Module 1 Challenge

Report

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

One conclusion we can pull is crowdfunding campaigns success is flipping a coin. The chances of success can vary greatly and can depend on many factors from popularity of campaign, participation interest, and even popularity of company initiating the campaign. The data set used, for an example, shows that roughly 64% of campaigns were successful in achieving their financial goal, which begs the question if the effort of starting a campaign is worth it. Another conclusion we can make about crowdfunding campaigns is that success does not translate directly to the number of backers. The difference in maximum number of backers from successful and failed (7295 successful and 6080 failed) campaigns is only 1215 backers, which goes to show that a successful campaign comes is only reliant on the amount given by backers. More backers does not necessarily translate to more successful campaigns. The last conclusion that can be gathered about crowdfunding is there could be an ideal goal range. According to the line chart I created that compares percentage of outcomes (percent successful, percent failed, and percent canceled are represented) to goal set, there is a roughly positive 60% increase in successful campaigns around the $1000 to $4999 range goal, then another positive 100% difference in the $20,000 to $24,999 range goal. Between those ranges is a negative 10% in the $10,000 to $14,999 range, then the gap grows closer together after another positive 100% difference at the $30,000 to $34,999 range. This showing that, at least in this survey, that there are more successful goal ranges than others.

What are some limitations of this dataset?

Some limitations of this dataset are the choices of categories and countries available. The kind of categories covered can paint a good picture but if more entertainment categories were presented then there could be change in the overall data. Another limitation is where the data comes from. The majority of the data comes from the United States and this datasheet could be skewed because of that amount of data.

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

We could also make a table that compared the date in which the campaign took place to the category. This could show if there is any relation result based on the time of year and season. There could be more backers that support backing a campaign focusing on music in the summer for example.

Statistical Analysis

I believe that the data would be better summarized by the mean of the because of the similarity of the successful and failed number of backers. In my opinion the median, compared to the maximum amount backers, does a good job of showing the range of amount backers. This showing that the amount of backers was not the only reason for success and that there could have been other factors at play, but because the median is so similar to the median of the failed campaigns it overshadowed this idea and shows more evidence that the mean is the better way to summarize the data.

According to my data there is more variance in the number of successful backers than failed backers. This does make to sense to me because of the comparison of the median to the maximum number. The median is 200 with the maximum being 7295 and shows that there was a large rise in backers as the amount of backers grew, almost exponentially.