1. The failure and success rate of each category is all relatively similar. There is no category that really outperform the other ones. One third of the project on Crowdfunding is all theater/play and it has about 60% success rate. It also seems like there is more project that gets launched in the summer. Something that I notice is that almost all project that is 100% funded end up being successful. The project that failed are the ones that didn’t get enough funds.
2. The sample size of some of the categories are relatively small which will result in inaccurate conclusion. For example, the success rate of journalism projects is 100%. This doesn’t mean journalism is an easy category where any future journalism project is guarantee to be successful.
3. I think a graph that we can make is a graph that shows the relation of project outcome to whether the project hits its goal. I would use a bar graph to compare the count of project outcome when a project meets its goal and when it doesn’t. This can show if meeting funding goal is a key factor to project success. Another graph that we can make is one that compares the countries of the project that it is in to whether or not is the project meets its target fund. This would also be done in a bar graph. This graph would show whether or not a project in certain country will be more likely to meet its target fund.

Bonus Statistic Analysis

1. I don’t think the mean and median of these data sets are meaningful. The data’s skewness and kurtosis are over what a normal distribution data should have. Both data are heavily skew to the left. There are lots of outliers. Just from the mean and median that we get from both data sets, we cannot confidently assume a correlation between the backers count and the outcome.
2. There is more variance with the successful data. The standard deviation is a lot higher compared to the failed data set. I think this makes sense because there are quite a lot of projects that attracts way more people and get more funds than their goal. However, these projects could have been successful anyways as long as they reached their target fund. Also, every backers could provide different amount of fund so 1 backer in 1 project may not be the equivalent of another backer in another or even the same project. The more backers there are the higher the variance there is. Successful project tends to have more backers than failed project so it makes sense for the successful dataset to have more variability.