|  |  |  |  |
| --- | --- | --- | --- |
|  | Jeff Pinegar | | jeffpinegar1@gmail.com |
| Assignment: | Excel Homework: Charting Crowdfunding | 717-982-0516 |
| Due Date: | September 29, 2022 |  |

**Crowdfunding Analysis**

# Overview

## Objective:

By analyzing a data set of 1,000 confounding campaigns, generalizable best practices and advice can be identified to help future campaigns achieve their goals.

## Data Source:

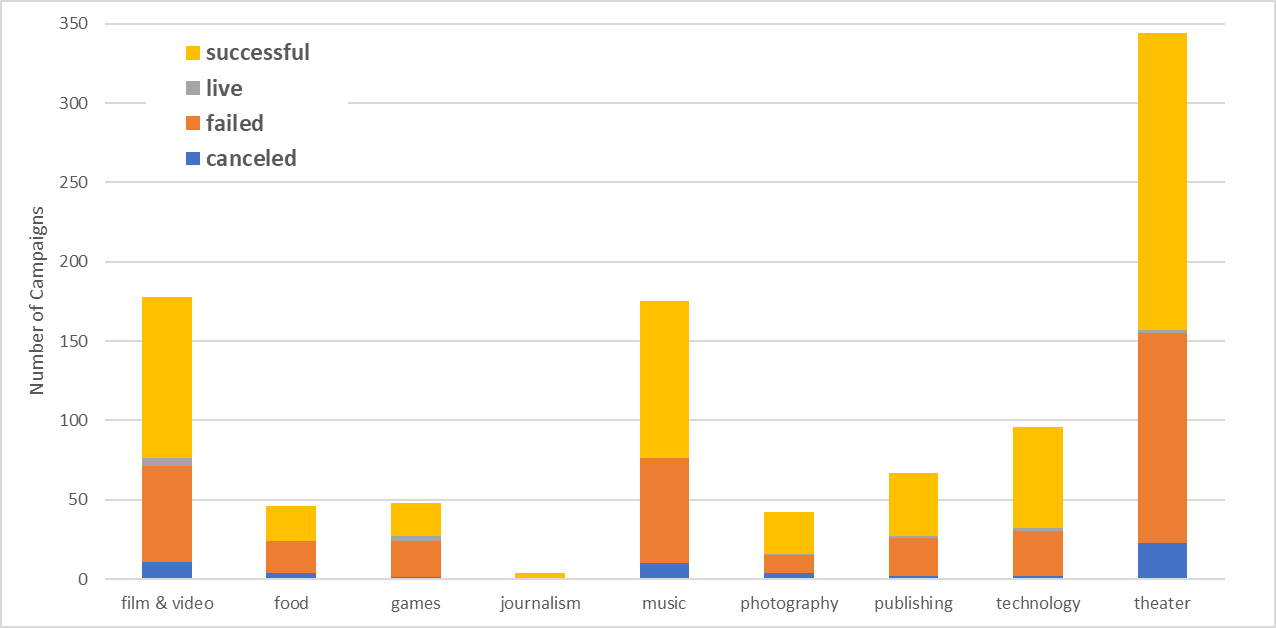
* Collection of 1,000 crowdfunded campaign (n=1,000)
* 57% of campaigns reached their goal, and 37% failed. The remainder were either canceled or still active
* Campaigns were conducted between Jan. 2010 to Jan. 2020 with an average duration of 15 days and a maximum duration of 59 days
* Campaigns are categorized into 9 categories and 24 subcategories
* Campaigns were conducted in 7 countries
  + 76% of campaigns were US based
  + The other 6 courtesies represent less than 5% each

# Crowd Funding Campaigns

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

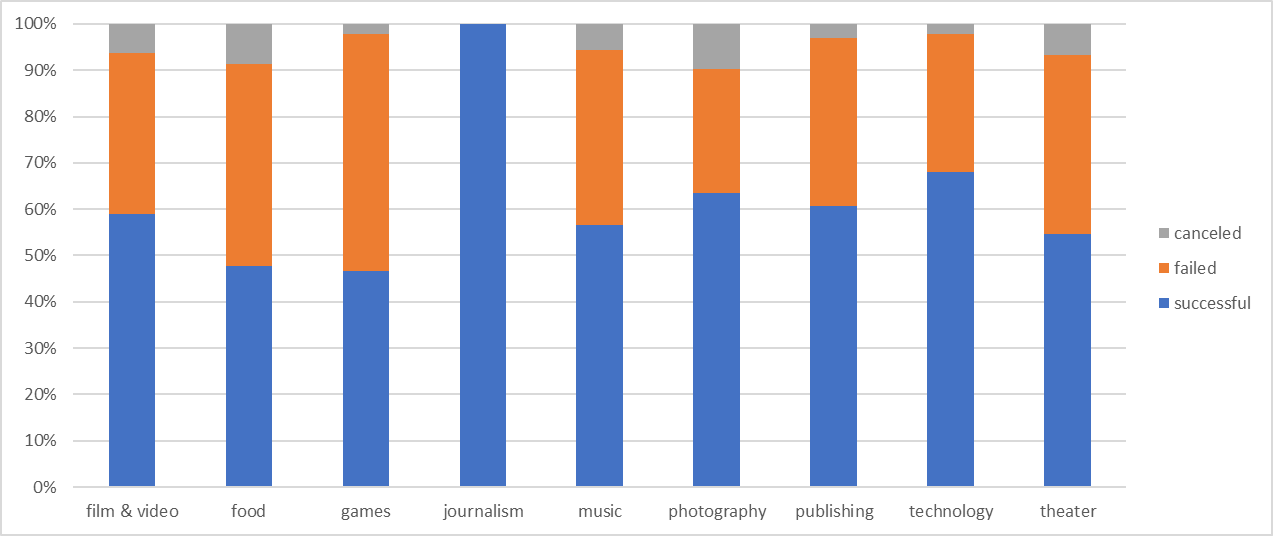
* The most popular campaigns are for the arts, including film, video, music, or theater. Together these categories represent 70% of the campaigns analyzed (see Figure 1). However, while more popular, these campaigns do not, on average, have higher success rates than other campaigns.

Figure : Campaign Outcomes by Category (Count)



* The campaign category appears to have minimal impact on the chances of success or failure. Most campaigns succeed between 45% and 65% of the time; see Figure 2. (Journalism was not considered because the sample size was n=4)

Figure : Campaign Outcomes by Category



* I only found one thing that would offer any insight into whether a campaign would be successful or fail. Campaigns that are both a Spotlight and a Staff pick appear more successful. See Figure 3. The other thing of note was that compared to either Spotlight or Staff Pick, there was a higher percentage of cancelation. I assume that combined Spotlight or Staff Pick contributes to a higher level of support from the staff. While there are only 10 campaigns that are both, we also observed results that are closer to the plan (less variability), perhaps another indication that these campaigns are managed better, see Figure 4. These campaigns also have higher cancellation rates and lower failure rates, indicating closer attention from management.

Figure : Outcomes of Campaigns that are Both Spotlight and Staff Picks

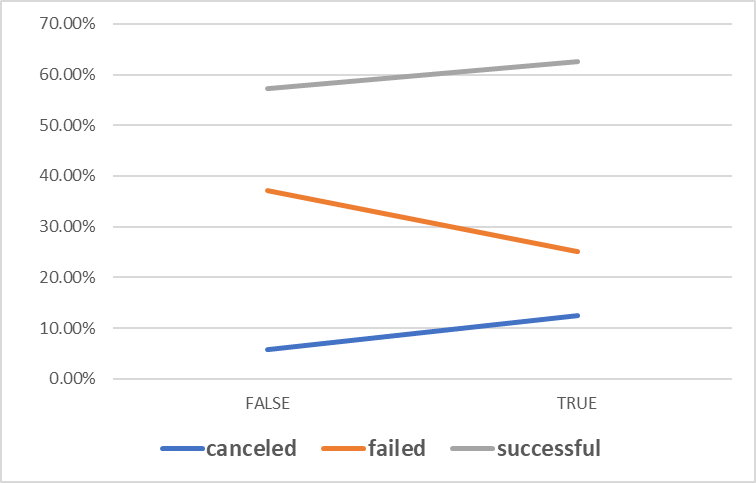


Figure : Analysis of Successful Campaigns that are both a Spotlight and a Staff Pick



* Successful campaigns have higher than average (mean and median) numbers of backers when compared to campaigns that fail, Figure 7. A Wilcoxon rank sum test confirmed that the medians differed (α=0.005). However, this difference is not clear simply by looking at the histograms (Figure 5 and Figure 6), with both distributions having long tails and significant variations measured by standard deviation. In any case, the Median is a more meaningful summary of the data because of outliers (a few campaigns with a very large number of backers).

Figure : Analysis of Number of Backers of Successful Campaigns

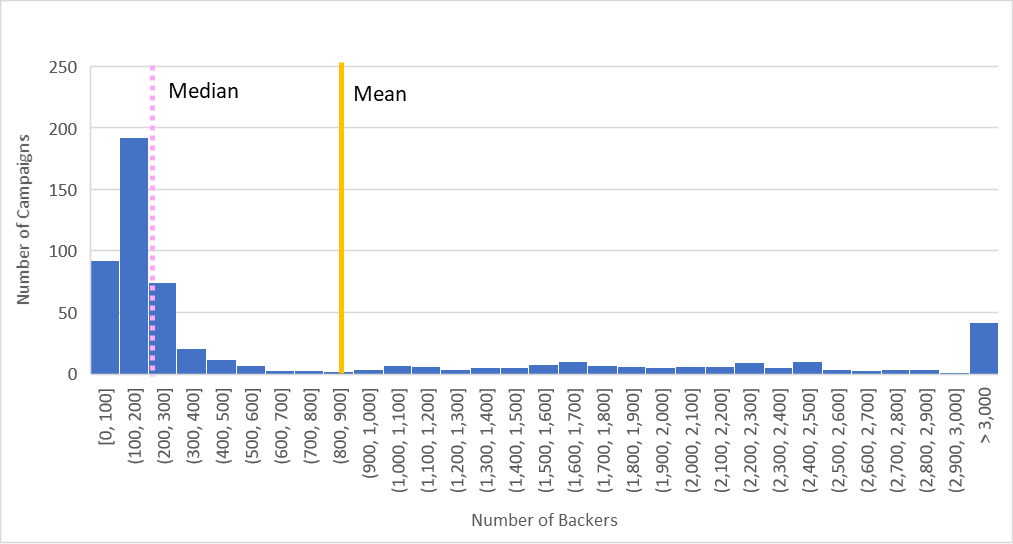


Figure : Analysis of Number of Backers of Failed Campaigns

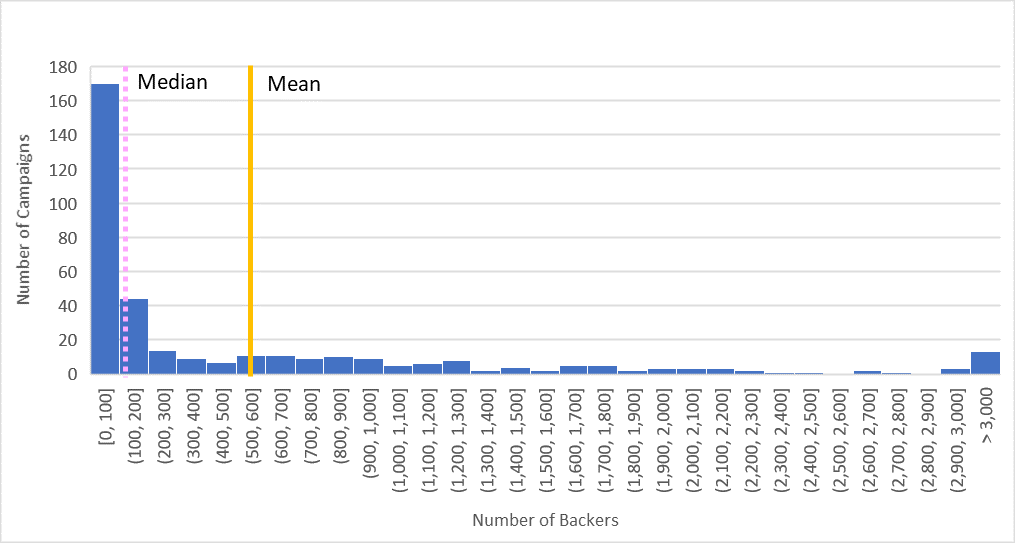


Figure : Statistical Comparison of Backer Participation Rates on Success



* There appears to be an exponential decay in the relationship between the Campaign goal and the percent of funding achieved. More specifically, there seem to be two different relationships. One that applies for campaigns up to $10,000 and one that applies for campaigns over $10,000. See Figure 8 and Figure 9.

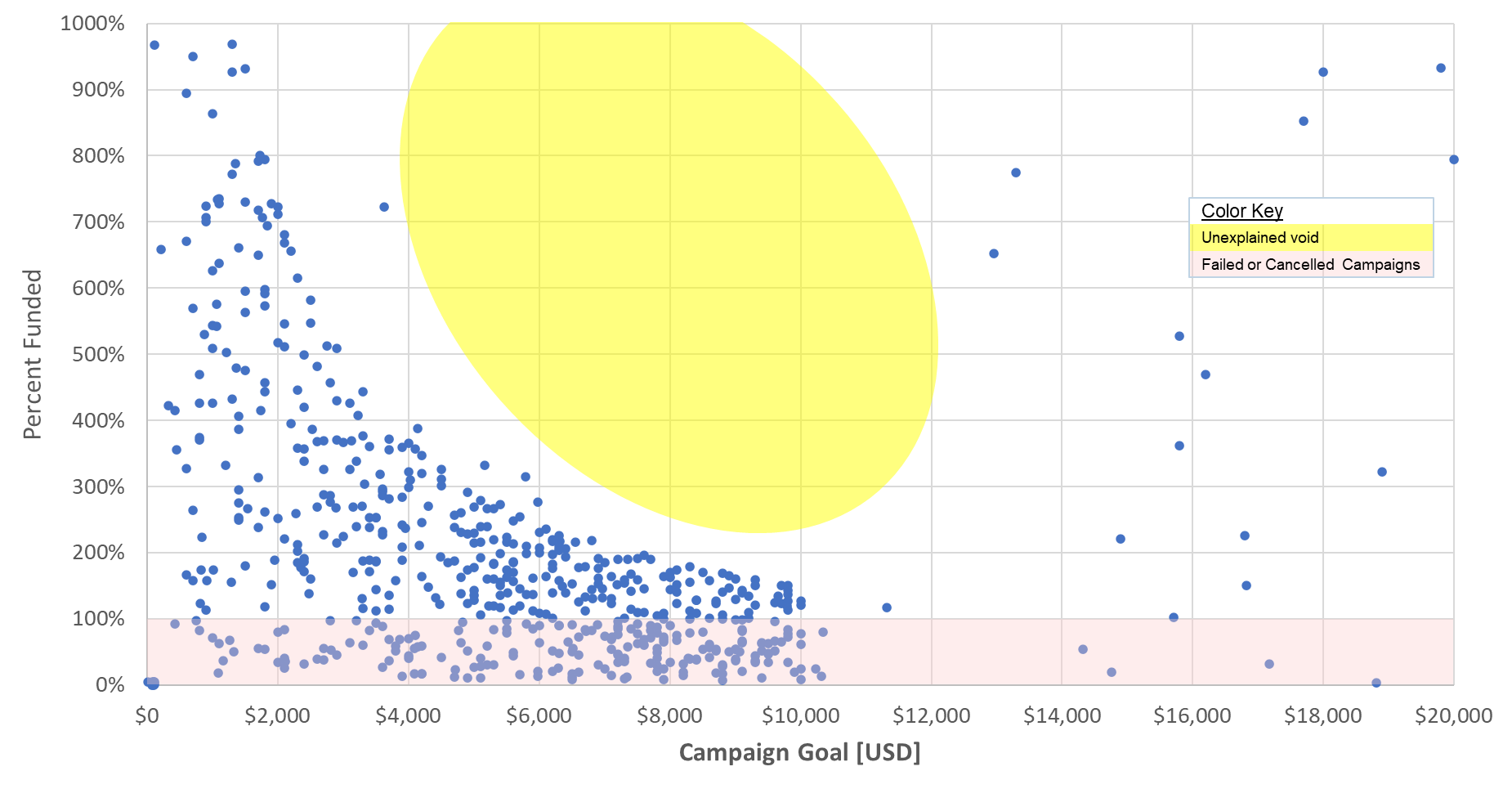
Figure : All Campaign in the dataset



Looking more closely at campaigns under $20,000 (Figure 9), we see a void of highly successful campaigns (those achieving more than 300% of their goal). Of course, more reach would be needed to explain this void, but several theories could be investigated.

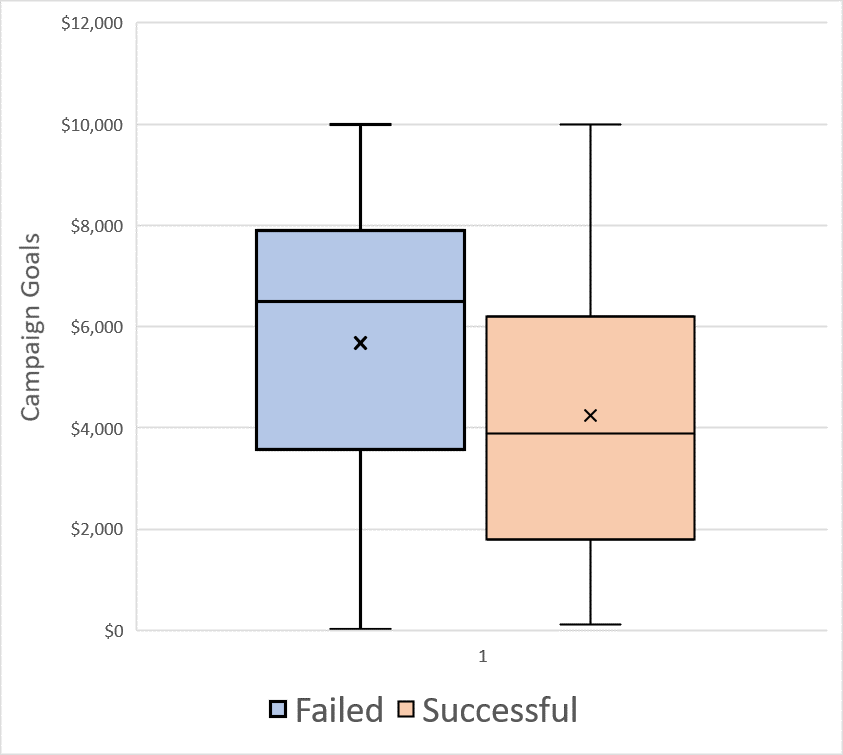
* + Smaller-scale campaigns are less emotionally engaged topics, so people stop donating once the goal is achieved. On the other hand, substantial medical bills, natural disasters, and mass casualty events are emotionally charged and widely communicated.
  + Smaller scale campaigns have less market momentum, so they rapidly decline once the goal is achieved
  + The threshold of $9,999 is a physiological barrier that is difficult for campaigns to cross.

Figure : Campaigns up to $20,000



Continuing to look at the campaigns with goals under $10,000 (61% of all campaigns), we see that 75% of the successful campaigns have goals less than approximately $6,000. See Figure 10.

Figure : Box Whisker - Campaign Goals less than $10,000



# Limitation of the Dataset and this Analysis

All data sets and analyses have limitations. The following are some of the limitations of this analysis.

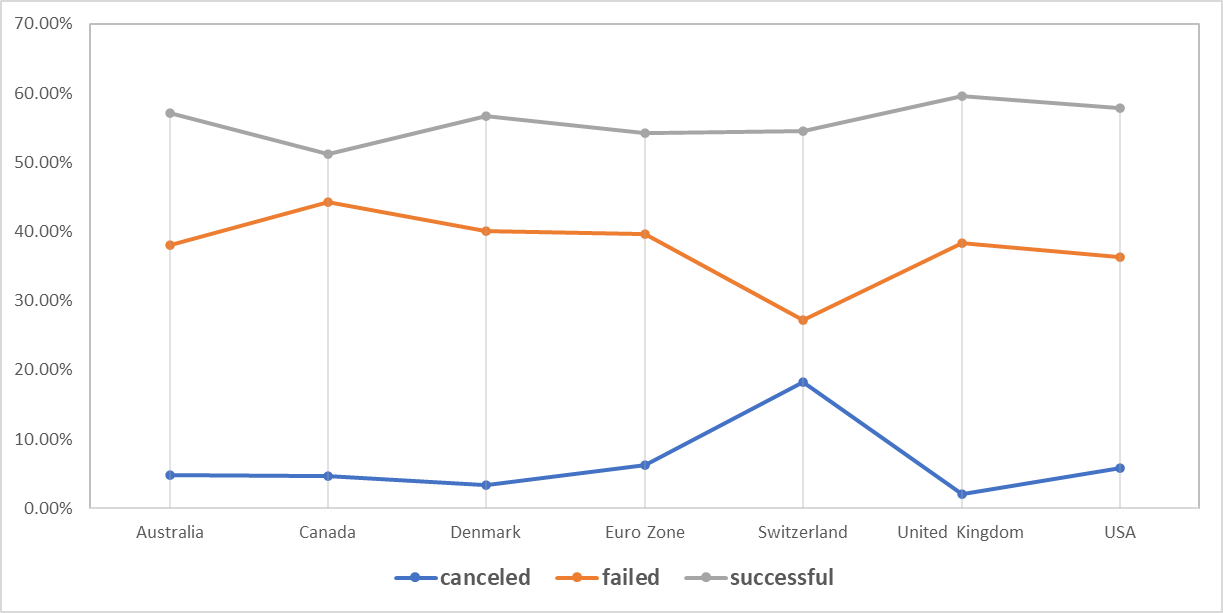
* We do not know the sampling methodology used to collect this sample. Without this information, it is impossible to determine if the sample is representative of any community, group, or organization (the market as a whole or even a single subgroup). For example, this could be a sample of 1,000 Kickstarter campaigns with goals under $200,000 and durations less than 2 months, promoted in Toledo, OH, to benefit minority-run charities. Or the sample could be a random sample of all campaigns run on the 20 largest crowdfunding sites with no limitation on the duration or goal size. The generalizability of these two data sets is very different.
* This data set has many limitations, making it risky (error-prone with high variability) to make inferences about the success or failure of a potential campaign. In general, you should not make inferences about campaigns outside the sample set’s bounds. Some of these limitations are:
  + We do not know which site or sites this data set came from, Kickstarter, IndieGoGo, Crowdfunder, or one of many other sites. Therefore, the results are likely to vary from site to site.
  + The largest goal in the set was about $200,000. Therefore, the data set should not be used to make inferences about campaigns with goals greater than $200,000.
  + The longest campaign in the set was 113-days. Therefore, the data set should not be used to make inferences about campaigns lasting longer than 113-days.
* There are a limited number of observations from countries other than the USA. No other country has more than 50 observations (or 5% of the sample). See Figure 11. With this limit, it is not possible to test for within-group differences. So, for example, you cannot say if being a Spotlight campaign makes a difference in Canada or not.

Figure : Country Participation



There is limited variability of outcomes across countries (see Figure 12); however, there is very little data for countries other than the USA. The noticeable exception is Switzerland. In this country, it appears they are more likely to cancel a campaign that is destined to fail.

Figure 12: Country Outcomes



* This data set does not include information on the distribution of donations to the campaign. We calculated the average (Total pledges / Total # of Donors), but we don’t know if these were normally distributed. The descriptive statistic indicated there was a difference in the average donations to successful and failed campaigns. See Figure 13. The distributions of donations (Figure 14 and Figure 15) were approximately normal, enabling us to run a Two-sample t-Test with an unequal variance assumption to confirm the statistical difference. See Figure 16.  
    
  While this is statistically interesting, it is not that helpful in practice since it proves that campaigns with higher average donations are more successful, which should be intuitively obvious.

Figure : Descriptive Statics for Average Pledges



Figure : Analysis of Average Donation to Successful Campaigns



Figure : Analysis of Average Donation to Failed Campaigns

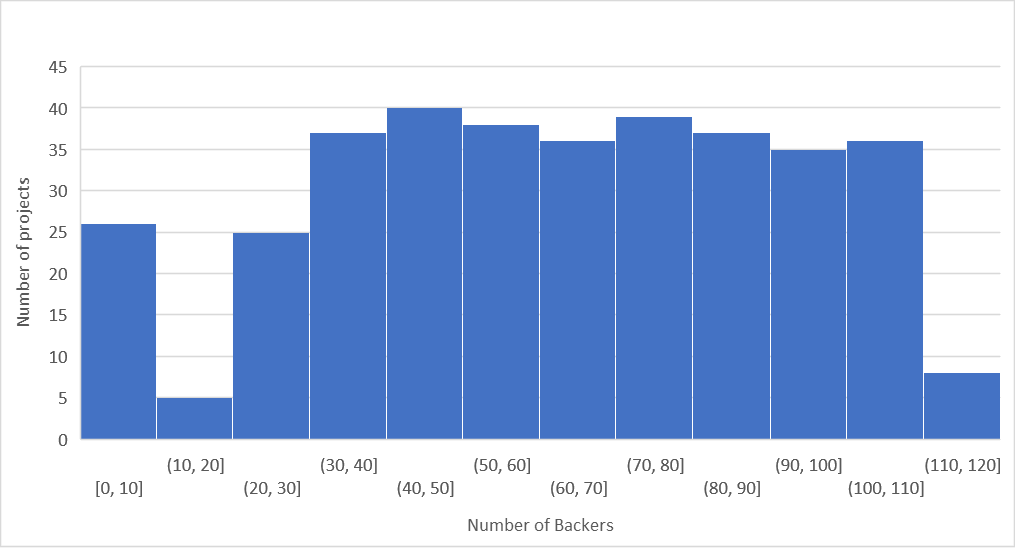


Figure : Two Sample t-Test for a difference in Mean



# Further Analysis

Crowdfunding success is often driven by either an emotional appeal or delivering a benefit to the backer/donor. An example of an emotional appeal would be raising money for hurricane victims in Puerto Rico. On the other hand, a campaign for the creation of a new technology/product would most likely be something that would benefit the donor. In either case, the description creates this connection to the potential donor. Therefore, I would consider the textual analysis of the “blurb” or hopefully the full text of the campaign description. I believe this would offer the most helpful advice for future campaigns.

To predict success/failure, I would attempt to do a discriminant analysis (similar to multiple regression, but the dependent variable can be categorical (success, fail, cancel). I would start with the independent variables of Goal, Country, Staff-Pick, Spotlight, Duration, Start Month, category, and subcategory. While individually, they did not appear to have predictive value, collectively, they may prove valuable.

Some of the additional tables, graphs, and observations that I did create include:

* Standardize all the currency data to US Dollars. Standardization is essential for analyzing the effects of goal size and average goal size. Alternatively, the research could have been done by excluding campaigns not denominated in USD.
* Histograms of the number of patrons for successful and failures
* Histogram of the average size of donation for successful and failures.
* Analysis of the effects of Spotlight status on success
* Analysis of the effects of being a Staff Pick on success
* Analysis of the effects of the interaction of Spotlight status and being a Staff Pick.
* The 80/20 rule of thumb was observed to be very visible in this data set.
  + 80% of the campaigns are less than $100k
  + 80% of the campaigns have less than 1,300 donors
  + 80% of the campaigns have an average donation of $100 or less.