Conclusions from the Data

1. Out of all the campaigns, theater and plays had the most campaigns submitted. They accounted for 34% of the data, which is significant. While this is the majority, it skews the data heavily. If we were to remove it from the data set, we would see a more normalized set to work with.

2. The projects that fall between $10,000 and $35,000 are the projects that are canceled the least. This could stem from them being larger scale projects for organizations. Projects over $50,000 were cancelled the most. Again, stemming from the cost. Either enough money could not be raised, or the cost ended up being too high.

3. The successful and unsuccessful campaigns have a spread of backers, leading to very high variances and standard deviations. The data is not normalized, so there is major right skew on the data with larger backers on campaigns. Normalizing this data set will give us a better understanding the spread and confidence levels.

Limitations and Suggestions

1. One suggestion I have to generate a percentage successful/failed/cancelled per Parent Category and Sub-Category to get a better idea of the overall success of certain campaign types. This could show those who are looking for backing what are some of the most popular things that can be funded by their audience.

2. One limitation that is not considered with the data set is the company and their own contribution to their project. Is the funding any of this on their own funding? Are these small businesses or startups? Before we can make a true and unbiased assessment, we would need to know something like that.

3. My final suggestion would be to gather more information based on the amount pledged by category to see where the pledges are going to, and to determine what areas may need more of a push for funding. Even an average per country