**Crowdfunding Goal Analysis Report**

Since the late 2000s, crowdfunding has been growing in success and popularity. As a result, many people are now using this practice to launch new products and generate publicity. In this challenge, I analyze crowdfunding data from a database of 1,000 sample campaigns to better understand its trends.

After assembling the required spreadsheets, I merged the pivot tables from the Primary Category by Outcomes and Sub-Category by Outcomes spreadsheets and calculated the successful campaign percent per category; but omitted primary categories and sub-categories with less than five campaigns: due to low numbers, these categories provide an insufficient sample space for evaluation (see Both Cat. By % Success). From my observations, a crowdfunding campaign in a primary category has a 56% chance of success with technology, photography, and publishing at the top; sub-category campaigns also have a 56% chance of funding where web (technology), translations (publishing), and television (film & video) are the leaders. In terms of sheer numbers, theater, film & video, and music are the most popular primary categories by significant margins; plays (theater), rock (music), and documentary (film & video) are the most popular sub-categories. Although this effort revealed some insights, there was still an incomplete picture.

To inquire further, I created a pivot table with both primary and sub-categories versus the sum of the campaign goal, the sum of the pledged contribution, and the percent of the pledged contribution over the goal (see Both Cat. By % Funding). Funding percent above the goal for all campaigns regardless of status or outcome is 97% with technology, music, and theater being the highest primary categories; the highest sub-categories are metal (music), wearables (technology), and jazz (music). For successful projects, the overall funding rate is 191% with the primary categories of food, games, and photography and the sub-categories of metal (music), jazz (music), and science fiction (film & video) being the highest; in particular, among the primary categories, successfully funded games and food campaigns receive 2.1 and 2.3 times their goals respectively; among the sub-categories, successful jazz (music) and metal (music) campaigns receive funding at 3.5 and 4.4 times their goals.

In the Crowdfunding Goal Analysis spreadsheet, I examined outcomes based on ranges of goals only to encounter specious results. In the graph, the successful campaign rate’s trend increases from $0 to $15,000, plateaus between $15,000 and $34,999, and then declines from $35,000 on. Overwhelmingly, the lowest three categories have the highest number of projects: all the categories equal to or above $10,000 have less than 15 projects compared to those below $10,000, which have two categories with hundreds of projects and one with about 50. For the larger goals, these low numbers of campaigns per category provide an insufficient sample space, which challenges the integrity of the analysis results; unfortunately, these circumstances are present in other parts of the dataset as well.

The Statistical Analysis spreadsheet exhibited distributions for number of backers for successful and unsuccessful campaigns with significant ranges between minimum and maximum values, large variances, and considerable differences between average and median values. When the average and median values deviate as they do in this case, the distribution is skewed, and the average no longer is an accurate representation of the typical or middle value of a set of values: in this instance, the median becomes the more accurate measure and better summarizes the data. Regrettably, the calculations in this analysis are averages when they should be medians. Moreover, the variances show that successful campaigns have more variability than unsuccessful ones. Specifically, variance measures heterogeneity: a higher variance means the values are more spread out, and a lower variance means that the values are closer together. This determination makes sense because there is a large spectrum of backers for successfully funded campaigns while unsuccessful ones have concentrated numbers of backers closer to and including zero.

Therefore, in addition to obvious deductions about success and funding trends, I made the following three conclusions about crowdfunding campaigns.

* The probability of a successful campaign in a particular category is not related to how popular that category is in terms of numbers: the categories with the most campaigns are not necessarily the most successful ones.
* The variability in successful campaigns is greater than that of unsuccessful ones because the values in the distribution are more spread out in the former and congregated and closer to zero for the latter.
* The two distributions of backers, for successful and unsuccessful campaigns, are heavily skewed, which the egregious deviation in mean and median values conveys. Hence, in this case, the median is a better measure of central tendency than the average, but the average is used instead. This situation coupled with insufficient sample spaces discussed in the report limits the usefulness of this dataset.

In addition to the analysis conducted, there are numerous other tables and/or graphs that we could create. For the sake of brevity, I will mention only three: countries vs. percent of successful campaigns, year vs. percent of successful campaigns, and outcomes vs. standard deviations of goals and pledged contributions. In countries vs. percent of successful campaigns, the table and graph show if campaigns are more likely to succeed in one country over another (see Countries by % Success). For year vs. percent of successful campaigns, the calculations demonstrate whether crowdfunding is becoming increasingly more popular or following a cyclical pattern (see Years by % Success). Finally, outcomes by the standard deviations of goals and pledged contributions will show if anything differentiates distributions for a successful campaign over the alternatives; there is: the standard deviation of the campaign goal is greater than the standard deviation of the pledged contribution in all cases except successful campaigns where the reverse occurs. What’s more, the standard deviation for a campaign goal is approximately the same for all outcomes except successful campaigns where it drops precipitously (see Outcome by StdDev Goal-Pledge).

Altogether, this investigation yielded practical results, but the limitations of the dataset render those results suspect. In the future, I would suggest a similar analysis with more tables and graphs on a much larger dataset using medians for calculations instead of averages.