Andrew Lane

Data Analytics Boot Camp

Module 1 Challenge

Crowdfunding Campaigns

Three Conclusions:

The bulk of crowdfunding campaigns were for Film & Video, Music, and Theater projects; with the most (34.4%) of campaigns for Theater. Out of the 1000 campaigns in this data set, 697 were in these categories.

The success rate for crowdfunded campaigns was 56.5% overall. While Journalism has a success rate of 100%, there were only 4 projects in this data set. Photography, Food, and Games campaigns performed below a success rate of 50%, representing higher risk campaigns.

Crowdfunding campaigns that begin in the months of June and July are slightly more likely to succeed than campaigns begun in other months. Looking at the chart that plots success and failure numbers by month, the greatest difference in the numbers is in these two months.

Limitations of the Dataset:

The dataset is limited in the parent categories and subcategories used. With the bulk of the campaigns in Theater, some subcategories would break the data down into more useable pieces. I also assume that this is a generated dataset due to the uniformity of the data. I would expect more variance in the data if it were actual. (The greatest variance seems to be in the percent funded, or pledged/goal, and the number of backers.)

Other Useful Tables and Graphs:

Looking at success rates in different categories would be helpful. I added this calculation to one of my pivot tables and found the rates a useful metric. This would help funders and fundraisers understand the risk of certain types of campaigns.

I’d also like to see the outcomes listed by country I added this pivot table and chart to a worksheet. The data shows that in this dataset, most campaigns were in the United States. This makes me wonder: 1) How do other countries fund these activities? and 2) Why is the US more active with crowdfunding than other countries?

I also did an analysis on the success of crowdfunding while looking at the amount of the campaign’s fundraising goal. I found that the most successful campaigns had a goal under $10,000. I addition, the data showed that the highest rate of success was with campaigns with a goal under $5,000.

Mean vs Median:

The median is a far better summarizer of the data due to the number of backers in both successful and failed campaigns being heavily skewed right. Evenly distributed data, represented with symmetry and without outliers, can be summarized with the mean. This data heavily skews right, and the median is the better choice. For example: The number of backers of successful campaigns has a mean of 851.15 and a median of 201. This 423% difference is a significant skew.

Variability and Success:

The range, variance and standard deviation of successful campaigns are significantly higher than failed campaigns. All these indicators point to a higher level of variability with successful campaigns. This higher level of variability is probably due to the campaigns being tailored to their region and targeted backer groups. I would assume that unsuccessful campaigns used a more general approach to their crowdfunding and while this is more uniform, this less varied approach was less effective at meeting the fundraising goals.