**Assignment 1 – Crowdfunding**

**Crowdfunding Report**

What are three conclusions that we can draw about crowdfunding campaigns?

1. Theatre generates the largest number of successful campaigns relative to unsuccessful campaigns. Furthermore, the subcategory ‘plays’ generate the largest number of successful campaigns compared to all other subcategories.
2. As the number of successful campaigns increase, so does the number of failed campaigns. Goals between 15,000 and 35,000, there are minimal to no failed campaigns, although the number of total projects is much smaller in comparison to other goals. There may be a threshold, that goals over 50,000 have more failed campaigns compared to successful campaigns.
3. The months June and July generate the largest number of successful campaigns relative to unsuccessful campaigns.

What are some limitations of this dataset?

1. The assessment of average donations is not all in the same currency. Therefore, it may appear that one country donated more than another, however, if the currency is not all the same, it is difficult to draw comparisons of this. Furthermore, in keeping with the currency, this would not take into consideration the country’s economic status. A country like the United States may have a higher economic status in comparison to Italy, where more money could be used for donations.
2. It does not consider potential campaign restrictions within countries. Some countries may be more heavily regulated, which may influence whether a particular campaign was successful or unsuccessful.
3. The dataset is only taking into consideration one point in time instead of a comparison over time (i.e., 6 months, 1 year). Therefore, a campaign that may have been initially successful, overtime may not be.

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

1. Box and whisker plot: As the data is right skewed, this graph would be able to better visualize the quantitative values of the data. It provides a visual in order to easily see where most of the data points land at a glance.
2. A table in which all the currencies can be converted into a specified currency. This would better be able to accurately depict how much money was spent on campaigns (i.e., average donations).

**Statistical Analysis**

In analyzing the data, both numerically and graphically, the median better summarizes the data. In this dataset, the data would be considered right skewed or has a positively skewed distribution, as the mean falls to the right of the median. The median better describes the data as it is not as influenced as the mean by the extreme values in the data set. In comparing variability between successful and unsuccessful campaigns, it makes more sense that successful campaigns have wider variability. This is likely influenced by not only the number of total campaigns, the wider range from minimum to maximum values, the variance and standard deviation of the data set.