Kickstart my Chart

1. The first conclusion that we can make from this chart is that if you wanted to start a Kickstarter campaign, your highest change of success would be in the music category in the rock subcategory. The rock subcategory had a 100% success rate with the largest sample size of all the subcategories that also had a 100% success rate.
2. The second conclusion that we can draw from this data set is that it is not wise to start a Kickstarter campaign in November or December as the data shows that the number of successful projects decrease in November until it falls below the number of failed projects in December for the first and only time all year.
3. The third conclusion that I came to is that a higher average donation does not necessarily mean a higher chance of success. The top 3 total averages of the average donations per categories listed from highest to lowest are technology ($133.74), film & video ($87.44), and photography ($81.97). Out of the top 3 total averages only 1, which is film and video, had a success rate of over 50%.

Limitations:

1. One limitation of this data is that the categories are not equally represented. Theater has close to 1400 projects whilst food and games only have around 200 each.
2. Another limitation I noticed is that most of the data is from the US. If you were trying to get a fair representation of Kickstarters outside the US, I would not use this data set as 73% of the data comes from Kickstarter campaigns in the US.

Extra Charts and Pivot tables:

1. A good pivot table that you can create is one that shows average of average donations and average of percent funded per category and subcategory. This allows you to see if a higher average donation or percent funded lead to a higher chance of success. This is the table I used for my third conclusion.
2. Another good chart and pivot table to create is one that identifies which categories and how many were featured in the staff pick. For the chart I would use the standard column chart that shows how many categories were true and how many false filtered by state. This would allow us to see if being featured as a staff pick gives us a higher change of success.

Statistical Analysis

1. For the successful campaigns I would use the median as the variance and standard deviation are quite large and there are more extreme numbers in the data set. For the failed campaigns I would use the mean as the standard deviation is quite close and there are not too many extreme numbers in its dataset.
2. There is higher variability with the successful campaigns. This makes sense as there are many more extremes within the successful campaigns and the difference between the min max is much higher. The SD and Variance also is much higher suggesting a higher variability