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

Given the provided data, there are a few things we can gather by looking at the data charts and tables we have made. One is that there are more projects trying to get funding through crowdfunding platforms than any other project category. Theater groups represent the largest number of successful crowdfunding projects within this data set. Second, more projects were starting their crowdfunding journeys towards the beginning of the year. January through July saw a higher number of projects asking for funding. As for the later months, September through December, saw a decreasing number of projects asking for funding. Lastly, film & video and music parent categories included the greatest number of sub-categories. This can draw a conclusion that those categories have the widest range of projects started compared to, say food or theater, which only had one sub-category and all fell within the same kind of project.

* What are some limitations of this dataset?

Some limitations of the dataset include the inability to see if there was a large donation that helped kickstart the project. A large sum of money could drastically affect whether the project was able to be successful or failed. Another limitation is the only contributing factor we can see from this data is how many people have given money. We do not see how many people came together to start the project beyond just those who contributed money, which could also be a huge part of whether the project failed or not.

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

Another helpful graph that we could create with this data would be charting the percent funded verses the outcome. This type of graph would show the correlation between hitting the goal with the amount of money pledged and if the project was successful.

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

I would conclude that the median better summarizes the data, because some of the projects were backed by a very large number which skewed the data.

* 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 because the standard deviation is larger, and there are far more successful projects in this data set. This does make sense because there is a wider range of backer counts in the successful category than in the failed category.