Madison Mihle

Module 1—Challenge 1

VBA Data Bootcamp

26 November 2024

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   1. Given the provided data, the three conclusions that we can draw about the crowdfunding campaigns are:
      1. We can easily tell which campaigns succeeded and which ones failed. If a pledge goal was set and not met, the campaign was marked as failed. Overall, this could present false data given that there was no universal line between the goals of all pledgers for the campaign.
      2. The dataset proves crowdfunding campaigns are more successful than they are not, providing insight into future campaigns. There is a severe lack of “canceled” and “live” within the outcome, which also goes to show future campaigns that there is potential for success.
      3. The dataset is extremely broad between all categories and countries. This, in turn, could present false data when analyzing because of the variance between countries and if one category is more successful than others. It’s expected that some categories will do better than others in various countries, but when grouped together as a whole, it makes the data skewed.
2. What are some limitations of this dataset?
   1. The limitations of the dataset are that the data is not even across the board, specifically for categories and countries. When the dataset is this broad, the variance and the standard deviation are extremely high, proving that the dataset is inconsistent as a whole. This could present us with false predictions when analyzing trends or forecasting. It could also show that some categories are more successful than others when they are not. Also, another limitation is the broadness of the data. If possible, it would be easier to break down by country and then, if the data allows, look at age group and maybe even location. Lastly, there is no universal set “pledge goal,” which makes the data uneven. If a pledge didn't meet its target, it was labeled as "failed," but every pledger had a different target amount.
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   1. There are more possible tables and graphs that we can create within this dataset, such as whisker plots and scatter plots. These types of plots will allow us to visualize the outliers and the potential trends through trend-lines. A pie chart is another example that we could use when reviewing the data within the different categories and subcategories. It could also be broken down further by taking the different countries and analyzing the categories there and see which ones were more successful and/or failed. This would help narrow down the data instead of having a broad overview.

Statistical Analysis Questions

1. Use your data to determine whether the mean or the median better summarizes the data.
   1. Here, the mean is better to summarize the data. This is because the mean looks at both the "Successful Campaigns" and the "Failed Campaigns" and sees the average of each. Whereas the median just looks at the mid-point of the data.
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. By using the data, I could determine that there was more variability with “Successful Campaigns,” which makes sense considering there are more numbers in the dataset and assuming that there would be more outliers. Also, just visually observing the data through the numbers pulled, "Successful Campaigns" exceeded their goals and again have more variability.