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

* The month a campaign is launched is has a significant correlation to its success/failure rate. While cancellation rate remains relatively steady throughout the year, December is the only month during which the success rate falls below the rate of failure. The success rate remains steadily higher than the rate of failure from Feb-June and falls closer to it from Aug-October and January.
* Film/Video, Music, and Theatre make up nearly 2/3 of the Kickstarter campaigns, and all have over 50% success rates. While Technology has more campaigns, the success rate for these is around 30%. We can therefore conclude that contributors are more likely to a Kickstarter in these types of art than to other campaigns.
* Kickstarter campaigns in the subcategories of Classical Music, Documentaries, Hardware, Metal, Nonfiction, Rock, Shorts, Small Batch, Tabletop Games, and Television have success rates of 100% or extremely close to it.

1. What are some limitations of this dataset?

There is no information about how the campaigns were advertised or how many people viewed the campaign page. We also do not know how campaign contributors learned about the project/s they decided to contribute to. Our data is also limited in terms of the categories of campaigns we are examining (for example there is no data about campaigns to raise costs that are associated will illness and healthcare). Since 2017 the last year we have data, we do not know if/how Kickstarter trends have changed over the last 4 years.

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

It would be useful to make a table/stacked bar or line graph comparing the campaign goal to the success rate. Doing a box and whisker plot for percent funded in categories and subcategories would be interesting as well, though there are some extremely high outliers in the percent funded column of our original sheet. We could also compare the bakers count and average donation to categories, subcategories, and success rates using pivot tables and scatterplots.

Bonus:

When studying the statistical analysis of the tables comparing the number of backers to the outcome of success/failure, it is important to note that the median of said datasets summarizes the data more meaningfully than the mean. This is true because of the large number of outliers, particularly in the number of backers of successful Kickstarter campaigns. Additionally, there is more variability in the number of backers for successful campaigns than unsuccessful ones. This makes sense because unsuccessful campaigns can only have x number of backers until they hit their goal and become successful, whereas successful campaigns can continue to get an unlimited number of additional backers even once the goal has been reached. Moreover, there are a substantial number of campaigns with very low goals, meaning that it is often unnecessary to have a large number of backers to be successful (leading to more variability in backer numbers for successful campaigns).