**Kick Starter Data Analysis Report**

**Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**

Given the data provided for these Kickstarter campaigns, one conclusion that can be drawn is as the financial value for a Kickstarter goal increases then the odds of failure for that campaign increases as well. Kickstarters that set a lower goal for funding seem to have a better chance of succeeding and reaching their set goal. Campaigns attempting to raise less than $20,000 have a higher chance of succeeding than attempting to receive more than $20,000 on average and any that attempt to raise more than $45,000 will have a much lower chance of hitting their target goal.

Another conclusion that can be draw from the data is that different times of the year warrant different success, failure, and cancelation rates for Kickstarters reaching their funding goal. For example, between January and February across the entire timeline of years given from the data displays a good chance of success and a low chance of failure within these months. Between March and May, chances of success for a campaign are high with failure rates lower but still growing. Between June and September, the chances of campaigns receiving their target funding goals for success decreases steadily with November and December being at the lowest chance for a Kickstarter to become successful.

The third conclusion drawn from the data is that the category of Kickstarter could correlate to the chance of its campaign becoming successful, a failure, or cancelling. For the music category, the ratio of successful campaigns to unsuccessful campaigns is roughly 4:1 (rounding down) with 20 cancelations. The food category seems to have the lowest ratio for success to failure being roughly 1:4. This could demonstrate that the type of category a Kickstarter chooses will warrant better results of success or results for failure.

**What are some limitations of this dataset?**

A possible limitation of this dataset is that the data could be incomplete in some sections. There could have been more kick starters that did better in some categories that were never reported given a larger sample size. Another limitation is if the data given in the spreadsheet was taken from surveys from each individual campaign. There’s the possibility that honest answers were not always given which would give a misrepresentation of the data. This would cause the data to be skewed and inflate or deflate funding outcomes for some Kickstarter campaigns results for their goals. Another possible limitation is if the data formatted across platforms caused it to change the results in the spreadsheet we were given. Differences in duplicate records, inconsistencies, and formatting across merged data fields could cause skewed data results giving inaccurate conclusions.

**What are some other possible tables and/or graphs that we could create?**

One possible table that could be created could be used to see a correlation between success/failure rates and the average project duration time for each campaign. This would help see if projects that took more time from start to finish using timestamps received a higher or lower success rate and vice versa.

Another graph that could have been beneficial is a scatterplot graph that reveals a trend between successful and failed campaigns in relation to their target goals for funding. This would hopefully reveal a linear, logarithmic, or exponential pattern based on the statistics collected to give a better insight in regard to analyzing the data.

**Bonus: Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

There are many duplicate integers for successful and failed campaigns in the backer’s report, thus the mean would be ideal for this situation. The mean gives the average number of backers from this large dataset for all the successful campaigns and failed campaigns. The median would be good to use if the outliner didn’t provide such a large sample of middle statistics that are duplicates for the amount of successful and failed integers representing each campaign.