Ryan Dekker

Module 1

9/21/23

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

The first conclusion drawn is obvious – most of these campaigns are for artistic productions. Theater is first at 344, and then film & video (178), followed close behind by music (175). Taking a bigger step back, a majority (56.5%) of these campaigns are successful, which importantly means that donations to these campaigns are not a pure money sink. Still, this sample shows that while most are successful, a notable contingent either fails (36.4%) or is simply canceled (5.7%). Interestingly, trends show that summer is a popular time for fundraising, as July is the most common month where campaign ends, and July has the highest rate of campaigns being successful (62.4%). Because most of these campaigns for the arts, it is not too surprising many campaigns end during a lull in the academic calendar.

*What are some limitations of this dataset?*

This sample is only for 1000 campaigns, a fraction of the population. The plurality of these campaigns analyzing theater ventures may not be reflective of all Go Fund Me campaigns. The dataset is also minimal for describing the actual proposed project, so finding trends for what makes a campaigns successful is not likely to be found from the present dataset.

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

Examining trends among ‘Staff picks’ from this organization, as well as the “spotlight” feature from the organization could be explored for whether those designations are meaningful predictors of successful campaigns. Finally, a comparison for central tendencies for length of campaign could be calculated for successful and unsuccessful ventures could be calculated, which could lead to whether length of campaign is another factor in making the proposed goal.

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

For both successful and failed campaigns, most have only a few hundred backers. However, both types had outlier instances where thousands of backers existed. The wide skew of different backers but vast majority being small indicates that the median amount is more reflective of reality. Additionally, when determining differences between successful and unsuccessful campaigns, the mean number of backers is not too different. The mean number of backers for a successful (851) campaign is only 45% more than the number of backers for a failed (586) one. However, the median number of backers for successful (201) campaigns is substantially larger than the failed (114.5) campaigns, being 75.6% larger. Failed campaigns can still fail, even with outlier cases of large number of donors. Successful cases tend to have a much large amount of median donors.

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

Successful campaigns have a higher variance score, higher standard deviation, and a larger difference between maximum and minimum number of backers. Successful campaigns show more variability than unsuccessful ones. This isn't too surprising, as some successful campaigns need large amounts of backers to reach their goal and some only need a relatively smaller amount of backers who each gave high amounts of cash. There are varying ways to have a successful crowdsourcing campaign and the data reflect that. Successful campaigns could have had a 'bandwagon appeal' where they were close to reaching the goal and gotten more backers at the last minute, where unsuccessful campaigns didn't get close enough to have those extra backers at the last minute. Overall, many unsuccessful campaigns likely had trouble getting traction and had trouble getting large number of backers. Many failed campaigns have little to no backing, which in its own way is consistent.