Katherine Kane

October 2020

Excel Challenge

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

Given the data provided, it is apparent that the success/failure of Kickstarter campaigns varies wildly based on numerous factors. Even within parent categories, campaign sub-category seemed to have an effect on overall success or failure of the campaign. In the film/video category, drama and science fiction campaigns were largely unsuccessful, whereas documentary and television campaigns exceeded their respective goals (Pivot Tables 1 and 2). Additionally, the data showed that goal amount was inversely related with success. Campaign success seemed to decrease as fundraising goal increased (Bonus Chart). Lastly, the data communicated that May was historically the month with the most successful Kickstarter campaigns, with highest number of successful campaigns vs. the monthly total (Pivot Table 3).

1. What are some limitations of this dataset?

Although blurbs were given for each entry, there was no data provided about the merits of each campaign (rating/acclaim, etc.). Therefore, it’s possible that someone could draw incorrect conclusions/assumptions about the reasons for the success and failure of the campaigns.

1. What are some other possible tables and/or graphs that we could create?

With the given data set, we could break down average donation by sub-category to see which projects drew largest donations. We could also look at unsuccessful projects to see if we notice trends in possible factors like campaign length and spotlight.

Bonus – Statistical Analysis

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

The median number summarizes the data more meaningfully. Unsuccessful campaigns had a median number of 4 backers, whereas successful campaigns had a median of 62 backers. I believe that the median value in this instance is more important because it shows that a large number of unsuccessful campaigns had fewer than 4 backers. The mean could be affected by a small number of campaigns with a large number of backers, but the median is more representative of the group as a whole.

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

There is more variability with successful campaigns. I would not expect unsuccessful campaigns to have a higher degree of variability, as they had fewer backers as a general matter (including multiple campaigns with 0 backers). There were large and small successful campaigns, so variability in number of backers makes sense.