Module 1 challenge report

**Given the provided data, what are three conclusions we can draw about crowdfunding campaigns**

1. The Arts (film, theater, music, etc) are most numerous in terms of seeking crowdfunding, though in many cases they are less likely to succeed. Theater seems especially unsuccessful.
2. Overall, crowdfunded projects, at least the ones studied in this data set, are slightly more like to succeed than not. About a 60% success rate when live projects are included, mid 50% range when they are not.
3. The number of backers tended to greatly influence the success of projects. The more backers, the better the chances that the project would succeed.

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

Limitations of this dataset might include that it is not large enough to include all categories. Some of the results contain so few examples that it is hard to argue they would be statistically significant. Another factor not in the data is any correlation between a project’s founder (whether it is a person, corporation, organization, etc) and the success rate. This might also be helpful to know.

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

* By creating a table that sorts by state we might be able to uncover trends in terms of success rate
* Creating a table that sorts by the time between start date and deadline date might tell us how likely a project is to succeed with regard to the time it takes to get them up and running.
* In the data, there was a column called ‘spotlight.’ It would have been interesting to see how spotlighting a project would contribute to its success. I am assuming that ‘spotlight’ implies some sort of marketing or advertising effort.

**Statistical Analysis**

**Mean or Median?** In this case, the median number of backers serves to better predict the success or failure of a project. It is nearly twice the number for successful projects than failed ones, and though the mean number of backers still correlates to the outcome, there is a much smaller difference here.

**Variability with outcomes?** There is more variability with successful projects than with failed projects, as both variance and standard deviation are higher in this category. This makes sense to me in that unsuccessful projects will be more likely to have fewer backers, bringing the averages down to a smaller range of numbers. Successful projects are more likely to have a much higher number of backers, which would result in a higher variance and standard deviation.