Crowdfunding Report

By David Burke

Conclusions

We can draw several conclusions from this dataset, based on data shown in the worksheets *Categories*, *Subcategories*, *Date Created*, and *Success by Goal*.

Starting with *Categories* and *Subcategories*, we can surmise that crowdfunding campaigns are overall more likely to succeed (56%) than fail (36%). This is good news for both creators and backers. While there is still a significant risk of failure, that risk is outweighed by hope of success 2:1.

When looking at the data categorically, artistic ventures greatly outnumber other groupings, with Theater, Music, and Film & Video representing the vast majority of campaigns. When broken down by subcategory, Plays considerably outweigh the other groups (though it should be noted that Plays makes up the entirety of the Theater category, whereas other categories have more subdivisions). By visual inspection, these categories and subcategories are largely similar in their rates of success, though Mobile Games seems a particularly risky venture with a much higher rate of failure than average.

On the *Date Created* worksheet, we see a graph showing outcome by month which has three notable shifts: a divergence around June, a strong convergence in August, and another convergence in December. These can possibly be explained by potential backers’ disposable income during these periods. In summer, without gifting holidays or financially-demanding social events, people may spend more freely on crowdfunding projects, represented by notable increase in the successful campaign rate. This is then curtailed in August when children return to school, as parents dedicate more money to school supplies, clothing, and extracurricular expenses, resulting in a sharp convergence of the two lines. Then in December the lines converge once more (though not as intensely), as individuals spend more on holiday gifts and perhaps less on their own interests. Regardless of the reasoning, it appears that a campaign that begins in June or July is much more likely to succeed than one started in August or December.

Moving to the *Success by Goal* worksheet, we see important shifts in success rate based on the campaigns’ initial goal. The first notable shift is a divergence between 1000 and 4999, suggesting that a creator may want to consider a goal of at least 1000, regardless of what is needed for the project. Likewise, campaigns with goals anywhere from 15000 to 34999 have very high success rates, while those between 10000 and 14999 and those exceeding 50000 show more failures than success. Based on this data, a creator would want to set a goal of, say, 15000 instead of 10000, as well as avoiding goals over 50000 – or at least acknowledging for such expensive projects that there is a considerably higher risk of failure. This stands to reason, as these projects will require more backers, more investment per backer, or both.

It should be noted, however, that the high success rate in the projects from 15000 to 34999 may be more indicative of campaigns run by experienced organizers who know how to run successful campaigns, rather than any psychological effect upon potential backers of the goal amount itself.

Limitations

The diversity of currencies is a notable limitation of this dataset. While 76% of the projects are based in the United States, and thus have currency amounts in US Dollars, the other 24% are from countries such as Great Britain, Italy, and Switzerland, with goal and funding amounts in their local currencies. This makes comparisons based on monetary amounts incompatible between countries. This is further conflated by the 10-year gap in campaign dates (ranging from 2010 to 2020) with no indication that these numbers have been adjusted for inflation. While the data could be converted to accommodate this limitation, as-is comparisons may have less value than they appear.

One unknown pertaining to this dataset is the collection method. Certainly this is not the full population of crowdsourced projects between the years 2010 and 2020, with only around 100 campaigns per year. What, then, was the sampling method? A simple random sample? A proportional stratified sample based on categories? Further, was the data collected from a single site such as Kickstarter, or an assortment of crowdfunding services? Without this information, any conclusions drawn from this data should be used with caution, as the sample may not represent the full population of crowdsourced projects.

The number of datapoints is also concerning. While a sample size of 1000 may be statistically sufficient (though we have not calculated that here), it leaves categories and subcategories with such small counts that any inference with that granularity may be dubious. For instance, the Audio and World Music subcategories show a 100% success rate, but they are represented by only n = 4 and n = 3, respectively. It is unlikely we can say that such projects are a “sure thing” based on this data. Even the aforementioned Mobile Games subcategory, with twice as many failures as successes, has only n = 13, which may not be enough to say that Mobile Games are actually high-risk projects.

Similarly, when looking at *Success by Goal*, the 15000 to 19999 grouping has a 100% success rate. But that is based on only 10 campaigns. So, while promising, it shouldn’t be considered a guarantee – though collectively, the 15000 to 34999 campaigns show a success rate of 92% over 38 campaigns, which seems more reliable.

Additional Visualizations

Several additional visualizations are provided in the Excel workbook to examine the relevance of certain data. Refer to these worksheets: *Backers Count*, *Percent Funded, Campaign Length*, *Staff Pick*, and *Spotlight*.

The *Backers Count* sheet shows histograms of the number of backers (in 500-count buckets) for both successful and failed campaigns. These illustrate the right-skew of the backers count, as explained above. While both graphs show a considerable majority of backers within the first bucket, the successful campaigns have larger counts in buckets further to the right, skewing the mean higher than the median.

On the *Percent Funded* worksheet is a chart showing the funding percentage for each parent category broken down by starting month. Based on this data, a new campaign would want to start at different times of year based on the type of project that is being funded. For instance, the Food category shows significant spikes in the months of June and October, suggesting that those months are best for starting a food-based campaign. Games, meanwhile, are shown to be massively overfunded on average with a March start date, but underfunded in April. So a matter of a few weeks may be important to the success of a campaign. But note that Percent Funded not only represents campaign success, but *how* successful a campaign is; the higher the number, the more money the campaign generated beyond its goal.

The *Campaign Length* sheet shows the distribution of campaigns based on the length of each campaign, in days. These charts show a similar distribution between successful and failed campaigns, hinting that the length of a campaign may not have a strong effect on its success, with one notable exception: the gap between the first and second buckets is proportionally larger for failed campaigns than successful ones, suggesting that it may be advantageous to let a campaign run more than five days.

On the *Staff Pick* sheet, we see a chart comparing the proportions of success and failure based on whether those campaigns were designated Staff Picks. The results are very close, with Staff Picks having success only 1.45% more often than non-Staff Picks. While further analysis would be required to know if this difference is statistically significant, this low percentage suggests that the impact of Staff Pick designation is negligible.

The *Spotlight* sheet shows a similar chart for Spotlight designation. Interestingly, this chart shows that Spotlight designations have lower success rates, though again, the difference is small (2.76%), and quite possibly within the margin of error should a statistical analysis be conducted. Therefore, Spotlight designation, like Staff Picks, likely has no meaningful bearing on a campaign’s chance of success.