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

There are multiple conclusions that can be drawn from the data but some of them include:

1. The parent category *Theater* constituted a 1/3rd of the data presented/analyzed
2. Due to such a high number of *Theater* campaigns, it naturally had the highest number of successful campaigns but also failed campaigns.
3. However, the parent category *Music* had the highest percentage of successful campaigns at 77%.
4. Majority of the currently live campaigns are for *theater* and *music*.
5. There were no successful or failed campaigns for *journalism*. All of them got canceled.
6. Overall success rate of all the campaigns is above 50%.
7. Of the sub-categories, within the *Theater* category, *plays* had the bulk of campaigns (26%).
8. Of the 4114 total campaigns, almost 3/4th of were in the US.
9. Of the successful campaigns, the most occurred late spring into early summer (May/June).
10. Number of campaigns built up year on year, maxed out in 2015 (1225 campaigns), before dropping again in 2016 (something must have happened in 2016).
11. What are some limitations of this dataset?

The biggest limitation I see is the bias towards the parent category *Theater*. It is tough to predict how the other categories would perform as they don’t have too much data. You would be assuming they would perform like *Theater.*

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

You could possibly use a pie/circle chart for the parent categories for successful campaigns, or pie charts for each category demonstrating the sub-categories. It may be also possible to use bubble charts to demonstrate successful vs. failed campaigns for all the parent categories. You could also possibly use a spline chart for analyzing dates of all the campaigns.

Bonus Statistics:

1. Does the mean or the median summarize the data more meaningfully?

It generally varies between data sets, however in this case it appears both successful and failed campaigns have such high standard deviation that median demonstrates the bulk of the data whereas mean is inaccurate due to the outliers.

1. Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or why not?

There is more variability with successful campaigns, which makes sense, because depending on the goal of the campaign it had a number of backers. The higher the goal, the more backers needed for it to become successful. The range is much greater for successful campaigns, from 1 to 26,457 backers, than failed campaigns, from 0 to 1,293 backers.