Jose Dominguez

February 14, 2023

**Module 1 Challenge – Crowdfunding Excel Challenge**

**Analysis**

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

* Looking at the pivot table for parent categories, at first glance it can be very simple to conclude that the most successful categories are Film & Video, Theater and Music. But this is not entirely true. If you evaluate the successful number against the total for that category, it will come to light that Technology, Photography and Publishing are more successful than Fil & Video, Theater and Music. It is very important to create different indicators that will make us, analysts, have different perspectives of the data.
  + Evaluating the Outcome Based on Goals, the most successful goal range is between 1000 to 4999, if we are considering only percentages as the most relevant value. But if we look at other ranges that don’t seem successful, such as Goals “Greater or Equal to 50000”, this range is the one that is bringing the most value, money wise.
  + Film & Video, Music and Theater and the parent categories with the greatest number of subcategories. This and revisiting the first conclusion for this question, is why at first glance the analysis can be biased only by perception of percentages if you don’t create enough comparison points or indicators.

1. What are some limitations of this dataset?
   * I think that the real limitations come from not creating enough comparisons between columns. It doesn’t matter how big or small your dataset is, in my opinion, it all comes down to the ability to be able to compare and integrate the different elements that you can find in a dataset.
2. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   * Pie Charts would make a good representation to display and put in perspective how big each of the categories and subcategories are. Also, a table comparing successful and failed campaigns against the total for each of the categories, will highlight truly how successful or a failure a campaign is.

**Statistical Analysis**

1. Use Excel to evaluate the following values for successful campaigns, and then do the same for unsuccessful campaigns:

* Successful Campaigns
  + The mean number of backers: 851.15
  + The median number of backers: 201
  + The minimum number of backers: 16
  + The maximum number of backers: 7295
  + The variance of the number of backers: 1606216.59
  + The standard deviation of the number of backers: 1267.37
* Unsuccessful Campaigns
  + The mean number of backers: 585.62
  + The median number of backers: 114.50
  + The minimum number of backers: 0
  + The maximum number of backers: 6080
  + The variance of the number of backers: 924113.45
  + The standard deviation of the number of backers: 961.31

1. Use your data to determine whether the mean or the median better summarizes the data.
   * Since the mean in both scenarios, Successful and Unsuccessful Campaigns, is greater than the median, the data is skewed to the right. Since it is not a normal distribution, the median is better for skewed number distributions.
2. 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. The standard deviation for the forementioned Campaign is higher than Unsuccessful Campaigns, meaning that the dispersion of the values is higher.
   * It makes sense because Successful Campaigns will have way more bakers count than Unsuccessful Campaigns. For example, the minimum value of bakers count for Unsuccessful bakers count is 0, and for Successful it starts at 16. Looking at the Kurtosis values for both campaigns, even though both are greater than 3, meaning that both will have tails, it is clearly evident that for Unsuccessful Campaigns the tail is more pronounced toward lower values.