Kickstart My Chart Questions:

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

We can conclude from the data that campaigns that are ‘Staff Picks’ have a much higher success rate than those without the staff backing. We can also gather that historically the month of May seems to have the highest success rate of all the months. My final conclusion is that the ‘Fine Art’ categories, such as: film & video, music and theater, primarily make up 77% of all the successful campaigns in our sample. Also, those 3 categories seem to have the highest probability of success when compared to the other 6.

1. What are some limitations of this dataset?

The limitation of this data, like most data, would be that it is strictly internal, and gives little to no insight into any external factors that may have played a part in a specific campaign’s success. The data shows us that being a staff pick increases the chances of being a successful campaign, so any number of external factors, such as internet virality, for example, could have played a part in any number of successful campaigns. Another limitation would be the lack of knowing what even deems a campaign a staff pick. Another limitation is we are only able to surmise how to create a successful campaign with this data as all we can truly conclude is that a successful campaign gets fully backed while a campaign that fails does not. We are limited by not knowing how much was spent on marketing per campaign or what platform each campaign used for marketing. In short, we have uncontrollable and unknown variables that could give us a better more meaningful insight into what this dataset represents, and this datasets inability to forecast the future.

1. What are some other possible tables and/or graphs that we could create?

A pie graph showing which successful campaigns was a “staff pick” vs campaigns that were not. Bar chart showing the distribution of successful and failed campaigns based on origin country. A scatter plot to see if there is any relationship between amount funded and backer count. Another interesting scatter would be to see if there is any kind of relationship between average donation and backer count.

Bonus Statistical Analysis:

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

I would say that the median would more accurately portray the number of backers due to the average being skewed by outliers in the top 1%. The main thing that led to this decision was the extreme variation between the mean and median, and another quick indicator was the 26456 difference in the minimum and maximum.

1. 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. This made sense to me especially when I was writing out the limitations of what this dataset could do. In the dataset there is no clear-cut indicator other than pledged was higher than or equal to goal, and everything in between that happens is variable and unknown.