Q1: Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

1. Projects that start / are created in May see greater success than failure.
2. Generally speaking, the more money needed to raise, the less likely the campaign will see success. (Based on Data from “Bonus” sheet). This is shown that sub $1000 projects succeed at 71%. The average percent of projects that succeed under $10,000 is 63%. The percent that succeed falls from there.
3. Music campaigns have a higher success rate – outside of Jazz, Faith, and World subcategories, which all failed or had projects currently live. The other subcategories performed well.

Q2. What are some limitations of this dataset?

1. We only have a small dataset. Some of the categories have too few projects to draw conclusions from, such as Journalism. Pulling more data for a larger sample size would be more helpful for categories with fewer campaigns.
2. Some of these projects may have successful Kickstarter campaigns, but it would also be helpful to see how many succeed after their campaigns so we know which ones will continue to perform after backing them. This would mean additional data from other data sources to augment this set. We would also need to define “success” and gather the measurable factors we need for that definition.

Q3. What are some other possible tables and/or graphs that we could create?

1. A box-plot would be helpful to see outliers in our backer-count and amount raised for campaigns that succeed, we could also do this for campaigns that fail. We could address the outliers accordingly. This would mean several box plots, but we could do two for each campaign state: Successful Backers, Successful Goal, Failed Backers, Failed Goal.

**Bonus Statistical Analysis**

Q1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

***Successful Campaigns*:** With such a high variance and standard deviation, the median is probably a better representation.

***Unsuccessful Campaigns:*** A much lower variance and standard deviation comparative to successful campaigns, but there are outliers. Again, median is probably the best representation.

Q2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

1. There does tend to be more variability with successful campaigns, perhaps because the different thresholds for Goals ($1,000, $5,000, etc.) can be met with different numbers of backers depending on the average amount each backer donates. It would make sense that with few or no backers for unsuccessful campaigns that they would fail, since they would require a higher average donation per backer.