* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  1. The number of campaigns related to the creative arts such as films, music, and theatre are higher as compared to other categories. The subcategory “plays” has the maximum number of crowdfunding campaigns, and their success rate is almost 50%.
  2. The success rate was high when the goal was in the range 15000 – 35000.
  3. The success rate is higher during May to August.
* What are some limitations of this dataset?
  1. The one limitation that I see in this data is that I’m not able to figure out the reason behind the higher success during May – August. Is it just a coincidence? What can we do to produce similar results for other months? The other thing that noticed is that the number of campaigns for theatre, music, and films is high. Is it because they have more backers or has it something to do with the funds required for such campaigns?
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. The other table that we could have included is the subcategory vs the average of goal. This would help us study how the success of each subcategory is related to their set goal. I have included this table in my analysis. A box plot can be included to study variance and the outliers.
* Use your data to determine whether the mean or the median better summarizes the data.

The median better summarizes the data as the mean is skewed by outliers. There is a large difference between the minimum value and the maximum value for both the successful and the failed campaign.

* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variability with successful campaigns. The difference between mean value and the standard deviation value is higher for successful campaigns. It doesn’t seem like a normal distribution curve for both the successful and the failed campaigns as the mean and median value are pretty different.