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

The first thing I notice is that there are more crowdfunding campaigns that are successful than there are ones that fail. Out of 986 campaigns 565 were successful at 57% compared to the 36% that were not successful.

Secondly, these campaigns are more effective in June and July, where these campaigns were successful 64% of the time in June and successful 62% of the time in July.

Lastly, campaigns that fell in the audio subcategory had a success rate of 100%, but there were only 4 successful campaigns. This is compared to campaigns that used the subcategory plays that were only successful 54% but had 187 successful campaigns.

1. What are some limitations of this dataset?

This data is limited by having outliers as stated above. If we were to just look at success as a percentage, we would think that crowdfunding campaigns with an audio subcategory would be the most successful type of campaign. However, there is not enough data on some of these categories to get an accurate read on how they perform within a larger sample size.

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

I believe looking at the average goal by subcategory versus average pledged by category we can maybe get some insight into why some categories were more successful than others. With this we can see that the “audio” subcategory was able to be more successful because their average goal was only $6k while “plays” had an average goal of $45k.

1. Justification for why mean or median better represents the data

I believe that the median best represents this data. We see that the median for both Successful and Failed campaigns are lower than the mean, which shows a positive skew.