Crowdfunding Analysis

Given the provided data, three conclusions can be drawn about crowdfunding campaigns: the highest rate of campaign success is seen in campaigns launched in July; theatre has the highest number of successful campaigns among the parent categories; and the parent category with the highest success rate is journalism. July has a total of 58 successful campaigns. According to this data, launching a campaign in July will yield the highest success rate. Theatre has a total of 187 successful campaigns. Although theatre has the highest number of successful campaigns, its success rate is 54.36%. Journalism has the highest success rate, which is 100%. It is important to note these statistics are not weighted; theatre has a total of 344 campaigns while journalism only has 4.

The limitations of this dataset are year relevancy and not being weighted. The dataset primarily includes older campaigns, ranging from 2010 to 2020. This is primarily reflecting pre-pandemic trends. There are only two campaigns that were launched in 2020; it is hard to draw definitive conclusions about 2020 campaigns. This poses a problem to stakeholders who want to know the current trends. In addition to year relevancy, the data is not weighted. The number of campaigns within each category vary greatly, ranging from 4 to 344. This means each category is not represented evenly. I believe more data should be collected; this would offer updated trends and encourage an evenly distributed representation among categories.

I created additional graphs to display the success rate per category and percentage of campaigns per category. The success rate per category graph allowed me to analyze how successful each category was, since the total amount of successful campaigns per category does not offer insight on this. Results for the number of successful campaigns per category can be a bit deceiving, due to the varying number of campaigns within each category. The graph depicting the percentage of campaigns per category showed me how each category is represented within the dataset. The number of campaigns within each category is not distributed evenly. Overall, the dataset offers a good foundation of knowledge regarding campaign success.

Statistical Analysis

Two conclusions can be drawn from the statistical analysis: the median better summarizes the data, and there is more variability within successful campaigns. There are many outliers among the backers count in both successful and unsuccessful campaigns, which means there is not a symmetric distribution. The median is more reliable when there is not a symmetric distribution. As for variability, successful campaigns have more variability than the failed campaigns. This is a reasonable conclusion, because the success of each campaign is defined differently; some campaigns have a higher goal or a higher average donation. Each campaign is unique, and the number of backers does not determine its outcome. Therefore, you can expect there to be high variability within the data.