Kickstarter Data Analysis

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?

As seen with the line chart in the Pivot tab, campaign success is high in May and low in December. One could speculate that general consumer spending is higher towards the end of the year, and they may not be as willing to fund campaigns then. Based on the first Pivot Chart comparing project category to campaign success, the categories of "theater" and "music" have the most campaigns with "successful" standings per group. This means that these projects are the most common on Kickstarter, with a greater than half success rate (60% of "theater" campaigns are "successful", and 77% of "music" campaigns are "successful"). However, in comparing those two arts categories to "food" or "games", there appears to be a difference in success trends. Food has a majority of "failed" campaigns, with a mere 17% success rate within the category. "Journalism" shows only "canceled" campaigns; there are no "successful", "failed", nor "live" campaigns from the original data.

The second Pivot Chart shows campaign success compared to sub-category. All sub-categories have fewer than 200 campaigns, with the exceptions of "rock" and "plays". "Plays" shows the biggest spike, with 694 "successful" campaigns, 353 "failed" campaigns, and 19 "live campaigns, totaling 1066 total campaigns. These two sub-categories explain the spikes in the categories of "theater" and "music", as "rock" and "plays" are subsets of that parent category. Some points of interest may include that all 260 "rock" campaigns were "successful", all 40 "pop" campaigns were "successful", but all 60 "jazz" campaigns "failed". This shows a correlation of certain musical genre campaigns and campaign success.

Based on the third Pivot Chart, it can be concluded that campaigns tend to be successful if they are created in the late Spring and early Summer months, with a low point in Winter. May shows the highest number of overall campaigns that were labeled "successful", while December shows the lowest number of "successful" campaigns. Overall activity on Kickstarter may be lower in the Winter because of the holidays, when people may not be as inclined to fund campaigns during this time.

2. What are some of the limitations of this dataset?

A limitation of this data set is that we don't know the person or organization behind a campaign. It could be useful to analyze and compare the campaign creator's popularity or marketing success, to funding success. Another limitation of the data set is uncertainty of the sample being representative of the true population. Categories and campaign results may be a bit skewed based on how the sample was selected. Another limitation is locational trends. There is not geographical data in the data set, and this could help analyze different trends among different countries/regions.

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

It could be useful to create a line plot which maps out the length of campaign and campaign success. A hypothesis that could be tested is if longer campaigns do better with hitting their funding goals. This would be an interesting thing to graph because shorter campaigns also may do better, because the focus is on a short period where urgency is a theme.