Kimberly Childers

Responses to Module 1 Challenge

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

Three conclusions we can draw about the crowdfunding campaign are:

1. Projects that fell under the journalism were all funded.
2. Projects dealing with technology as the parent category have a 60% chance of being funded. That was the highest average of being funded out of the other categories.
3. Most proposals submitted for crowdfunding relate to theater.
4. Applying through Crowdfunding gives a project a little over 50% chance of being funded.
5. What are some limitations of this dataset?

This dataset could include more data points to show trends based on race and gender to see if the majority of Crowdfunding applicants are of a certain demographic. If that is the case, then the data will be skewed to represent only that portion of the population. The categories could have been broken out a little more to cover a wider spread of topics. Most of the categories centered around liberal arts.

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

I ended up making a table to show how many projects were successful compared to failed ones to get an idea of the odds of a project being funded. This could be used to convince applicants that their opportunity to get funded through Crowdfunding is likely. I would also like to see if there is a relationship on whether the factor of an established company or an individual shows a higher rate of getting funded. Another viewpoint I want to check out is reviewing the average time it takes to complete the project. It may show a trend if shorter projects are more likely to be successful compared to long term ones.

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

Looking at a graph of the data points, it looks like the median better summarizes the data for the successful ones and the mean for the failed ones. The successful data points have a smaller standard deviation meaning the datapoints are less dispersed in relation to the mean. Hence, it makes me think that the mean will be within the range of the median to summarize the data okay. As for the failed datapoints, there is a larger standard deviation, so I think the median better represents the datapoints.

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

When I reviewed that datapoints in a visual depiction below and threw in the deviation, you can see that there seems to be more variability in the successful datapoints compared to the failed ones. This does make some sense to me when comparing the charts below because the successful datapoints show a greater range of points compared the failed ones. The failed datapoints look like they all fall within a more concise range.