**Crowdfunding Questions**

* **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?**

Three conclusions I can draw are the following: Journalism had the fewest campaigns (4) while theater had the greatest number of campaigns (334). The campaigns most likely to succeed were “journalism” at a 100% success rate. The campaigns most likely to fail were “games” at a 47.9% failure rate.

* + **What are some limitations of this dataset?**

Some limitations of this dataset are the following: There is not an equal distribution of campaigns in this dataset with journalism only having 4 campaigns at the low end, while theater makes up 344 campaigns at the high end.

The dataset is heavy weighted to campaigns from the *United States making 76% of the dataset* (763 of 1000), leaving the 6 countries combined making up the last 24% of the dataset.

The donation amounts are measured in their country of origin’s currency, and there are 7 different countries in this dataset.

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

Other useful visualizations that I could create to add value are the following: A clustered column chart comparing “percentage funded” and “Outcome” by category. A line graph comparing campaign outcomes with campaigning time (deadline date - launch date).

**Statistical Analysis**

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

The data distributions for both the successful and failed projects are greatly left-tailed skewed, the median being several times that of its respective outcome project group. Given these extremes, going by the median is the better representation of the data.

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

Looking at my data, it would seem the successful campaigns have both a higher variance and standard deviation of backers compared to unsuccessful campaigns. This does not make sense, as I would expect the opposite to be true, but this is only one of many factors at play. Not all backers contribute funds equally—a few billionaires vs a thousand average Joe’s. A better measurement would be the variability of funding amount—follow the money.