Module 1 Challenge Analysis

Pivot table analysis

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   1. We can see from the third pivot table that during the summer time (June and July) that there is an uptick of successful campaigns that were created.
   2. We can also see that there were much more campaigns that were started for theater compared to any other categories of campaigns.
   3. When looking through all of the data, there were about half and half of successful campaigns compared to the failed and canceled campaigns.
2. What are some limitations of this data set?
   1. Since each campaign may be set very differently so we don’t really know what was the main reason why the campaign was successful or failed.
   2. This data is too broad in terms of the specs of the analysis. We would need to get more data sets in the same categories to see why they may have succeeded and failed.
3. 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 pie chart instead of a bar chart for all of the different outcomes filtered by categories to show what percentage of the category’s campaigns were successful compared to the other outcomes.
   2. We could create a different line chart that graphs out all of the different months and years instead of grouping them together by months. This will allow us to see if there is a increase or decrease in the likelihood of the campaigns succeeding or failing.

Statistical Analysis

1. We can see that the median would definitely be a better way of summarizing the data. The outliers of some of the successful and failed campaigns move the “middle” of the data much higher for the mean, while the median always takes the middle value.
2. We can see that since the variance and standard deviation for failed campaigns are lower than the successful ones, the successful campaigns are much more variable. This does make sense since if some campaigns did very well, it could make the whole data set much more variable. If it failed, we know that it did not do up to par in terms of funding, which means they were always going to be a lower value.