Analysis of Kickstart My Chart Data

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

1. First conclusion: In the first tab titled “Campaigns Category”, theater focused categories made up a majority of the Kickstarter campaigns and they seem to of been domiciled mostly in the US as a country. The US led this category as a whole when compared to other countries. We comprised 65% of the total amount of these campaigns which is quite significant.
2. Second conclusion: “Plays” as a sub-category were successfully funded more so than any other sub-category if you look at the tab titled “Campaigns Sub-Category”. Of the 2185 Kickstarter campaigns that were successful here “Plays” made up 32% of the total.
3. Third conclusion: The month of May seems to be the best time of the year to start a Kickstarter campaign as it leads to the best chance for your goal to be 100% or more funded. Please look at tab titled “Campaigns Dates”.

Question 2. What are some limitations of this dataset?

1. A limitation that I could see in this dataset is since we are collecting this information from various countries, I would imagine that there are cultural differences that would play into a campaign being successfully funded. For example, theater as a category in Italy (IT) is not as popular as technology focused campaigns when compared to the US which is more of an entertainment consuming culture.
2. Another limitation that I hear often about these Kickstarter campaigns, is yes, they could be successfully funded but could take years for a product as an example to come to fruition. Sometimes promises are made but not kept. And by the time these products do come out they could change from what was originally promised.

Question 3. What are some other possible tables and/or graphs that we could create?

Answer: I line graph could be used to interpret data over a course of years to see which KickStarter campaigns were the most successful, this would help you see trends in the data as to which Categories seem to be popular during which years.

Bonus Statistical Analysis (“Bonus2” Tab)

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.
   1. I would say that median is the best measure to be used with these 2 datasets dealing with successful and failed Kickstarter campaigns because overwhelmingly there are more successful campaigns then failed. The mean is dealing with a bigger dataset on the successful state versus the failed state. This would in my opinion skew the numbers because failed campaigns seem to have a lower number of backers overall.
2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
   1. I would say there is more variability with successful campaigns due to the fact that the # of backers could be greater due to goals on each of these campaigns being overfunded by a dramatic margin.