**Ida Astaneh 01-Excel-Homework-Challenge**

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Based on the limited data that was made available to us, we can come to the conclusion that while we have the most campaigns within the theater, we will have a larger variance of % successful and failed subcategories.
   2. Journalism was the only category within the dataset to prompt a 100% “cancel” rate, which causes the analyst to assume a combination of limited interest, limited funds, and/or limited backers that could support the campaign.
   3. The technology campaigns had an even spread, creating very little deviation from the median. This shows that the campaigns were “statistically” unreliable to come to any meaningful conclusions about how to best proceed moving forward.
2. What are some limitations of this dataset?
   1. As an analyst, its very difficult to come to solid conclusions when you cannot compare the data points on an even playing field. For example, the country codes were a delimiter because there are several unknown factors that could cause data to be skewed unfavorably (country population, government regime throughout the years, category favoritism as it relates to the people).
      1. Not all currency was converted to USD
      2. Data was scraped from almost ten-year time span
      3. Population size affects backer count understandably
      4. Different country; different campaign focuses
      5. Political countries have differing governments which can control allocation of funds
      6. Unbalanced goals across countries
      7. Limited knowledge of additional columns
         1. How do we categorize a “successful” campaign?
         2. What does “staff pick” entail? Is there bias?
         3. Can we assume “spotlight” means that there is more attention surround the campaign?
3. What are some other possible tables and/or graphs that we could create?
   1. Based off of the information provided, I think it would be beneficial to bring light to the percent funded, backer count, and goal columns to analyze the allocations to determine if the spread of “success” is measured equally amongst the countries. It’s important to always compare like items within your data set to come to meaningful and actionable conclusions.

**Bonus Statistical Analysis**

Within this dataset, its important to separate the failed and successful campaigns to come to a meaningful conclusion. After calculating the statistical factors, as an analyst, one can conclude that the median helps to paint a powerful picture. The successful campaigns have a larger spread (with a max of 26,457 backers), and so the mean tells us that while we have a higher average of backers per campaign, causing the “success” rate to the outkick the number of backers supporting a “failed” campaign.

Successful campaigns: 2,185

Failed campaigns: 1,530

Successful backers: 424,819

Failed backers: 27,096

% of backers supporting a successful campaign: ~.5%

% of backers supporting a failed campaign: ~5%

As far as the variability of the campaigns, again, it is difficult to truly compare as there might be several unknowing outliers within the data set. The standard deviation of the successful campaigns is much larger than the failed campaigns, meaning that the spread of the data is quite larger (and thus further from the mean). The bigger standard deviation of the successful campaigns also tells us that the number of backers per campaign is larger, which can ultimately affect the goal/pledged results.