**Crowdfunding Challenge**

**By: Sophie Effy Vazquez**

The three conclusions we can draw about crowdfunding campaigns are as follows:

1. There are more “Theater” crowdfunding campaigns than the other crowdfunding categories. The “Theater’s” sub-category, “Plays,” has three-hundred and forty-four (344) crowdfunding campaigns out of the one thousand (1,000) crowdfunding campaigns. The remaining crowdfunding categories had less than one hundred (100) campaigns with the majority having less than fifty (50) campaigns.
2. The “Journalism” category has the least crowdfunding campaigns with four (4) projects, however, all were successful.
3. There are more successful crowdfunding campaigns than the failed, canceled, and live campaigns.

The crowdfunding campaigns should be analyzed further to get more insight since there are some limitations to this dataset which are not solely exclusive to the limitations listed below:

1. The data is limited to the select categories provided and not all possible categories available. Using this dataset may not yield accurate results to determine what the more successful crowdfunding campaign categories are to possibly look into them further.
2. The data is also limited because the source of the crowdfunding was not provided (i.e. Kickstarter, Indiegogo, GoFundme, etc.). The source can be useful to see if there is a common crowdfunding platform to use for the campaign or multiple ones, amongst other variables.

Other possible tables and/or graphs that we would create to analyze the data are:

1. Pivot table to show the correlation between the duration of the crowdfunding campaigns and the success or failure of those campaigns.
2. Scatter plot graphs to visualize the findings, one for the successes and one for the failures. The value these would provide are more insight of whether or not the duration of the campaign is a major factor to a particular category succeeding or failing and if either is something that should be analyzed further.

Statistical Analysis:

An analysis was made on the number of campaign backers the successful and unsuccessful campaigns had. The median better summarized the data as opposed to the mean because of the variability of the dataset in both the successful and unsuccessful campaigns. There is more variability in the backers of successful campaigns than of the backers of the unsuccessful campaigns. Having more variability in the backers of successful campaigns versus the unsuccessful campaigns makes sense because there are simply more backers to count in the successful campaigns than the unsuccessful campaigns.