Objective:

To analyze the popular rating site like rotten tomatoes, Metacritic, IMDB and Fandango to assist in making the decision for picking the best rating site to rely on for buying most popular movies distribution rights.

Preface:

This analysis is conducted on the sample of rating provided by popular movie rating sites like Fandango, Metacritic, Rotten Tomatoes, and IMDB for 146 movies in the year 2014 and 2015. The data include ratings by user and critics.

Initial analysis of the data shows that the rating is done on different scales by the rating sites.

Scales used:

- Rotten Tomatoes: 1-100
- Metacritic: 1-100
- Metacritic Users: 1-10
- IMDB: 1-10
- Fandango: 0-5

Hence the final analysis is done on the data that is normalized to the scale of 0 to 5 to provide an unbiased result.

Analysis:

The analysis is primarily based on the ratings by the critic but the user rating have been also considered for the purpose of comparison to check that the critics rating are not purely biased or influenced.

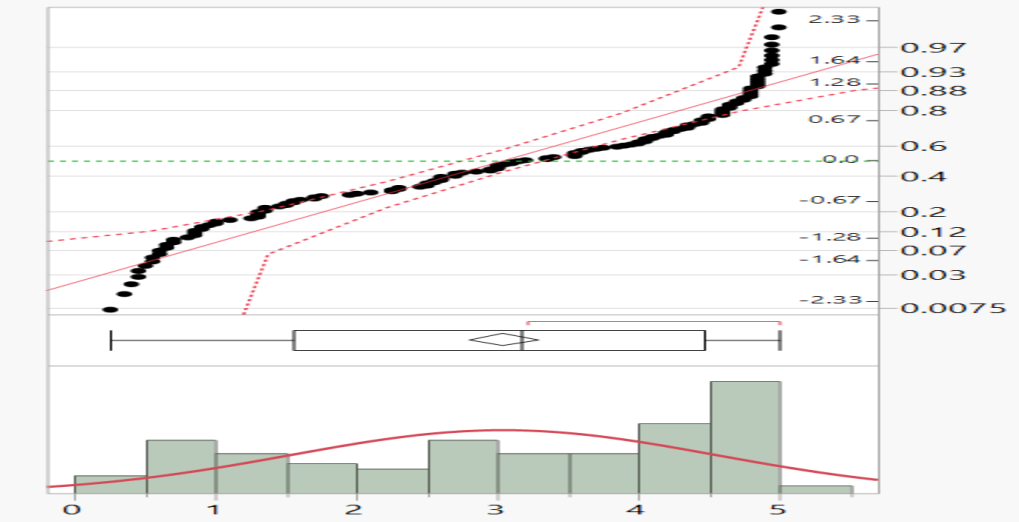
Descriptive Statistics and Normal Distribution

	Fandango _Stars	Fandango _Rating value	Rotten Tomatoes	Rotten Tomatoes _Users	Metacritic	Metacritic _User	IMDB
Mean	4.09	3.85	3.04	3.19	2.94	3.26	3.37
Standard Deviation	0.54	0.50	1.51	1.00	0.98	0.76	0.48
Minimum	3.00	2.70	0.25	1.00	0.65	1.20	2.00
First Quartile-25%	3.50	3.50	1.56	2.50	2.18	2.85	3.15
Second Quartile - 50%	4.00	3.90	3.18	3.33	2.95	3.43	3.45
Third Quartile - 75%	4.50	4.20	4.45	4.05	3.75	3.75	3.70
Maximum	5.00	4.80	5.00	4.70	4.70	4.80	4.30

Fandango Stars:

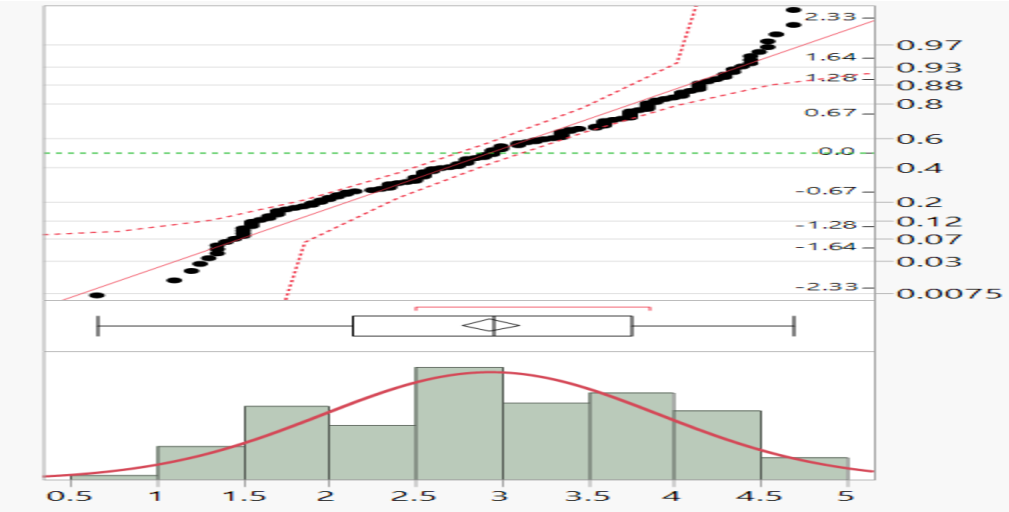
The above statistics shows that the mean value is 4.09 which states that this the average rating given by Fandango Stars. The minimum rating given is 3.0 and standard deviation tells us about the spread of data along the mean which is 0.54. The above figure shows the distribution of the data is near to normal distribution but left skewed. Thus, 68% of movie ratings fall within the range of **3.46 to 4.54.** From the above scenario we can conclude that the fandango ratings are above average, so it can be biased, and we cannot rely on it.

Rotten Tomatoes



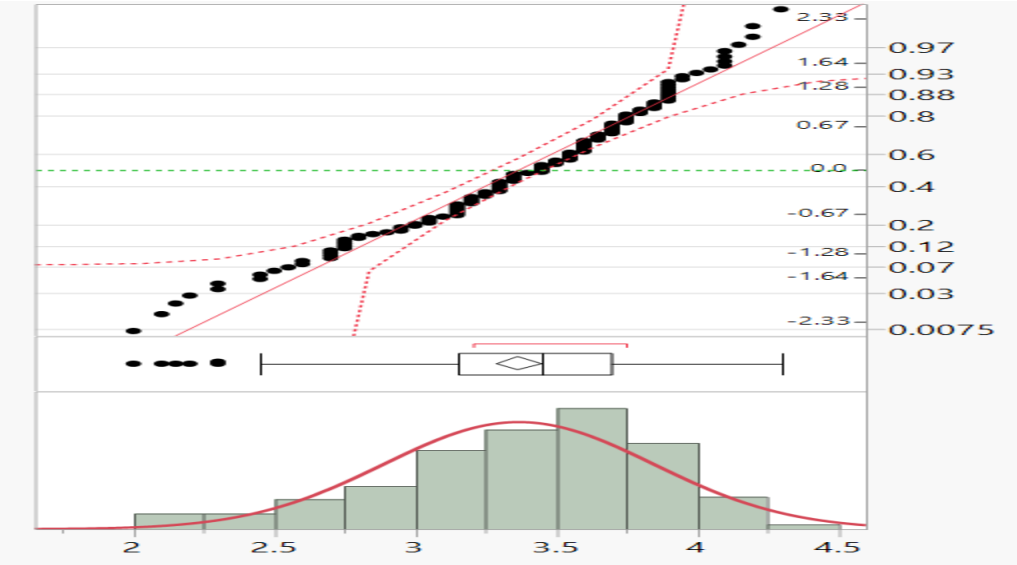
The higher standard deviation (1.51) tells us about the wide spread of data and from the distribution curve we can say that the distribution of the ratings doesn't have a bell-shaped curve or normally distributed. As the center is flattened, we could say most of the movies are rated either below average or above average and higher standard deviation tells us that data is widely spread. Thus, we can conclude that as data is not normally distributed we cannot take it into consideration.

Metacritic



From descriptive statistic we know that the mean (2.94) and the median (2.95) are almost same and the graph shows that the data is normally distributed also the skewness is near to zero. The rating is also between 0.65 and 4.70 so we can conclude from it that the ratings are not biased and genuine ratings has been given.

IMDB



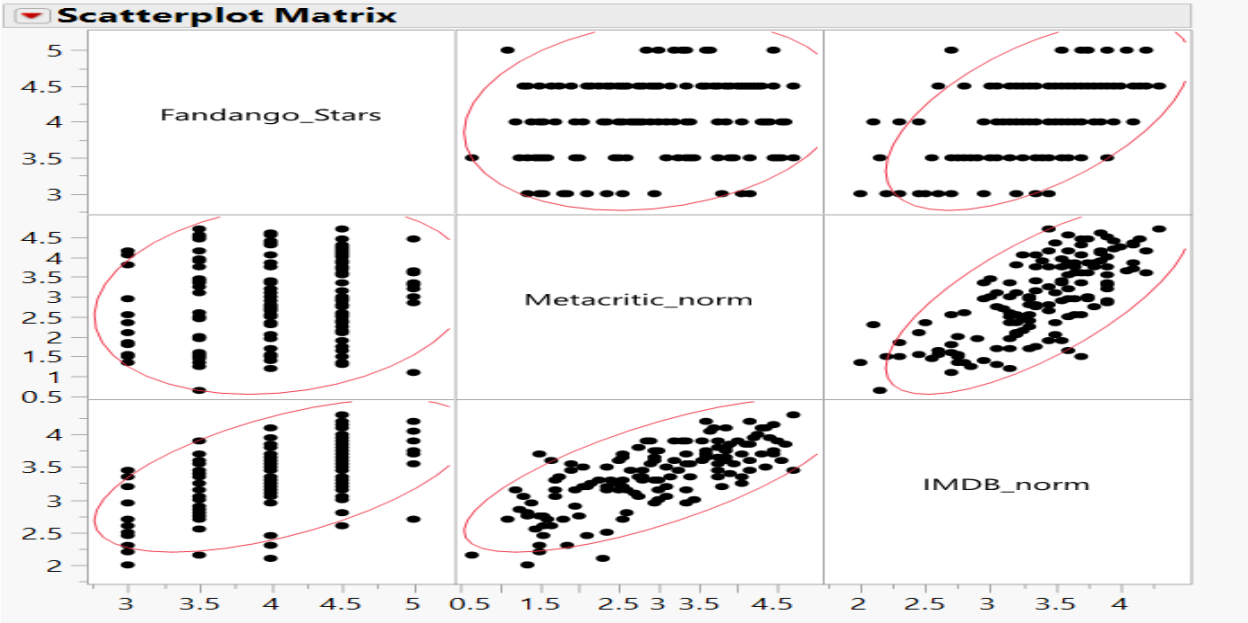
By comparing the quartile values, we can say most of the ratings fall almost near to average and above average. The figure shows the frequency distribution for IMDB rating. Looking at the graph, we can say that the data is near to normally distributed data. The Mean (3.37) is less than the median (3.45) which tells us that the it is negative skewed (-0.62). As per the quartile value stated in the descriptive statistics, it could be concluded that most of the movies are rated as above average. Thus, there is a chance that rating can be biased.

Correlation

From the above analysis, Fandango stars is least preference compared to others. We are not taking Rotten Tomatoes as it is not normally distributed hence, now we compare IMDB and Metacritic with Fandango to find which is least corelated.

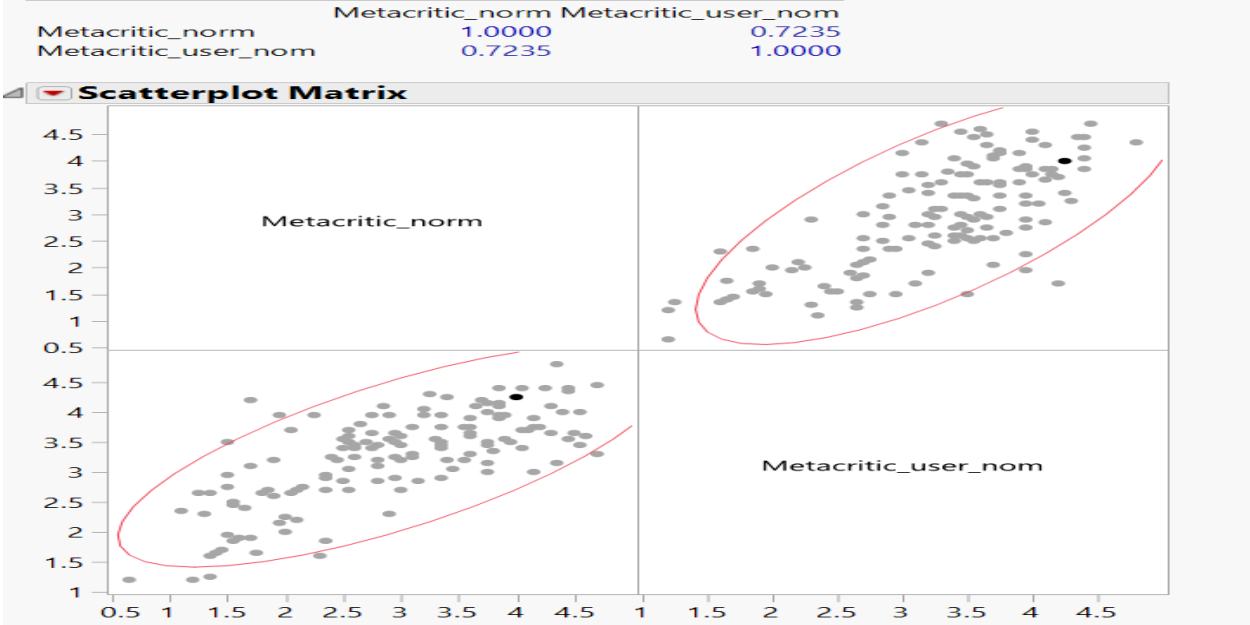
Below is the correlation matrix:

	Fandango_Stars	Metacritic_norm	IMDB_norm
Fandango_Stars	1.0000	0.1811	0.5873
Metacritic_norm	0.1811	1.0000	0.7273
IMDB_norm	0.5873	0.7273	1.0000



As per the above table, Metacritic has the least correlation value compared with Fandango Stars.

From the above analysis, we can say Metacritic is best among all. Hence, to further substantiate the conclusion, let's compare the critics rating with user rating to see the correlation between them.



As per the above correlation table, the value is **0.72** which shows they are correlated. This proves that the critics rating balances the user rating also.

Conclusion

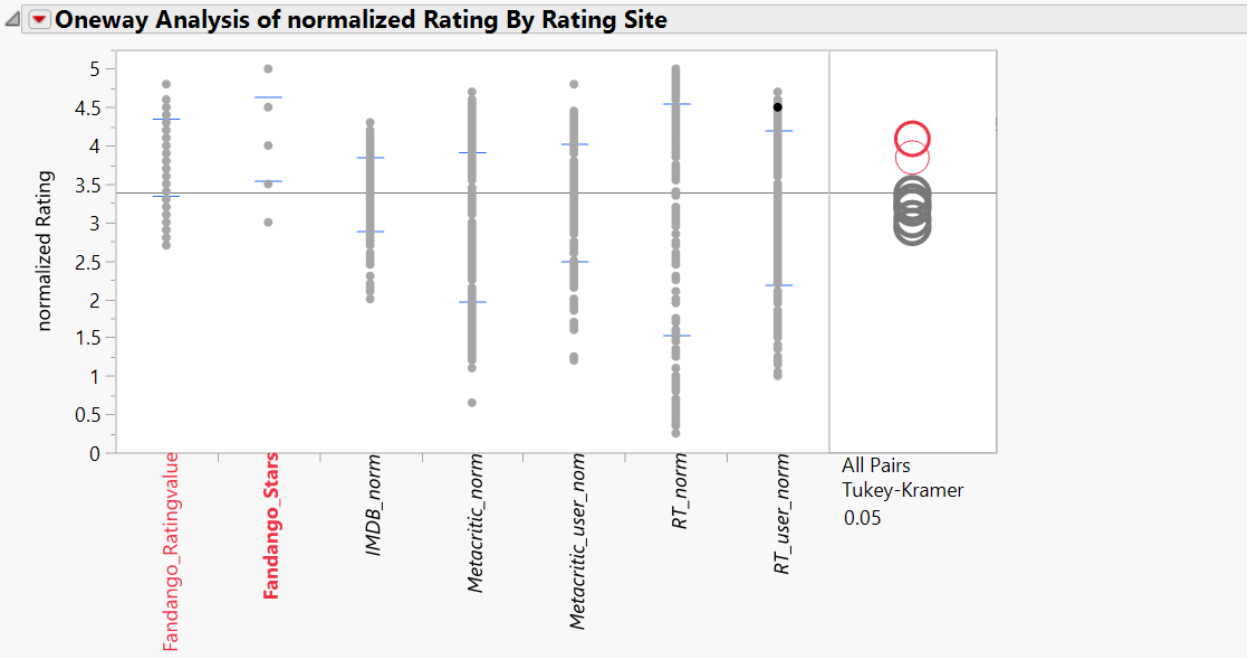
From the above analysis, my recommendation would be ***Metacritic***, the reason being

1. The critic's ratings are spread across below average to above average with more rating falling under the average category.
2. The data is normally distributed compared to others.
3. Metacritic is least correlated with Fandango which is least preferred and looks biased in rating.
4. Metacritic critic's rating and user ratings are also correlated, which shows that critic rating matched with user rating as well.

Methods and Tools Used:

- Descriptive Statistics
- Normal Distribution
- Correlation
- JMP

Appendix



HSD Threshold Matrix

Abs(Dif)-HSD

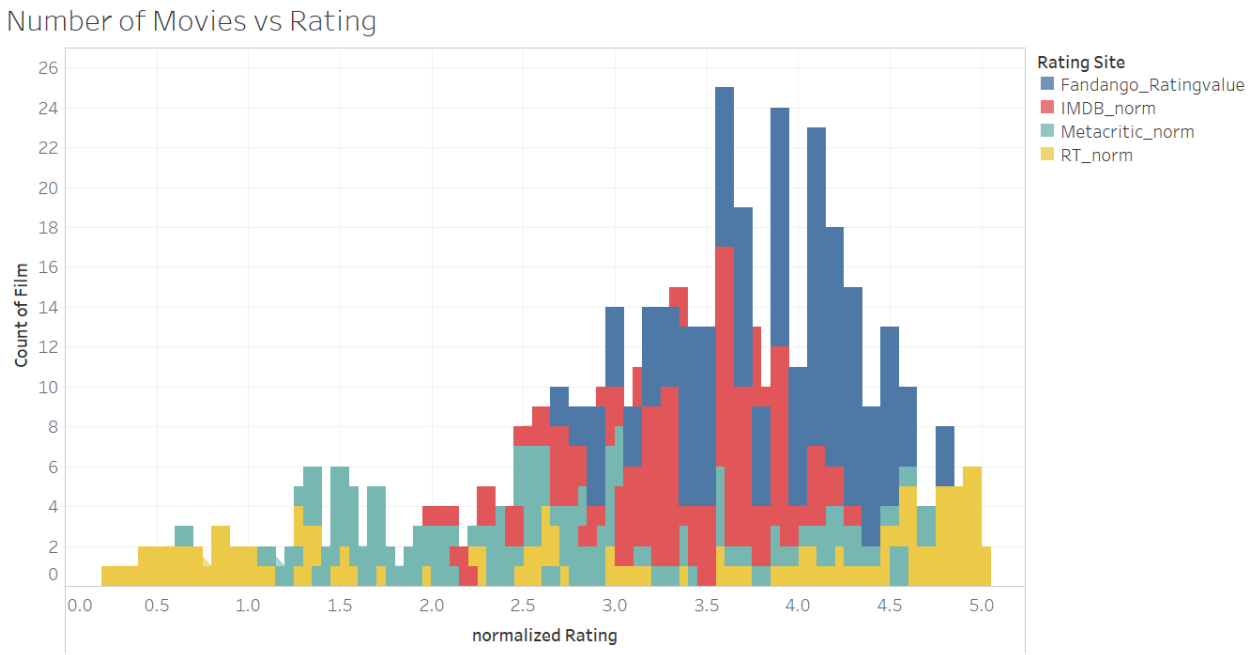
	Fandango_Stars	Fandango_Ratingvalue	IMDB_norm	Metacritic_user_nom	RT_user_norm	RT_norm	Metacritic_norm
Fandango_Stars	-0.30859	-0.06475	0.41196	0.52087	0.58662	0.73799	0.84004
Fandango_Ratingvalue	-0.06475	-0.30859	0.16813	0.27703	0.34278	0.49415	0.59621
IMDB_norm	0.41196	0.16813	-0.30859	-0.19968	-0.13393	0.01744	0.11950
Metacritic_user_nom	0.52087	0.27703	-0.19968	-0.30859	-0.24283	-0.09146	0.01059
RT_user_norm	0.58662	0.34278	-0.13393	-0.24283	-0.30859	-0.15722	-0.05516
RT_norm	0.73799	0.49415	0.01744	-0.09146	-0.15722	-0.30859	-0.20653
Metacritic_norm	0.84004	0.59621	0.11950	0.01059	-0.05516	-0.20653	-0.30859

Positive values show pairs of means that are significantly different.

Connecting Letters Report

Level					Mean
Fandango_Stars	A				4.0890411
Fandango_Ratingvalue	A				3.8452055
IMDB_norm		B			3.3684932
Metacritic_user_nom		B	C		3.2595890
RT_user_norm		B	C	D	3.1938356
RT_norm			C	D	3.0424658
Metacritic_norm				D	2.9404110

Levels not connected by same letter are significantly different.



The plots of count of Film for normalized Rating and normalized Rating. Color shows details about Rating Site. The view is filtered on Rating Site, which keeps Fandango_Ratingvalue, IMDB_norm, Metacritic_norm and RT_norm.