Paige Singleton

Excel Challenge

GA Tech Bootcamp

December 2020

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Over 50% of all campaigns are considered successful (2185/4114)
   2. Theater Campaigns accounted for 1393/4114 (approximately 34%) of all campaigns; there were more successful theaters campaigns that total campaigns in any other parent category.
   3. Campaigns have a higher rate of success during the month of May than during any other month of the year. Campaigns have a lower rate of success in December.
2. What are some limitations of this dataset?
   1. Backers are not divided into categories (such and major and minor backers). This limits the amount of analysis on the effect of backers on the success rate of campaigns.
   2. States of campaigns (successful, live, failed, cancelled) are not well defined. Campaigns are not divided into types based on pledge amount. This limits the analysis of campaigns based on size.
   3. Average donation is not necessarily a useful metric, in and of itself. Additional metrics, such as minimum, Maximum and median are needed regarding backers and their donations in order to place donation amounts in a context.
3. What are some other possible tables and/or graphs that we could create?
   1. The line graph that we graphed showed a monthly trends. Another graph to consider is a line graph that showed yearly trends.
   2. A scatter plat could compare the Goal with this percent funded to see if there is a relationship between the goal amount and the level of backing for the campaign.
   3. A table could be used to show the state of campaign given certain ranges of goals and/or total number of days of the campaign from date creation to date conversion.

Bonus – Statistical Analysis

**Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

1. Successful Outcome: The mean and the median are working together to help us analyze the data. By themselves neither is completely meaningful. With a min of 1, a max of 26457, a median of 62, and a mean of 194.4, we can infer that there are some outliers, such as 26457, in the data on the high end of the dataset. The median of 62 tells us that half of the data points are less than or equal to 62. The lower end of the dataset is dragging the mean down. This data is skewed right.
2. Unsuccessful (Failed) Outcomes: The mean and the median are working together to help us analyze the data. By themselves neither is completely meaningful. With a min of 0, a max of 1293, a median of 4, and a mean of 17.7, we can infer that there are some outliers, such as 1293, in the data on the high end of the dataset. The median of 4 tells us that half of the data points are less than or equal to 4 The lower end of the dataset is dragging the mean down. This data is skewed right.

NOTE: Together, these metrics seem to indicate that the number of backers may be a predictor of campaign success. More analysis is need to test this hypothesis.

**Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

The range of values for Unsucessful campaigns is 1293. The range of values for Successful Campaigns is 26546. An initial look at range as a measure of variability suggests that variability is more significant in Successful Campaigns than in unsuccessful campaigns.

Additionally, with a mean of 194.4 and standard deviation of 845, successful campaign outcomes based on number of backers appears to have more variability than successful campaign outcomes whose mean is 17.7 and standard deviation of 61.4.

This does make sense. Backer count appears to predict success and project with high backer counts appear to be more successful: the more backers, the greater the likelihood of success. The data is skewed right for both successful and unsuccessful campaigns, indicating that outliers in the data occur when the backer count is higher. The higher backer counts are contributing to the variability of the data.