Kelly Curtiss

Excel-Challenge

KickStart My Chart Report

June 13, 2020

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Given the provided data, I can conclude that there were more successful (2185) campaigns than failed (1530).
   2. I can state that the highest number of campaigns fall in the Theater category (1393) and the fewest campaigns fall in the Journalism category (24).
   3. If you look at the percent rate of successful campaigns for each category, you will see that Music had the highest success rate with 77% successful. Theater came in second with 60% success. The least successful campaign would be food at 17%. I would not say that journalism is the least successful, because they all still listed as live campaigns.
   4. You can also determine which sub-category was most successful according to percent rate. The most successful theater sub-category was plays at 65% successful. To determine this, I looked at the percent rate of success when compared to the total number of campaigns in that category and/or sub-category.
   5. When looking at the data by the time of year the campaign occurred, you can conclude that the number of campaigns is relatively consistent by month except for December. They all fall closely within 300-400 campaigns.
2. What are some limitations of this dataset?
   1. One limitation of this data set in regard to determining success, is that there is no data on marketing or more specifically how much money was spent on marketing the campaigns outside of Kickstarter.
   2. Another limitation when trying to determine success was somewhat addressed above. How are you measuring success? If it is just by the highest count, you could be misleading. Theater is the perfect example. By count, theater has the highest success, but by looking at the percent rate of success, even though music has many fewer campaigns, it’s rate of success is higher.
   3. A third limitation of the data is that there are campaigns being included that are still live. The results or conclusions regarding overall success could be skewed based on their final outcomes.
3. What are some other possible tables and/or graphs that we could create?
   1. I would look a table that displays the percentage rate of success and the percentage rate of failures for categories and sub-categories.
   2. I would also look at a chart/graph that displays the percent funded by category or sub-category.
   3. A third chart/graph could show the average donation by category and/or sub-category.
4. Does the mean or median summarize the data more meaningfully?
   1. When trying to determine if the mean or median provides a better summary or representation of a data set, you need to consider items like sample size and outliers. It is also helpful to look at the graph of the data to see which value provides a better representation. In this scenario we are looking at the number of backers per campaign. The mean or average number of backers for successful campaigns is 389 whereas the median or the middle number of backers is 62. The maximum of value for successful campaigns is 424819, but the upper bound is 309 and the lower bound is -139. This clearly demonstrates at least one outlier. When there are outliers this far from the data set, it will greatly impact the mean value. This maximum value/outlier has caused the mean value to also be larger than the upper bound value. This tells me that the median of 62 is a better representation of the data as it falls between the lower and upper bounds.
5. Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or Why not?
   1. Data sets that are more spread out have high variance values. The successful campaigns dataset has a much higher value for variance than the failed campaigns. Therefore, it does make sense to say that the successful campaigns have more variability. The standard deviation also measures how close or spread out the values in a data set are. The standard deviation for successful campaigns is much larger than the failed campaigns. This is also an indicator that there is more variability with successful campaigns.