

Module 1 Challenge Reflections

Crowdfunding Goal Analysis

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

From the given data, it can be concluded that crowdfunding campaigns are generally an effective way to raise money; being successful in passing set goals roughly more than half of the time in each category. The three most popular categories, film & video, music, and theater indicate that projects that have the potential to draw crowds will have a wider pool of public interest and willing donors. This could mean that these categories are the safest projects to launch in a crowd-funded campaign. Although the previously mentioned three categories have the higher number of successful campaigns and total donors, photography, journalism, publishing, and technology have a higher percentage of successful campaigns relative to total campaigns launched in those categories. This could be a sign that these categories have more untapped potential than the three leading categories.

What are some limitations of this dataset?

The outsized impact that theater/plays hints at a limitation of the dataset. Just because theater/plays find crowdfunding to be the most convenient or effective way to raise funds for that specific type of project, does not necessarily mean that theater/plays are the key to success in obtaining funding for projects via crowdfunding. The dataset is also limited in that the goals set by different projects can differ greatly. When data is sorted by the goal amount, it shows that many projects, although able to pledge a great deal of money, are counted as failures due to not meeting ambitious goals, and vice versa. There are too many variables to consider; and simply labeling a project a success/failure based on their own goal amount is not a strong indicator of any special “trick” to crowdfunding.

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

The dataset could benefit in several areas. First, there are several projects that seem to have been inactive which should be excluded from the data. This would clean up the data and make it more representative of projects that were earnestly pursued. The dataset could also benefit from lumping similar projects together. For example, there are several musical performances that could be consolidated into one category. This would reduce the number of categories to look at, and when combined, would likely produce a category more comparable to plays/theaters so it would look like less of an outlier category. More attention should also be paid to the number of backers in relation to the goals/pledges being set. This could give the client a better understanding of what projects bring in what amount of backer interest, and what kind of projects bring in more support per backer.

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

Because the standard deviation and variance is so high in both the successful and failed categories, the median would better summarize the data since the median is less impacted by extreme values in the dataset.

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 more variability in successful campaigns. This makes sense as there are more successful campaigns than failed campaigns to get data from. It also makes sense because successful campaigns tend to skew towards having more backers on average due to the nature of a campaign being successful, as opposed to a failed campaign which would have failed due to not having enough backers.