Crowdfunding Base Analysis

One of the odd conclusions we can draw about crowdfunding campaigns is that most campaigns originate from the US. However, a country-to-country comparison may raise issues or concerns about sample size compared to other countries. For example, Great Britain holds a greater success rate for crowdfunding campaigns, but the total number of campaigns is much lower than in the US, which may raise sample size questions per CLT.

The following two conclusions draw from my experience in sales and the use of indices to compare different groups. The consolidated definition of an index is that anything over 100 represents % more likely, and anything less than 100 represents % less likely (propensity-wise).

The second conclusion we can draw from the parent categories is that theater is the most common crowdfunding campaign type; however, it does not perform as well as other categories compared to the success of all crowdfunding campaigns. Using an index calculation with the total number of campaigns as our base, we can see that campaigns in the theater category are 4% less likely to be successful. In contrast, the technology categories are 18% more likely to be successful than all successful campaigns across all categories.

My last conclusion is that campaigns with goals under 40,000 are 63% more likely to be successful than those with goals greater than 40,000. I include my index calculation as a separate mini table in the Crowdfunding Goal Analysis tab. I take the same calculation steps for the previous conclusion.

Regarding limitations, two of the metrics I would have liked to see within this dataset are:

1. Did the campaign team produce a roadmap for their project at the campaign's creation?
2. What types of outreach or advertisement did the campaign team perform for each project? The staff pick and spotlight columns represent a small widget of outreach. A significant crowdfunding component is getting the word out about the project, so a metric indicating outreach is crucial.

I would include tables and graphs in this analysis that render and compare the staff pick and spotlight data correlating with the number of backers and campaign success. We could use both as binary indicators within a regression as well.

Statistical Analysis

We can see more variability within successful projects by looking at the statistics for the number of backers between successful and failed crowdfunding campaigns. After looking at a scatter plot and a bin analysis, I hesitate to suggest using the mean to represent the data accurately. Within both groups, failed and successful, there are far too many grand outliers to indicate a normal distribution.

Nonetheless, the two-tail value on the heteroscedastic t-test indicates that the means between the two groups are statistically significant. Regarding the average, we can step away from the null hypothesis and suggest that there is a difference between the two groups.