Kickstarter Analysis

Zachary Owen

The data collected here looks at Kickstarter campaigns from May of 2009 through March of 2017. This data allows for the analysis of a variety of different factors. One of the first conclusions one can draw from this data set, is that Kickstarter was most popular as a fundraising platform in the year 2015. The number of campaigns run on Kickstarter climbed steadily from 2009 to 2013 before taking a significant leap up in 2014. It crested in 2015 before dropping back down in 2016. If 2017 continues through the rest of the year at the same rate of new campaigns as it had through March, then the popularity of the platform will have fallen once again.

The data can also be used to look at the popularity of different categories of campaigns. While the most common types of campaigns across Kickstarter all fall into the arts categories, Plays are the most common sub-category with almost four times the number of the nearest sub-category. From this we can conclude that theatre groups view Kickstarter as a valuable fundraising tool, since over 65% of the plays that use the platform have been successfully funded. While that is a significant success rate, we can also conclude from this data that if one specifically wants a successful campaign, one should pitch a project in the Rock Music sub-category. Not only is Rock the second most popular sub-category, it has also been successful 100% of the time.

While this data is useful for seeing these trends in popularity based on category, there are many factors missing from this set. For one thing this data does not provide us with any information on the group pitching the project. Is the project being pitched by a new entity that has no promotional network, nor track record to stand behind? Is this project being pitched by an established creator with a significant cultural presence, and a list of successful projects to their name? These factors significantly skew the likelihood of success or failure, but we have no indication of that in the data set. Some creators are likely to have already had successful campaigns on Kickstarter already, so even gaining a knowledge of the number of previous projects the creator has pitched, would make the data more well-rounded.

While the data does list the geographical base for each campaign, it gives all of the data in the dollar amounts of the respective countries. As currency conversion rates change on a daily basis, there is no good way to truly compare the currency amounts from different countries in terms of how much each campaign needed to be successful. If the data was converted in to one standard unit of currency as it occurred on the day the pledge was made, one would be able to more accurately analyze the average contribution per backer, and the ratio of amount requested to success.

One additional piece of analysis that would be valuable, is a pivot table based on the country that the project originated from. Not only would one be able to see what countries have the highest level of engagement from the creator side, but one could also analyze the success rate of projects based on its geographical location, and whether certain categories are more popular in specific countries. Another valuable analytical tool would be a scatter plot of number of backers vs. percent funded. One could use that to see how much the percentage deviated based on that number of backers and also how engagement of a variety of backers affects the likelihood of success.

In performing a statistical analysis of this data, one will find that the median is a more accurate representation of the number of backers on successful and failed campaigns. The outliers within this data, are the campaigns that have a significantly higher number of backers. Those data points skew the mean, thus making it less meaningful for the data set as a whole. Overall, there is a higher degree of variability in the number of backers of successful campaigns than there are in failed campaigns. This makes sense as there is no ceiling to the number of backers a successful campaign can have, while there is a floor to how few backers a failed campaign can have. The mode of the number of backers for a failed campaign is zero, and sense it is impossible to have a negative number of backers, it logically follows that there would be less variance in the number of backers of the failed campaigns.