Aaron Rotman

03/07/2020

Excel Challenge Homework Report

We can draw several conclusions from our dataset. The first conclusion we can draw is that our dataset has a higher than average proportion of successful campaigns compared to all Kickstarter campaigns. Over half of the campaigns in our dataset were successfully funded. That is compared to the ~33% success rate for all Kickstarter campaigns according to the statistic given in the instructions for this assignment.

We can also see that among the categories examined, theatre had the highest numbers of campaigns in general, as well as the highest number of campaigns that were successfully funded. Within the category of theatre, plays were the sub-category with the highest number of successful and overall campaigns. This could indicate that Kickstarter is a good funding source for theatre related projects, especially if they are plays.

The least successful category was journalism with all (24) campaigns in this category being cancelled. With no successful journalism campaigns we could conclude that Kickstarter is not a likely source of funding for journalism related projects and that those seeking funding for journalism may want to look elsewhere.

One limitation of this data set is the small number of live campaigns included in the data. With this small sample size, we should be cautious about any conclusions that we draw as it may not be representative of all live campaigns. Another limitation of this data set is the lack of current data. The most recent data is from 2017. It is possible that the criteria for a successfully funded campaign has changed in the three years following the end of this data set.

Another interesting metric that we could look at would be the percent of campaigns that were successful/failed/canceled/live per category and sub-category. This would allow us to see the success rate of campaigns in each category or sub-category. The success rate as opposed to the simple count of outcomes that we calculated would be a more useful metric to compare the success of campaigns in different categories. These values would be best visualized with a summary table and a stacked bar graph that could be filtered by category and another that could be filtered by sub-category.

We could also create a summary table and boxplot to compare different metrics for the amount pledged per backer. This would be interesting to see how the mean compared to the median and the mode. This might help us to understand whether successful campaigns received a high number of low dollar pledges, or if a few high dollar pledges made up the majority of their backing. A small number of high value pledges could skew the mean in which case the median might be a more accurate metric to gauge the average backer.