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

One conclusion that can be drawn about Kickstarter campaigns is that campaigns having to do with theater will have the most success, as well as the most failures.

Another conclusion that can be drawn is that within the sub-categories, Kickstarter campaigns about plays are leagues above the others, in both terms of success and failures.

Lastly, another conclusion that can be drawn when looking at the dates of these Kickstarter campaigns is that there is a drop in successful campaigns and an increase in failures when December arrives. This can be due to the amount of shopping that people have to do for the holidays, and the less likely they are to contribute.

1. What are some limitations of this dataset?

A limitation that can be seen from this data is that despite what may look like a correlation between the amount of success a campaign has and the category it is in, doesn’t always mean that you’ll have success. This is evident in the theater category that has a lot of success, but also a lot of failures. So thinking that starting a campaign in a category of theater will be successful is not exactly true.

Another limitation is that despite seeing an increase in success and a decrease in failures during the month of March to April, that doesn’t mean that you’ll have the most success during that time.

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

A possible graph I would like to see would be one that shows the average of money donated in comparison to the categories. Would be interesting to see just how much people are willing to donate in each category.

Also, would like to see a graph that compares the amount of time a campaign is live to its success. Does the amount of time being live matter or is success just dependent on the category?

Bonus:

Mean is better at summarizing the data, as the median only considers the amount of numbers present and not the total. While mean is the average of all the numbers, thus gives us a better understanding of the data.

There was more variance with the successful campaigns as they had a larger maximum when compared to that of the unsuccessful ones. The difference was about 26,300 which is very huge, and thus, had a larger variance. This makes sense, as the successful campaigns had a larger range when compared to the unsuccessful ones.