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

1. One thing I noticed is that when you look at the start date of a campaign there seems to be a great deal of successful campaigns in the summer months. From May to August. Then tends to level out throughout the rest of the year. As for failed campaigns it would seem like starting in the later months of the year (after August) isn't a good time to start a campaign. This could be implying or showing that the best time to start a campaign is in the summer months.
2. Another conclusion to come to is which campaign is best to start or has the greatest chance of success. We can see in the category pivot table that the top 3 categories that have the greatest chance are Technology, Music, and Theater. The percentage of success of calculated by (successful/total)*100. The percentage of success can also be seen on the graph as there are bigger green bars in these categories compared to the others. It is notable that film & video also have a good chance at success. (Not counting the success rate of journalism because there is not a big enough population to determine how successful that category can be.)
3. We can also see the popularity of campaigns. Looking by category we can see that theater is the most popular campaign that people go after. Then when we look at the subcategories we can see that plays are the most popular campaigns to go after, which aligns with our category graph.

What are some limitations of this dataset?

One limitation I found was that all the categories do not have an equal amount of campaigns per category. So it can be argued that some categories don't have enough information to determine their chances of success. So for example. All the journalism campaigns have been successful. Meaning that journalism has a 100% success rate for campaigning which doesn't seem realistic. Speaking of another data set that limits us, in this case I don't see how the "live" campaigns help

us to tell a story here. It can skew our chances of success per campaign and also give no context on where or not the “live” campaigns are doing well or are being canceled or are failing.

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

I think a combination of backer engagement and funding over time would be a really good graph to add because we can see how much of an investment on average a campaign by category would need to be successful. We could see how much backer engagement is needed to have a successful campaign. It would just be useful data to have so we can determine the factors of what makes a successful campaign.

Another factor could be the politics or limited resources of countries. I think it would be useful to have a bar chart showing the success rate of a campaign per country. This would be very helpful for people deciding to work on international campaigns allowing them to decide which countries to focus more of their funding and attention on.

Please refer to backer analysis sheet for questions below:

Use your data to determine whether the mean or the median better summarizes the data.

The median better summarizes the data because for both the unsuccessful and successful backers data we can see that the median is much lower than the mean. The median is a better measure of central tendency in this case because there are influences of extreme values dispersing the data. The presence of outliers has more effect on the mean than the median, which is why it’s better to use the median in this case.

Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

The data tells us there is more variability with successful campaigns being that the variance (1,606,216.59) and the standard deviation (1,267.36) are very high. What this means is that successful campaigns have a variety of factors that will contribute to its success factor, you can not make an assumption that a successful campaign's contributing factor is the number of backers it has.

When it comes to the unsuccessful campaign data set, it has a lower variability. Meaning that we can have more confidence that without enough backers it is more likely that a campaign will be unsuccessful.

These explanations make sense logically.