1.1 Challenge

Crowdfunding Analysis

# Part I Q&A

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  1. Although there is no direct correlation between campaign goal and success rate it seems that there is a lower limit. There were zero successful campaigns with goals of less than $600. It’s possible that such small amounts of money are not appropriate for a crowdfunding platform and should look for a small loan from a friend or family.
  2. Performing arts are the dominant presence in the data set. This may indicate that people are more willing to invest in a crowdfunding project that is entertainment based. Folks are giving money to things they are enthusiastic about.
  3. The total number of campaigns drops off in the 4th quarter of the year. This may be due to the financial strain of the Holidays. Most people may not have a lot of extra money for passion projects.
* What are some limitations of this dataset?
  1. We are not given any descriptions of the play or film themes. This may help us dig further into what people like in our most popular category of campaigns.
  2. We do not know the amount of individual donations. It is possible that some campaigns have been primarily funded by a handful of very large donors that are skewing the average donation amount numbers.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. We could create a table showing the total amounts raised for each category. Most crowdfunding sites make money by taking a small percentage of the donations so this information would be very useful.
  2. We could make a chart displaying the average number of backers per category to show which categories have the most visibility. This is an important predictor for success.
  3. We could make a bar chart showing success vs. failure by country. This would tell us if there was any significant difference in the success rate based on the geography of the campaign.

# Part II Statistical Analysis

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

The Median better summarizes the data set due to the amount of campaigns with multiple thousands of backers. Although these are not the majority of the data set, there are enough of them to make the mean irrelevant.

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

There is more variability with successful campaigns which is to be expected. Successful campaigns will generally have a high number of supporters. This lends itself to higher variability. There are also simply a much larger number of successful campaigns. The sheer amount of examples gives more chances for variability.