Chris Krokus

29MAR20

Kickstarter Data Report

General:

The dataset was made up of over 4000 Kickstarter campaigns, launched between 2009 and 2017. Our goal is to use the data from these eight years to uncover any tricks for finding success when we launch our very own Kickstarter campaign in the very near future. We do know that Kickstarter is a valid way to launch a business because over $2 Billion has been raised for entrepreneurs but the odds are stacked against us as only a third of them have made it through the funding process with a positive outcome – we’re hopeful that with this dataset, we can avoid joining the less fortunate majority…

This stacked column bar chart does not do this large dataset justice, especially as it is cut and paste from Excel. The largest columns on this chart were theater/plays, film, music/rock, and music indie rock. After filtering these campaigns out (all of which achieved great success with Kickstarter), the remaining columns were enlarged and trends emerged. Strictly successful campaigns, based on the entire column being gold was evident in not only music/rock, but also technology/hardware, games/tabletop games, food/small batch (some still live), music of the classical/electronic/metal/pop varieties, and publishing of the nonfiction and radio type. At the other end of the spectrum, we had several categories of campaigns that lacked gold/success entirely; food/food trucks, food/restaurants, games/mobile games, games/video games, music/faith, music/jazz, photography/nature, photography/people, photography/places, publishing/children’s books, publishing/fiction, publishing/translations, technology/gadgets, and technology/web.

From these two lists delineating between successful and not successful, we can make a few determinations. If you are a musician and considering launching your career via Kickstarter, there is a tremendous amount of data that will support your decision on whether to move forward with a campaign or not, purely on the genre of music you are in. If you are in rock, indie rock, classical, electronic, metal or pop then you should absolutely leverage Kickstarter, but on the other hand if you are a faith or jazz musician, I do not believe a Kickstarter campaign would benefit you. A similar thing can be said in the food industry, food trucks would not fair well but small batch on the other hand would. If I were a musician, I would eliminate the non-musical categories and scrutinize this dataset with a larger number of filters pertaining to whether or not my band would fit more successful trends or failing trends.

Similar to the previous stacked column bar chart, ‘plays’ dominates the total number of campaigns as a sub-category. When temporarily filtered out, you can get a better chance to see which sub-categories of campaigns are golden/successful and which are orange/blue/not successful. Largely successful campaign generalizations include Rock, Documentaries, Hardware, Indie Rock, Photobooks, Tabletop Games, Television, and Shorts. These columns have enough data points and appear to have more than 50% gold. On the other end of the spectrum, we have campaign generalizations that tend to fail or cancel like Animation, Drama, Faith, Fiction, Food Trucks, Jazz, Translations, Video Games, Wearables, Web, and World Music. These sub-categories are lacking the gold that we are looking for and it would be foolish to launch a Kickstarter campaign if the company falls under one of these genres.

This line graph displays the cumulative data over the course of 8 years and the purpose of it is to determine when the best time to launch a campaign is. My lines do not appear to accurate in relation to my pivot table data so I made a side table simply containing the difference between what should be the gold (success) line and should be the orange (fail) line.

|  |  |  |  |
| --- | --- | --- | --- |
|  | success | fail | diff |
| Jan | 182 | 149 | 33 |
| Feb | 202 | 106 | 96 |
| Mar | 180 | 108 | 72 |
| Apr | 192 | 102 | 90 |
| May | 234 | 126 | 108 |
| Jun | 211 | 147 | 64 |
| Jul | 194 | 150 | 44 |
| Aug | 166 | 134 | 32 |
| Sep | 147 | 127 | 20 |
| Oct | 183 | 149 | 34 |
| Nov | 183 | 114 | 69 |
| Dec | 111 | 118 | -7 |

The larger the number in the Difference column, the higher the better your chances of success will be if you launch your Kickstarter campaign during that month. May (108), February (96), and April (90) have a noticeable advantage over December (-7), September (20), and August (32). Obviously it depends on the product or service your company is launching, but my theory on the launch date success rate is that as people are preparing for summer with more happiness and daylight, they are more likely to contribute to these entrepreneurial campaigns.

Bonus Statistical Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Backers Successful | |  | Backers Fail | |
| Mean | 194.4 |  | Mean | 17.7 |
| Median | 62.0 |  | Median | 4.0 |
| Min | 1.0 |  | Min | 0.0 |
| Max | 26457.0 |  | Max | 1293.0 |
| Variance | 712841.0 |  | Variance | 3773.2 |
| Std. Dev | 844.3 |  | Std. Dev | 61.4 |

Variance: how far values in the data set are from the mean

Standard Deviation: how spread out the data is from the mean

For both Successful and Failing backers, the Mean is greater than the Median. This means that this data set is skewed to the right. The Median is not as affected by the outliers and therefore summarizes the data more meaningfully. The lower the Standard Deviation, the closer the data is clustered around the Mean. The Failing backers are therefore much closer to their mean of 17.7 than the Successful backers are to their mean of 194.4.

The Variance is the average of the squared errors. Taking the square root of that we have the Standard Deviation. If this were a normal distribution, then about 68% of the number of backers backing Successful campaigns would be 194.4 +/- 844.3 = between -649.9 and 1038.7. For Failing campaigns, 68% of the number of backers backing them would be between 17.7 +/- 61.4 = between -43.7 and 79.1 It is not possible to have a negative number of backers though…

In terms of variability, range is the simplest measure for it. The range for successful backing is 26456 whereas the range for failing backing is 1293. The range is obviously sensitive to outliers in general, so I calculated the Interquartile Range (IQR) for both. Successful backer’s IQR was 111.75 and Failing backer’s IQR was 11. Successful was more variable which makes sense because there is no upper limit for a campaign that succeeds (therefore yielding more large outliers) but there will be for a campaign that fails.

BONUS

There is a pretty steep drop off in success percentage and increase in fail percentage once the Goal Amount exceeds $45,000. The number of total projects between $30,000 and $49,999 is low so having such few data points could be a contributing factor to the sudden movement line movement at this p