Crowdfunding Analysis-Module 1 Assignment Summary

Many companies run crowdfunding campaigns to secure funds to start a new business or get additional money to grow their company. The goal of this project is to see factors that affect the success or failure of such crowdfunding campaigns and see if the data is representative or not. Detail analysis is done in the excel file.  Below are some points to answer specific questions asked by the instructor. See answers under each question asked to address.

Create a report in Microsoft Word, and answer the following questions:

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

Among the main categories, theater companies run the largest number of campaigns and proportionally they have the largest number of successful and failed campaigns followed by music and, film and video. On the other hand, journalism runs the least  number of campaigns followed by photography and food companies. In general, most of the campaigns are either successful or failed and there are few campaigns that are live or canceled.

Among the sub categories, the most successful and failed campaigns are run by play companies compared to other subcategories. From the date conversion data, successful and failed campaigns showed high variability based on the month when the campaign was ended, and campaigns that ended in December have the most success.

Looking at the data percent success against goal, campaigns with small or very large amounts of goal have more failure and less success compared to companies with moderate campaign goals. The most successful campaigns with less failure are those with goals of 15000 to 35000.

* What are some limitations of this dataset?
* One of the limitations of the data is it didn’t give the background of the backers if this has influenced the outcome of the campaigns. So, it is difficult to dissect the data and examine the reasons for the variability and success or failure of the campaign goals. Moreover, the number of backers for each campaign is very different and difficult to measure the outcome of the campaigns of the different industries with different sample sizes.

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

Probably using a histogram chart could help more to understand the distribution of the dataset and see how much the data is skewed and determine what measures of central tendency to use. Moreover, calculating inter quartile ranges can help to identify outliers and exclude them from the analysis.

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

Based on the backers count summary statistics, the median looks better measure (indicator) of the data than the mean. Because the mean number of successful and failed companies is very high compared to their medians indicating that the data is skewed (not normally distributed) and not representative of a good data set (normal population).

* 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 among successful campaigns than field ones. Because the difference between the median and average number of backers is very high in successful campaigns compared to failed ones.  This is also revealed by the higher variance in the successful campaigns than failed ones.