Brian Silfer
Excel Challenge
Written Report

Finding financing for creative projects is difficult, especially with how broad of a topic range these projects fall under. After analyzing 1000 sample projects, I discovered 3 trends that would give future creatives and entrepreneurs a leg up on their next passion project by increasing the chances of a successful campaign before the launch date.

My first conclusion from my analysis of the data set focuses on the data collected regarding the total number of campaign outcomes (canceled, failed, live, successful) over the period 2010-2019. Successful campaigns occurred more frequently than failed campaigns across all months within that time frame. When looking at the total number of campaigns broken up by outcome, displayed as totals for each of the 12 months of the year over the 10 years, campaigns that launched during the summer months of June, July, and August, were more frequently successful than campaigns launched in any other month of the year. This insight would help an entrepreneur set the launch date of their next venture. Launching campaigns between June and August could lead to more successful campaigns as people are possibly more willing to spend money during the summer months. Although there could be many interpretations of this point, it is clear that more successful campaigns are launched between June and August than any other months during the year.

My second conclusion from analyzing this dataset is that when studying the count of each outcome per sub-category, plays had the highest number of successful campaigns, as well as the highest number of failed campaigns. Analyzing the pivot chart for this data could lead an inexperienced analyst to conclude that plays are the most successful campaign subtype, as plays also had the most overall number of campaigns. The data shows a large sample size for plays in comparison to other sub-categories. I wanted to investigate whether plays had the largest success rate compared to other categories. After analyzing the success rate of each sub-category, I found out that the Web sub-category had the highest success rate of 70.6 percent while plays had a success rate of 54.4 percent. This is a drastic difference and would be a great piece of information for a prospective entrepreneur looking at the data and understanding that even though the majority of campaigns are plays, there are more successful campaign sub-categories. I removed audio and world music from my analysis because they had a success rate of 100 percent. Both subcategories had a combined sample size of 7 campaigns whereas web had 51. This is an interesting detail, as there is possibly an underlying trend, showing audio and world music as two potential sub-categories to look into for a creative's next crowdfunding venture.

My third conclusion involved additional analysis. I analyzed the success rates of parent categories. Again, even when looking at visualizations of this data, it is easy to look at the stacked columns and say theater is wildly popular among crowdfunding campaigns. The nature of a stacked column pivot chart is that they aren't very good at describing how successful a campaign category is compared with other categories, especially when most categories have a

dramatically lower campaign count than theater. My initial analysis can be misleading as the theater had a success rate of 54.4 percent while the most successful parent category was Technology, with a success rate of 66.7 percent. My analysis of parent categories helps prospective crowd funders realize that just because theater seems to attract the most crowdfunding campaigns, doesn't mean there aren't other parent categories that are more successful.

Some limitations of this dataset include the massive differences in sample size category to category. Journalism has 4 total campaigns in the dataset whereas theater has 273. This creates several limitations when trying to visualize this data as well as creating analysis based on the count of outcomes and not the average of each outcome. Another limitation is that data from 2020 only has 2 entries. If I was trying to use this analysis for a crowdfunding campaign today, I would be forced to use data that could be out of date, not accounting for any shifts from years 2020-2023. The other limitation is that the data set includes data from several countries, yet has a sample size of 736 for the United States, and a combined sample size of 264 for the other countries. Any crowd funder from outside of the US would not be able to gain a lot of useful insight from this dataset. If the dataset is going to include data from other countries, there should be more equality in sample size to make the data more useful for an international audience.

Some additional tables and charts that could be created are a pivot chart showing success rates filtered by country, parent category, and sub-category. This pivot chart would allow an entrepreneur to easily filter through success rates of categories giving them insights into the road ahead for their campaign. This chart would also help even the playing field for all categories, as now campaigns can be viewed with the difference in sample size in mind. Another chart that would be useful would display each outcome, filtered by range, allowing you to analyze the outcomes in more detail. The ranges are fairly broad and are not designed well for understanding why certain categories fail miserably compared with others that barely fail. Especially when looking at the campaigns that have a low percent funded value and therefore a failure, it is not shown well in any visualizations that portray counts of outcomes as false. Being able to look closer at campaigns that failed could allow a crowd funder a better understanding of where their last campaign went wrong.

Analyzing crowdfunding data showcased the need to dig deeper and expose more useful metrics such as success rate, rather than making decisions based on the count of outcomes alone. Looking at the success rates of each parent category and subcategory allows us to uncover that technology and web campaigns are by far more effective than theater and plays, but receive a lot less campaign traffic from the crowdfunding community.

Statistical Analysis:

The median summarizes the data best, as there are several outliers in the successful and unsuccessful data. Several of the outliers are several times higher than the values within the IQR and would skew the mean more towards the upside. This would severely impact the summarization of the data.

There is more variance in the successful data as there are several positive outliers for backer counts for successful campaigns. Most of the successful outliers occur within the range of 3000 to 7000 whereas unsuccessful outliers occur in the range of 2000 to 4000. Higher outlier values create a larger variance in the data.