Lyndsay Wylie

September 16, 2022

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

Given the provided data, we can draw the following conclusions:

Theater is the most popular type of campaign; Theater also has the largest number of successful campaigns but also the largest number of failed campaigns.

Journalism is the least popular type of campaign; Journalism may be the least popular type of campaign, but it has the highest percentage of successful out of all campaigns, and no failed or canceled campaigns.

Another conclusion we can draw is that campaigns see more rates of success in the months of June and July. Failed campaigns are in a range from January to August, but campaigns see less failure in the fall, while September has the least failed campaigns. Canceled campaigns stay within a range throughout the year, seeing between 0 and 10 each month.

1. What are some limitations of this dataset?

One limitation would be percentages in comparing the different types of campaigns and outcomes against one another, not just individually. I would also like to, in addition to months, be able to compare the data by outcome and year, and year plus the type of campaigns. I’d also like to have datasets on the strength of the economy to see if there is a correlation with the percentage of successful campaigns.

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

A graph that I would be useful for the limitations I outlined above would be a pie graph. It would allow the data to be compared in terms of a whole instead of individually, and having to run the percentages of outcomes.

Additionally, being able to compare in a line graph data by year and outcome, and type of campaigns and campaigns it would allow me to draw conclusions about a particular year or years having more success or failures, and there is a trend to type of campaigns being more popular in specific years. These two sets of data can also be displayed in a scatterplot with a linear trendline.

Bonus

In both instances, the data is skewed with outliers; That being said, the data can be summarized more meaningfully if we use the median.

There is more variability with the successful campaigns. It makes sense given the greater number of backers and higher standard deviation.