**Crowdfunding Goal Analysis**

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

1. Parent Category – Journalism has a 100% success rate; however, due to only 4 data points, you are unable to draw any meaningful conclusion from this set of data.
2. Sub-Category – The sub-category “Plays” makes up nearly 1/3 of the total dataset.
3. Monthly Compare - ~57% of campaigns have been successful out of the total of 986 completed campaigns.

**What are some limitations of this dataset?**

The definition of “Pass” vs. “Fail” does not really draw a clear picture of what went right or wrong within each project. I do not think we really have much of an understanding of what the dataset we are comparing and would like to know how Goals were determined, time spent marketing and what went into finding backers. Other columns such as “staff\_pick” or “spotlight” do not provide much context. I think some sort of key incorporated with data would make for a much stronger deal of analysis.

The raw dataset was limited by the formatting originally presented. Dates were in an unreadable form and had to recalculate these dates so they could be properly viewed. Same goes for Parent & Subcategories. Instead of separated columns, we had to create a formula to separate these 2 items for comparison.

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

I would like to see a pie chart added to the visualization. If comparing different categories & sub-categories, it would be meaningful to see the data represented in percentages. The range of datasets in each of these parent/subcategories varies greatly. Without viewing the data in Totals & Percentages, I find it difficult to draw any conclusions or understand if one category is typically more successful than others. Additionally, We only analyzed the data on a “Pass” or “Fail basis if it hit 100%. There are several different projects that were well over 1000% and seems like we are missing the full story comparing a project at 1000% funded compared to 100% as the same.

I would like to see a breakdown on Average Donation has an impact on success of a project. My hypothesis would be that the higher the average donation, the more likely the crowdfunding is to succeed. This would be one of the first things I would look after reading the columns within the dataset to look for trends and useful information to help future crowdsourcing campaigns.

I would like to see a breakdown by “Year” to see if there is any correlation with success. Like monthly compare line graph, I am questioning if year has any impact on success. I would think many factors would go into this but economic state I would think is a key driving force of year-by-year success for funding campaigns. As a hypothesis, in a strong economic year, I would expect successful projects to be greater than during a poor economy.

**Statistical Analysis**

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

Median would be the better metric to use. Outliers are increasing the mean and skewing the data to think that the average is much higher. If you removed these outliers, then the average would be way closer to the median of the dataset.

**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 greater variability in the successful campaigns. Both the variance & standard deviation are greater for successful campaigns over failed campaigns.