**Weekly Challenge #1 Summary**

**Crowdfunding Analysis**

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

Answer to query:

* First observation – the number of crowdfunding campaigns are predominantly in film/video, music, and theater
* Second observation – the ratio of failed to total projects rarely rises above 50%.
* Third observation – the best success rate in relation to sought backers appears to be in the range of 15k to 35k
* Fourth observation – there seems to be a consistent downward trend from roughly July until September in the project tally
* Fifth observation – plays are predominately attempted over any other subcategory

1. What are some limitations of this dataset?

Answer to query:

* We don’t have any information on the success of the crowdfunding campaigns after launch (e.g., profitability two years or five years after launch). There could be some early indicators which could help us predict success. Which our clients would be very interested in.
* We could add some economic “health” data to see if we can corroborate this with our “prime start time” for crowdfunding campaigns (listed below).

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

Answer to query:

* We could add a graph that analyzes the amount raised in proportion to number of backers in various countries to see if there is an advantage to one country over another.
* We could look at the date started and see if some years were better to start a crowdfunding campaign than other and see if a pattern emerges.

**Backers Analysis**

1. Use the statistical information to determine whether the mean or the median better summarizes the data?

Answer to query:

* Since both the mean and median of our failed and successful set are within one standard deviation of each other. Both are acceptable to use with relatively safe assurances of accuracy. In this case I favor the median for both, since that is where the majority of our values lie and in both set’s the extremely large outliers are pulling the mean higher. Interestingly both sets have a mean and median that are roughly half a standard deviation from one another.

1. Determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Answer to query:

* There is significantly more variability in successful campaigns versus failed. This does make sense since due to the fact that the nature of startup businesses is the weed out factor will mean many programs flounder, while a few rise quickly to the top and see large amounts of success. Whereas the failed projects likely have number of similar structural issues which keep their numbers low. A successful project could be a failed project that set a low bar.