

Crowdfunding Report

Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

We can see that, in general, more backers mean more successful campaigns. On average, there were 851 backers for successful ones, and 585 for failed. With the median being 201 for successful campaigns, and for 114 failed, the company or organization running the campaign should know it has more than 200 backers before starting a campaign.

The time of year and amount of the pledge goal also play a factor in success rates. Campaigns should set their funding goals between \$1k-\$5k to have the mostly realistically high success rate of 83%. The best time to start a campaign is in Jun and July with a success rate of around 62% to 64%, and the worst is in August at 49%.

The most successful category was Journalism with a 100% success rate, although, with only 4 campaigns, when the median number of campaigns was 67 and the mean 111, the success seems like a bit of an outlier. Outside of Journalism, Technology was the most successful at 67%, and Games was the least with 44%. The most successful subcategories were audio, and world music as 100% of their campaigns were successful. Again, they seem like outliers, with only 4 and 3 campaigns respectively. The median number of campaigns was 21, and the mean 42. Given that, Web was the most successful subcategory at 71%, and Mobile Games, with 31%, was the least. A final note is that while Theater may have the most campaigns overall, it is not the most successful – only 54% of campaigns succeeded.

What are some limitations of this dataset?

We don't know the target audience or the average income of the backers – that could help a campaign know how high to set the pledge goal. We also don't know how, and how well, the campaigns were marketed and publicized. Especially in the instance of outliers. For instance, how was Journalism successful every time, and why were there so few – only 4 - campaigns? Another example of outliers was in the crowdfunding

Rosensweig, Ben
Data Analytics and Visualization Boot Camp
April 2024

goals – why and how did the ranges of \$15k-\$25k, and \$30k-\$35k reach 100% success? Not to mention, we also don't know why there is a dip in successful campaigns between the \$5k-\$15k range. We also don't know why certain months are more successful than others for categories. For instance, we can easily see that Theater campaigns were most successful in June and September, and the least in May and October – why?

What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

We could create a table comparing the duration of a campaign to the outcome; that could provide useful insight for how long a campaign should prepare to be active before seeing success.

We could further compare the duration to pledge goal – one could assume a higher pledge goal with fewer days would result in more failure, but knowing this could help set more realistic pledges and durations of time.

A table comparing the average donation, backers, pledge goal, and outcome could also give an organization an idea on the most achievable pledge goal.

Regarding the comparison of the number of backers to successful or failed campaigns, use your data to determine whether the mean or the median better summarizes the data. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

In statistical analysis, the median better summarizes the data. For both successful and failed, the mean skews heavily to the right. The majority of successful campaigns had 16-556 backers, the median was 201 backers, and the mean 851, falling well outside. There is also more variability in successful campaigns. This makes sense, as not only were there more successful campaigns (565 versus 364), there was a greater range of the number of backers per campaign (the max being 7295 backers versus 6080).