Hiromi Miyata

2/9/2023

Assignment 1

**Crowdfunding Goal Analysis**

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

The three most popular crowdfunding categories are theater, music, and film & video. The most popular subcategory for crowdfunding is plays. There is a spike in success overall for the campaigns during July.

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

This dataset does not include the percentage success based on category. This data will be helpful because it can help investors predict which categories have a higher success rate and lead to profit. The average donation rate may be also inaccurate because not all profit may be returned to the investor.

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

It will be useful to make a graph/table that correlates the average donation rate with the category so we can pinpoint which category is likely to create a higher return. To do this, we will need to average the donation rate based on category and then create a pivot table that corresponds to each category. It would be also useful to have a graph of the duration of the campaign (Date Ended Conversion – Date Created Conversion) corresponding to the outcome of the campaign. This will help in seeing if the length of the campaigns may have correlation with the outcome of the event.

**Statistical Analysis**

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

The standard deviation is very high for both successful and failed campaigns. This indicates that the data set differs a lot from the mean. The variance also seems to be very high. These indicators show that the mean can be skewed a lot by outliers. Therefore, the median is better to summarize the data than the mean.

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

According to the data, the successful campaigns have higher variance and variability. This makes sense because there more values in the successful campaigns that have a large difference from the mean. The data shows that there is a trend of having higher backer counts for the successful campaigns. However the high variance may indicate that the successful campaign’s backer count may be more catered to each campaign than failed campaigns.