**Crowdfunding Report**

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

1. Based on the data that we have one conclusion that can be drawn about crowdfunding campaigns is that they are more likely to succeed then to fail/get cancelled. Of the 986 completed crowdfunding campaigns we looked at (not including the 14 still active campaigns) 565 were a success (57%), 364 failed outright (37%), and only 57 were cancelled (0.058%).
2. We can also conclude which kick-starters were most common, most successful, and least successful. The kick-starters in the parent categories that are most common are Theatre 344/1000 (34.4%), film & video 178/1000 (17.8%), and music 175/1000 (17.5%). The most successful parent categories were technology 64/96 times (67%) and photography 26/42 times (62%). The least successful parent categories were games 21/48 (44%) and food 22/46 (48%) which both finished at a total success rate of less than 50%. The kick-starters most common in the sub-categories are plays 344/1000 (34.4%), rock 85/1000 (0.085%), and documentary 60/1000 (0.06%). The most successful sub-categories were audio and world music which both had a success rate of 100%. The least successful sub-categories were science fiction 5/14 (36%) and mobile games (31%). With this information investors can better conclude whether a certain type of kick-starter is more or less likely to be successful.
3. Based on our monthly data we can conclude that the best months to start a kick-starter would be June & July. With 55 & 58 successes respectively. This is likely the result of potential investors having more free time during the summer months in order to do research on the product/utilize the product they are investing in. Based off this data any company looking to start a kick-starter may want to consider doing so during this peak time.

What are some limitations of this dataset?

1. One limitation with our dataset is certain parent categories/sub-categories aren’t large enough to get accurate (or at least as accurate compared to the other more prevalent categories) information from. For example, the audio and world music sub-categories both had a success rate of 100%. However, because the total amount of kick-starters was 4 and 3 respectively that 100% doesn’t give us a very clear picture. Plays had only a 54.4% success rate which on paper seems much lower but it was out of 344 total kick-starters so it actually gives us a much more accurate representation of how a kick-starter in that category might go.
2. All of our data is from kick-starter but it would be helpful to look at other crowdfunding campaigns to compare results and get a broader view of how successful they are as a whole.
3. We are taking data/currencies from multiple countries but when we are looking at the values we are not adjusting/accounting for exchange rates in the currency which could massively affect how much someone is paying for a Kickstarter with say GBP vs CAD.

What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

1. We could have made a scatter plot to better show the relationship between two parts of our dataset. For example, we could have plotted the kickstart outcome on our Y-axis (independent variable) and see how it is affected by different aspects of the dataset. if we used backer count as our X-axis (dependant variable) then we could see how the # of backers affects the outcome of the kick-starter. We could then do the same for average donation, country, etc…
2. Could use a box-and-whisker plot to visually identify potential outliers within the dataset.