1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   1. The first conclusion I would draw is that there are a lot more campaigns being done in the theatre category compared to categories like journalism and photography. When we take a deeper look at sub-categories, we can see that plays dominate the campaigns and have a grand total of 344 outcomes, with the next closest total being 85 outcomes from rock. From the line chart we can see that campaigns trend towards success starting in May, then they take a dive down from July to August, so it would be best to start a campaign in May.
2. What are some limitations of this dataset?
   1. The number of tests each year varies, so if we did any analysis using a trend chart or something similar, then we would have to be aware of this. Another limitation is that we do not know if certain campaigns had advertisements going before data was taken. We also do not know if the successful projects had failed before, and then tried multiple times to get the right kickstarter going.
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   1. Another possible table that we could create would be a trend line on the data vs years. We would have to consider the difference in sample size for multiple years, but it would still be interesting to see how the data changed through the years. Another possible graph we could create would be to see what the relationship between the outcome and the goal column. This would allow us to see if the goal was met, and if so, was it passed, by how much, etc.…

Statistical Analysis

1. Use your data to determine whether the mean or the median better summarizes the data.
   1. I would say that the median better summarizes this data. The mean for both datasets was hundreds of values higher than the median. This is because of the max of both data sets being very large. This causes the mean to not represent the whole data set.
2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
   1. There is more variability in the successful group. This does make sense to me. The mean of the successful group is 851.15 and the median is 201. There is a large gap between the average of the data, and the middle number. The variance is how far each value is from the mean, so it makes sense that the successful values have a higher variability.