**Crowdfunding Report**

By Natalia Mitchell

Given the provided data, three conclusions that we can draw about crowdfunding campaigns are:

1. Entertainment categories (film & video, music, theater) had the most crowdfunding campaigns within the dataset, making up about 70% of the total. This tells me that more than likely, more crowdfunding campaigns in general will be started for entertainment ventures.
2. Crowdfunding campaigns with goals between $15,000 and $34,999 had the highest percentage of success rates, with cancelation and failure rates increasing with most goal ranges above and below those numbers. This data suggests that if a campaign keeps their goals within that monetary range, they have a higher chance of success.
3. Campaigns that ran between May and July also had a higher rate of success, with the percentage of successful campaigns peaking in July and then dropping back down again in August. We can assume that crowdfunding campaigns run in the summer months may have a better outcome.

Limitations of this dataset are, firstly, the geographical location of the campaigns. There are more campaigns listed in this dataset from the United States than any other country (763 out of 1000). Getting a dataset with an even number of campaigns across all seven countries, I believe, would represent a more accurate analysis globally. Secondly, we do not know how any of the campaigns were marketed, and to what audience. For example, campaigns for the technology category, which had one of the more successful outcomes, may have been marketed to an older, wealthier population, whereas campaigns for games may have been marketed to a younger, lower income population, which could have contributed to the higher failure rates in that category. Lastly, the range of dates that these campaigns took place starts in 2010 and ends in 2020, with between 2 and 109 campaigns listed per year. I think a more accurate dataset would be to look at campaigns run within a reduced range of dates, or to have an even number of campaigns analyzed within each year.

Some other possible tables and/or graphs we could create to add value to this analysis would be to look at the correlation between the timeframe of the campaign (date created and date ended) to the outcome and percentage funded. This could show us if campaigns that had longer funding windows did better overall. We could also look at average donations to see if there were any categories that had higher averages, and we can filter them by country to see which campaigns had higher or lower averages in each location.

Statistical Analysis: I believe that using the median better summarizes the data because there are very large variances within both successful and failed campaigns. There is more variability with successful campaigns. This makes sense because looking at the data, the total campaign and backer count is much higher, there is a larger gap between minimum and maximum backer count, and the standard deviation is higher.