UCI Data Analysis BootCamp

Module # 1 Challenge

Data Report

Submitted by: Ronald Felix Alfelor

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

Conclusion # 1: Out of all the campaigns being funded, most funders prefer to fund the theater category, particularly the plays. Based on the Campaign per category table, a large portion of the crowd funders go to the theater category. Likewise on the Campaign per Sub-category table, the dominating sub-category is the plays.

Conclusion # 2: Majority of crowd funders donate the amount starting from 1000 to 9999 as well as the greater than 50000 range. But the goal ranges of 15000 to 19999, 20000 to 24999 and 30000 to 34999 all have 100% success rates.

Conclusion # 3: Crowd funders usually successfully donate during the months of June and July. Based on the Campaign per Month Table, the months of June and July have the highest number of success as well as a higher percentage of success.

*What are some of the limitations in the Dataset?*

There were a few missing data that limited the extraction of conclusions from the data set. One would be the company's nature. It was obvious that they were technology-based companies, but it would be better if we categorized the companies to determine whether a certain company type would tend to choose a specific category for reference in the future.

Another one would be the absence of reason why the outcome was successful, failed, or canceled. Being able to determine the reason of the outcome would give us a guide in the future of any strategies on what to do or not to do that could improve the next crowdfunding process.

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

One table that we could create is a pivot table and/or chart on the location where the company was based and the category they selected. This could determine the tendencies of the companies in selecting a certain category and sub-category depending on the location they are based.

It would be a great help if they provided additional values such as the size of the company. Categorizing all companies as small, medium, and large companies depending on income, would be a great help in determining what category of campaign that they tend to fund.

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

The data shows a skewed distribution where both successful and unsuccessful campaign’s means and medians are unequal. It is possible that either the data is not properly filtered, or the data is not evenly distributed.

*Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?*

The last sheet shows that the data has a lot of variability with both the successful and unsuccessful campaigns. This means that the outcomes are unevenly dispersed around the mean. With this data, it can be concluded that it is difficult to make any predictions.