Kickstarter Campaign Report

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   * The category of theater campaigns had the largest amount of *total* campaigns, while the music category had the highest percentage of *successful* campaigns (79% of completed music campaigns were successful, while 60% of completed theater campaigns were successful). If you were a backer who wanted to back only one campaign and maximize your chances of a successful investment, the data suggests a music campaign would be the safest bet.
   * Food, games, and publishing have the highest percentage of failed campaigns and lowest percentage of successful campaigns. The sub-category stats for these groups show that their individual campaigns have struggled; for example, video game campaigns failed 100/100 times. The data suggest these campaigns have the greatest likelihood of failing. Technology campaigns had the largest amount ofcancelledcampaigns, which suggests it could also be a risky investment. All journalism campaigns were cancelled, so this would also be a category to avoid.
   * December is the month where the least amount of total campaigns was launched, and also the month where the least amount of successful campaigns was launched. This suggests December is the worst month for launching Kickstarter campaigns.
2. What are some limitations of this dataset?
   * We do not know the reason why a campaign was canceled, failed, or successful. This is particularly a limitation when it comes to canceled campaigns, because there is no way to tell why the campaign was canceled. For example, technology campaigns had a large number of cancelations, but the reason behind those cancelations is not included. This makes it impossible to tell if cancelations are a reliable indicator of a campaign’s success chances.
   * Live campaigns do not have outcome data yet, and the outcome of those campaigns could affect the overall numbers of their campaign’s category.
3. What are some other possible tables and/or graphs that we could create?
   * A pivot table of failed campaigns and the number of backers they had, split up by categories/subcategories.
   * A summary statistics table for the outcomes of campaigns broken down by category, and a summary statistics table for the outcomes of campaigns broken down by subcategory.

BONUS

* For both failed and successful campaigns, there is a large amount of variance in the number of backers and both have significant outliers in the data set. Therefore, median would be a better metric for describing the number of backers for both successful and failed campaigns. Since the median number of backers for a successful campaign is much higher than the median number for failed campaigns, this indicates that number of backers could be a valuable predictor of a campaign’s success.
* There is a large amount of variance for number of backers in both successful and failed campaigns, but the variance is higher for successful campaigns. This indicates that there is more variability in the number of backers per campaign and that the data points are very spread out from the mean and each other. Based on the median, it seems that a greater number of backers is a positive indicator for the success of a campaign, but the variance shows that the number of backers per successful campaign varies greatly. This makes sense because we are looking at data for all of the campaign categories together; for example, a theatre campaign may require more backers to be successful than a technology campaign. Because we are analyzing multiple different types of campaigns, it makes sense that there is greater variability in our data set.

NOTE: For the date ended/date started conversions, I applied my formula down the entire column. This entered some random dates into the blank cells (like 1/7/70) and my computer kept freezing when I tried to clear them all because it filled in so many cells. I have unselected this date on all of my pivot tables so it does not interfere with anything, but I apologize for that mistake/extra cell values. I will be sure to not make this mistake again!