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

Crowfunding campaigns under the category of theater have a greater response than other categories as the most failed projects and most successful projects are within that category. This trend continues as we break down categories by sub categories. I can see that the greatest response is under Plays with having the biggest failed and successful projects. Crowfunding campaigns success/fail outcome can also be correlated with the time the campaign is launched. The most successful campaigns/projects were usually launched during July while the most failed projects happened during January and cancelled projects were during August, this gives us an idea of when it would be the best timing to launch a campaign.

What are some limitations of this dataset?

Some limitations of this dataset are that we might not see the full picture when it comes to campaigns and there might be other factors that affect the outcome aside from the one’s we are provided. The dataset is also providing 1000 samples which might not be representative of the whole population.

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

Another table/graph we could have created that would have added value to our understanding of crowdfunding would have been between outcome and backers counts. This would have helped us understand how the outcome is affected by how many backers the project has and this would give us a margin of at least how many backers a project should have in order to be successful. The most appropriate way to display this would be with a line graph were we can see the correlation between outcome and number of backers.

**Statistical Analysis**

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

The mean better summarizes the data as it gives us a clear average of how many backers, we will at least need to be successful. In the case of an unsuccessful/failed campaigns we can also see on average how many backers were on that campaign that made it unsuccessful. Overall, an average in this case gives a better picture of at least how many backers you should have to be successful/unsuccessful.

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

There is more variability with successful campaigns than with unsuccessful campaigns. Yes it makes senses as the larger the standard deviation, the more variable the data set is. This is also consistent with the variance as successful campaigns are reflecting a greater degree o spread than unsuccessful campaigns. Successful campaigns have more spread than unsuccessful campaigns, thus it has a larger variance in relation to the mean.