Samantha Rose Pabalate

Challenge 1 Crowdfunding Questions

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

Based on the data provided, we can observe that the parent category “Theater” seemed to have received the most successful campaigns, but also received the most failed campaigns. We can conclude that the theater category, overall, holds the largest stake within the nine parent categories.

Based on the sub-category chart, we can see that the only sub-category to theater is plays. Because the theater category is less diversified, all the donations are allocated to one place. One could conclude that a less diverse portfolio of projects could be more enticing to contributors, as it allows them to know exactly where their donations are going.

Conversely, parent categories, such as film/video and music, each have 6 sub-categories making the pool of ideal contributors more niche. One could also conclude that having a smaller pool of potential contributors may result in less likelihood of a successful campaign.

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

This dataset does not include the dollar amount of each individual contributor or backer. It only provides the average donation. However, the average donation only reflects central tendency and doesn’t acknowledge any individual outlying donations.

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

Another table and/or graph that would be valuable to analyze could be the outcomes by the duration of each campaign. The length of time put into each campaign could also play a part in the campaign’s success rate. This could help plan for an ideal duration for future campaigns.

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

When comparing the number of successful and failed backers, I believe the median better summarizes the data. The two sample sizes vary greatly in size with 365 failed campaigns and 566 successful campaigns. The number of successful backers consist of several significant outliers which greatly skews its mean.

**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 much more variability with successful campaigns than unsuccessful campaigns. This makes sense because the sample sizes vary greatly.