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

There were three distinct conclusions about Kickstarter campaigns based upon the data provided:

* 1. Theaters were a favorite among other campaigns, but music Kickstarter campaigns were less likely to fail than theaters. Music had a success rate of approximately 77%, whereas theaters had a success rate of 60%. We can also conclude that journalism campaigns are less likely gain traction, as we saw no successful campaigns. Additionally, food and game campaigns were the highest failure rates at 70% and 64% respectively. See Pivot1 worksheet.
  2. Plays, a sub-category of theater had the greatest number of successful campaigns across all categories. We can also see that rock music campaigns are widely successful, where all (100%) campaigns started were successful. Each of these points align with statements raised in the first (i.) conclusion. See Pivot2 worksheet.
  3. The largest number of successful campaigns occurred in the month of May, whereas there is a sharp decline in the number of successful campaigns occurring in the month of December. The number of cancelled campaigns were consistent month over month with only a couple of peaks in July and October. The number of failed campaigns varied and had high points during the months of January and October. It is also interesting where the number of failed campaigns exceed successful campaigns in December. See Pivot3 worksheet.

1. What are some limitations of this dataset?

If someone would want to analyze why campaigns fail or get canceled, more data would be needed to determine if there is a common cause and effect.

Also, we are not given the definitions of the meaning of ‘state’ values in the data. For example, the term ‘live’ could be determined to be a type of successful campaign, however we are not accounting for such states when evaluating successful campaigns.

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

I believe pie charts could be used to show the percentage of the various states of campaigns. See below for an example.

Also, scatter plots could be used to show the relationship between two variables, such as the table below. This can also highlight any potential outliers.