**Summation of Results**

After a preliminary analysis of the data provided, some conclusions are evident:

* While campaigns fail in every category, successes tend to outnumber failures in both primary and secondary categories.
* Performing arts (film & video and theater & plays) have a large presence in the crowdfunding data but also tend to fail at a higher rate than the other categories, whereas journalism has a low presence in the data (four data points) but a 100% success rate.
* Crowdfunding goals are either attained or forfeited in a small period of time, usually in a few days and seldom more than a month.

Some limitations of this sample could be:

* Selection bias – Could the data have been selected with an emphasis on completed or successful campaigns? While there is a category for canceled campaigns, we don’t always know if the collection of this data was focused on live, canceled, failed, successful, or unbiased, mixed data (as would be optimal).
* Representation – Is the sample representative across campaigns, or was the collection biased towards one of the categories? Comparing rates of success or failure would have little value if you were planning to raise capital in an underrepresented industry in the data, since the rates for your industry could be far different than those of the survey.
* Missing data – Could missing data contribute to biases in the data? If a survey or form was half completed, how would the data be represented and would it change our results?
* There is no listing for the amount of time that projects existed prior to this crowdfunding effort, including but not limited to attempted rounds of funding and financial successes or failures.
* Outliers could be present in the donation data due to one or two large investors, particularly donations that would create a successful outcome that would not have been possible otherwise.

Some additional analysis that would bring value to our view of the information:

* Charts detailing statistical analysis, such as plotting the mean, median, standard deviation, and z-score.
* Amount of time a project existed before the crowdfunding effort.
* Number of prior projects attempted by the entrepreneurs listed in the “name” category and the outcome of those projects.

**Statistical Conclusions**

Whether the median or the mean is optimal for our analysis depends on our use case for the study. In a normal distribution the mean and the median should be close in value, but in our study in both successful and unsuccessful categories, the median is much less than the mean. (Additionally, maximum values for successful and failed are 7295 and 6080, respectively, demonstrating that there are extreme values in the set.) For scientific reporting purposes then, the median should be used because it gives us a better idea of the interest in a typical crowdfunding project. However, if our analysis is based on choosing whether or not to use crowdfunding for our own project, outliers, and therefore the mean, have a lot more value, since these raw numbers represent discrete investors who can be expected, in general, to act to maximize their own return, and could easily allocate their capital to other projects if they saw value in them. If we have a project that shares similar characteristics with the most successful funding projects, then there is no reason not to expect higher investor interest.

Based on a higher standard deviation among the successful campaigns (1266 vs. 960), variability among successful campaigns appears to be higher. One would expect to see a higher variability among projects that will ultimately be completely funded than among ones that will not reach the funding hurdle. This makes sense. Since successful projects, and therefore perceived good investments, tend to attract more attention as well as capital, we would expect to see higher variability among them.