Conclusions

Based on the dataset provided, Kickstart campaigns were generally successful for the film & video, music, and theater categories. Kickstart campaigns were not very successful for the food, games, photography, publishing, and technology categories. The Kickstart campaigns were generally more successful in the United States and Great Britain where the largest number of campaigns were run.

Data limitations for this dataset include the

There are data entries that provide little information or are not defined, like “Spotlight” for example. All successful campaigns show “TRUE” for this data field while failed, canceled, and live campaigns show “FALSE”. It would be interesting to learn more about what the “Spotlight” field means and whether or not it predicts success. Data fields like “name” and “blurb” are difficult to analyze using Excel. The quality of these items may contribute to the success or failure of a campaign, but it is difficult to look at these fields when there are so many and each is unique and the quality of this content is subjective. In addition, we lack detail regarding the demographics of the specific location of each campaign. Beyond country and currency, culture, affluence, and population density would be a few additional parameters that may provide insight. Perhaps the number of people each campaign reached could predict a campaign outcome – we have the number of backers but not the number of people who were solicited. Regionally, perhaps different music genres are successful in one culture over another. Finally, the canceled campaigns do not provide additional rationale as to why the campaign was canceled. I would lean towards grouping the cancelation with failed campaigns but need more information, like why it was cancelled, to conclude this.

Additional Analysis

It would be beneficial to create a chart showing the relationship of campaign success and how many people donated to a campaign by looking at the number of backers compared to campaign success, failure, and cancelation. Another analysis of campaign success may be to examine the relationship between the campaign goal and outcome to see if the target amount correlates to campaign success. A new table of summary statistics for the mean, median, minimum, maximum, and standard deviation of average donation, backers count, and goal by campaign outcome could be created to look at these factors. To better assess the global information, a tool like Tableau would be useful in showing campaign concentration and success/failure/cancelation/live rates in countries represented on a map rather than looking at individual country data from a drop down filter.

Bonus

Looking at the summary statistics for the count of backers with respect to campaign outcome, the median value for the successful and failed datasets more appropriately summarizes this data because there are outliers in the data that skew the mean, making it higher than the typical datapoint. The standard deviation is very high relative to the mean, indicating that there are data points that differ from the mean significantly. The median value is nearer to more datapoints than the mean because it is taken from the middle of the dataset where most of the data is represented. There is greater variability in the number of backers for successful campaigns as is shown by the standard deviation of 844 compared to a standard deviation of 64 for the failed campaigns. This makes sense because a few of the successful campaigns had very large donations which offset the data, increasing its variability.