Kickstarter Data Findings

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   * According to the pivot table counting categories, it can be concluded that ‘theater’ related campaigns are the most common on Kickstarter, based on this dataset. Journalism related campaigns are by far the least common and have never completed. Music related campaigns seem to be the most successful.
   * According to the pivot table counting sub-categories, plays are the most common campaign and are successful more than half of their campaigns. That said if you want to raise money for a play, Kickstarter is a good way to do it. Something else that I think can be determined here is that campaigns that would be better funded by a VC or a bank tend to fare poorly. Some examples of this are food trucks, video games, wearables, and web categories seem to not do well.
   * The best time to launch a Kickstarter campaign is in the second quarter and the worst time to launch is in the fourth quarter.
2. What are some limitations of this dataset?
   * I think the more complicated the ask, the more prevalent any limitation would be. That said, there seems to be quite a bit of data in this dataset, but this is a list of observations that I see.
   * First off, something that I feel like is missing would be some sort of metric to indicate what the backers get for being backers. It would be really nice to know what the various donation amounts that the campaigns were requesting from backers.
   * The data is between 2009-2017 so anything after 2017 not available and Kickstarter trends could have changed since.
   * When we look at currency, there doesn’t seem to be a means of applying exchange rate. This means that there would need to be some sort of mechanism to convert everything into a common currency to compare monetary metrics across different countries and geographical regions.
   * There is potentially valuable information in the blurbs but trying to parse through that column would be difficult with Excel.
   * When working with the backers\_count column, I needed to add some conditional logic in order to eliminate NULL values on campaigns that don’t have any backers.
3. What are some other possible tables and/or graphs we that we could create?
   * What category or sub-category will generate the highest average donation among backers.
   * What category or sub-category have the highest goal.
   * Judging the efficacy of being a spotlighted campaign or a staff pick with respect to the following items.
     + Does being a staff pick or spotlighted campaign have any effect positive or negative effect on success?
     + Does being a staff pick or spotlighted campaign have an effect on average donation amount?
   * What percent of campaigns happen in the US?
     + What are the top 5 countries that have the most campaigns?

Bonus Statistical Analysis

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.
   1. I would say in this case that the mean is a more meaningful summarization of the data because median seems artificially low. There is a huge variance between the largest and smallest value of the data and it seems to me that even the outer limits of the backer count should be taken into consideration when analyzing successful verses unsuccessful based on backer count in this data set.
2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
   1. There is way more variability with a successful campaign. This absolutely makes sense to me for several reasons.
      1. 60% of completed campaigns are successful. This data set contains a total of 3715 completed campaigns and that means that there are an additional 655 campaigns that can act as data points which makes a difference. So just the total amounts can create variability.
      2. It’s totally possible to have one backer and still be successful. There are several examples of this in the data. At the same time a successful campaign will generally have a lot of backers and the higher the goal, the more the backers.
      3. The smaller a goal is, the fewer backers it requires for success. It just doesn’t take that many people to raise $1000.