Kickstarter Campaign Analysis Report

1. Given the data provided there are many conclusions we could draw. The most successful campaigns were those in the category of music, having a higher success to failure ratio than other categories, like theater which had the highest number overall of campaigns, but also had a pretty high failure rate, the worst being food which had a much higher rate of failure than success. All journalism campaigns were canceled which does not give us enough data to compare it to other campaigns completely. The most successful subcategory of music was rock music have no failed campaigns and the most successful campaigns within the category. The least successful subcategory of music was jazz have the most failed campaigns and no successful campaigns. There appears to be no correlation with success or failure of campaigns based on month, however most live campaigns appear to be during the months of January, February, and March. The highest percentage of successful campaigns are those with a goal less than $1000 and the lowest percentage of success are those with a goal great than or equal to $50,000.
2. There are some limitations within this data set that could be pointed out. The most obvious one is the question of why. You can make many correlations within this data set as to which campaigns were successful, failed, or canceled, however there is not any data provided which can explain why those campaigns were so. If we had some quantifiable from things such as marketing strategy, target audience, reach, social media presence, etc. it maybe be possible to make a better determination as to why. The only data close to this we have available is whether they were a staff pick or a spotlight, which is helpful, but not enough. Another limitation within this data set which is missing is the individual(s) behind the Kickstarter. Success or failure of a Kickstarter can easily be determined by the level of celebrity of the individual(s) starting it. Someone who already has a certain level of notoriety would likely have a higher chance of success by something as simple has having more followers on social media or having the ability to market their Kickstarter through outlets unavailable to someone with no celebrity.
3. There are some additional tables or graphs we could create to determine more possible correlations. We could create a pivot table with the number of successful, failed, canceled, and live campaigns based on staff pick and spotlight with a stacked column pivot chart or scatter chart to display this data. We could create a pivot table to break down the data into success vs failure based on currency and display this again with a simple stacked column pivot table, or chose to break down the data into successful, failed, canceled, and live, each being displayed in a pie chart showing each currency.

Bonus Statistical Analysis:

Given the data available in the summary statistics tables created, the best choice to summarize the data available would be the median number of backers for both successful and failed campaigns. This is because the mean is heavily skewed by the fact that in both successful and failed campaigns, we have at least one campaign with a high number of backers relative to the rest. This is easy to see in a histogram where we can see that most of our data is on the lower end as far as number backers. This could also be easily determined by looking at the mode and seeing that the mode is close to the median than the mean is.

There is a much higher level of variability in successful campaigns than in failed campaigns. This make sense given the data because in general, the failed campaigns have far less backers therefore decreasing the opportunity for any variability.