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

One conclusion that we can draw is that crowdfunding campaigns are more successful in the summer and less successful in the winter. We can see on the line graph that the number of successful campaigns is higher in the summer months and that there is a rise in cancelled and failed campaigns in the fall to winter months. Another conclusion that we can draw is that the most common categories of crowdfunding campaigns are theater, film & video, and music respectively. A third conclusion that we can draw from the data set is that August is the worst month for crowdfunding campaigns, as this was the point where the number of successful campaigns were the lowest and the number of failed was the highest.

What are some limitations of this dataset?

One limitation of the dataset is not being able to see the methods of crowdfunding used and thus not being able to analyze which method works better than others. Another limitation is that only a handful of countries are included, and they are all western countries. Perhaps countries like Japan would also have data about crowdfunding that could be analyzed.

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

One other table we could create is average donation by category, this would show which categories have the most likelihood of receiving large amounts of funding. We could also create a backers count table by category, which would also provide some insight into which category has the most support.

Does the median or the mean best summarize the data?

I believe that the mean best summarizes the data because it shows a greater difference in between the number of backers for successful and failed campaigns. Successful campaigns had an average of about 851 backers while failed campaigns only had an average of nearly 586 backers. The median values were within 100 backers of each other, so I do not think this would best summarize the data.

Is there more variability between the campaigns that were successful or the ones that failed? Does this make sense? Why or why not?

There is more variability between the campaigns that were successful than the ones that were unsuccessful. This makes sense because there was a larger average in successful campaigns, and a higher total number of successful campaigns. Variance also shows preference to outliers, in which case successful campaigns are also favored since there cannot really be an outlier of number of backers in failed campaigns. Really popular campaigns, however, would appear in this data set as an outlier.