Model One challenge

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

While viewing the data sets, I can see that there are three main categories. The most crowdfunded group would be theater followed by music tied with film & video. We can conclude that Theater was heavily crowdfunded over all other data points. We can also consider plays in the theater category to be an outlier due to the drastic difference between all other categories. We can see the difference in the second pivot table, although there are many failed crowdfunding attempts, the success rate is way above the rest. We can see a in success and fail rate between July to September, a few reason that come into mind is the change in weather. Due to colder weather people are less likely to be outside supporting groups. We can also see a slight increase in failed data towards the end of the year followed with a decrease of successful data most likely due to holiday.

* What are some limitations of this dataset?

Due to the outlier being extremely drastic, most of the smaller fields of information tends to be overlooked. Small data sets are almost invisible and look smaller compared to an outlier. Due to information being smaller compared to the outlier it’s harder to view the actual progress smaller businesses made so far. When broken down into subcategories you can see this effect further. Another limitation would be how drastic the variables are within the data. This can be seen towards the end of the assignment by viewing the mean median and range.

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

I believe a pie chart, or a scatter plot would help aid in viewing the information. Using a pie chart allows for a even visualization of all categories so that even if a category is small it’s still considered in the data to have an adverse effect. Scatter plots are also beneficial because it removes the outliers of the data set, and you can visualize all other points in correlation with one another.

* Statistical Analysis

I don’t think the mean or median justify the data set that well, the values are far in between from the maximum data value. The mean although shows the average, the data point can be skewed by large outliers as seen in the data.