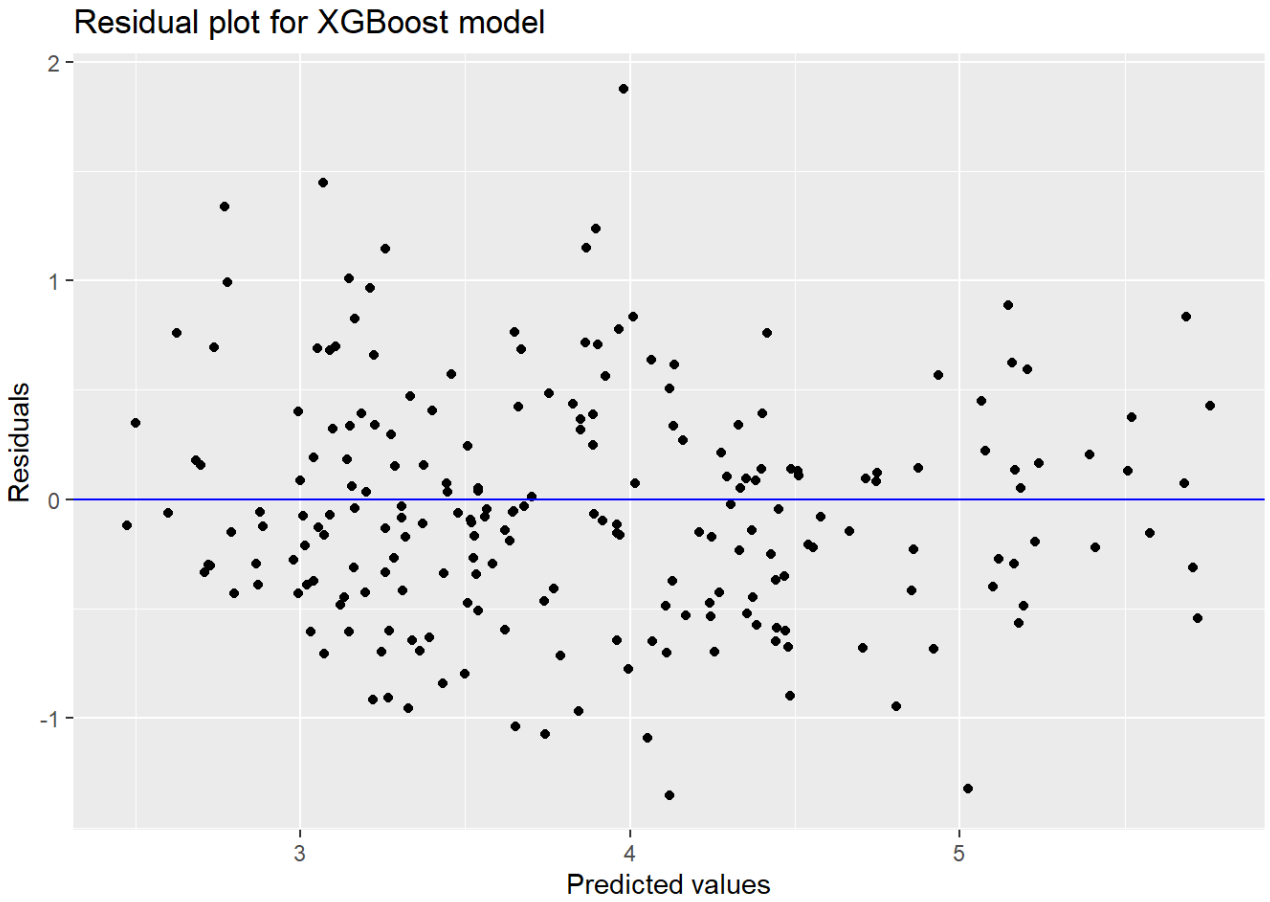
In other words, my approach to building this XGBoost model is to allow for less pruning for each tree (low gamma) but reduced sensitivity to individual observations (high lambda) combined with many trees (high nrounds) incrementally combined (low eta). I tried other variations (not shown here) and determined that this approach results in the best predictions and similar average errors for the training and test sets.

(NOTE: I use the xgboost package for R and all other hyperparameters for the xgboost function are left at their defaults)

The XGBoost model I construct includes all the variables from Model 1: the two Rotten Tomatoes score variables, month, genre, year, number of theaters on release and studio status category (minor, mid-sized and major). The root mean squared error (RMSE) for the training data is 0.398871, while the RMSE for the test data is 0.531061. The RMSEs are pretty close to each other, which is desirable.

The following plot shoes the importance of each "feature" or variable included in the XGBoost model. Importance is essentially a measure of how often a variable is used by a leaf node, that is, to make decisions about how to minimize the gap between an observation's actual score for the outcome and its predicted score. We can see that the number of theaters on release is far more important than any other variable. Both Rotten Tomatoes score variables are next in importance, with the audience variable the more important of the two. Of the three studio status categories, the minor and major categories are substantially more important than the mid-sized one. Finally, December, October and September are the most important months when it comes to predicting box office performance. No genre seems to play a major role nor does the year variable. These findings more or less match those we obtained from the OLS models presented above.





The above plot demonstrates that the residuals of the XGBoost model's predictions are fairly evenly spread out across the zero line, as we would hope. There does seem to be some bias in favor of above the line for values on the lower and upper ends of the x-axis, but this pattern seems fairly weak. No other patterns are apparent.

Predictive performance

Now that we have several models to use for predicting box office performance, let's compare them directly to see which is best at prediction. I use a folding method that divides the data into five groups (or folds) and trains each model on four of the five folds and then tests each model on the fifth fold This process is repeated five times such that each fold is used once for testing purposes. I do not include Model 1, since Model 2 is identical except for the inclusion of three interaction effects. The table below displays the results for each model, indicating average errors (specifically, root mean squared error) and r.squared statistics, as well as minimums and maximums for each.

Root mean squared error and r.squared statistics for three models

Model	Average error	Minimum error	Maximum error	Average r.squared	Minimum r.squared	Maximum r.squared
Model 2	0.5856	0.5520	0.6311	0.6201	0.5498	0.6816
Stepwise model	0.6003	0.5400	0.6607	0.6080	0.5432	0.6653
XGB model	0.5164	0.4741	0.5419	0.7092	0.6914	0.7323

We can see that the XGBoost model performs the best across all seven measures. Model 2 performs second best across most, though there seems to be a somewhat wider range for RMSE compared to the stepwise model. This may be an artifact of the way in which observations were allocated to folds, however.

Lessons for the film industry

What does all this mean for producers looking to decide which film projects to invest in? For one, timing matters, but only to a limited degree. Outside of December and (to a lesser extent) November, it does not appear that when a movie is released matters all that much. Whereas in the past, the summer months and December were when films made the most the money, this is no longer the case, as recent box office smashes such as Black Panther (released in February), Joker (released in October) and American Sniper (released wide in January) attest. Instead, studios could benefit from spreading their releases (hopeful blockbusters or otherwise) across the calendar year. It is unlikely that such a strategy would result in much lost revenue and the films may encounter less competition from others.

Similarly, genre does not seem to matter as much, with the major exception of science fiction (which includes speculative fiction of many kinds). Fans seem to flock to science fiction movies recently in much greater numbers than they used to (see the relative lack of successful scifi movies in the 1990s). Much of this is due to the rise of Marvel Studios and its immensely popular cinematic universe. Even so, films like Jurassic World, Beauty and the Beast, and Wonder Woman illustrate that Marvel does not have a monopoly on smash scifi hits (though it seems Disney almost does). And while the great majority of major studio hits are based on well-established intellectual property, there are notable exceptions, such as Interstellar and The Martian, both of which made more than 180 million USD in North America. Both of these films were helmed by highly respected directors (Christopher Nolan and Ridley Scott) and had prominent leading men (Matthew McConaughey and Matt Damon).

How many theaters a film is released in, not surprisingly, has a major impact on how much money is earned. But this relationship is stronger for major studios, especially for releases of more than 2700 theaters. Producers working for smaller studios may be justified in releasing their films less widely (2500 theaters or less) since the added value of wider releases for these types of studios is ambiguous.

Critics and audience reception matters, but not as much as studio executives and producers may think. A higher audience score is associated with better box office performance for all studio types, but much more so for major studios. These latter studios should therefore focus on making high-quality films, even if they are big-budget scifi actioners. Audiences likely having increasingly high expectations for acting, dialogue, editing and special-effects quality. As movie tickets continue to increase in price, fans will expect a better overall product. The weak reception of many DC Comics films and the much stronger reception of Marvel comics films speaks to this. Major studios also have to pay attention to the critics. Whereas the relationship between critics score and box office performance is essentially nonexistent for mid-sized and minor studio films, it is very much existent for major studio films. The preferences of critics should be taken into account when producing a major studio film to maximize ticket sales.