Azriel Manuel Tamayo Written Report Excel Challenge / Module 1

One of the main three conclusions that can be drawn from these crowdfunding campaigns is that “Theater” is the most popular parent category with “Film & Video” and “Music” behind in second and third place. For “Theater”, there was a total of 344 campaigns out of the 1000 in total. Out of the 344 campaigns, 187 of them were successful out of the 565 total successful campaigns. In other words, “Theater” had the highest number of successful campaigns out of any category, but also had the highest number of failed campaigns as well. Overall, “Theater” still is the most popular parent category no matter the result and that is the first conclusion that can be drawn from this dataset of crowdfunding campaigns.

The second conclusion that can be drawn from these crowdfunding campaigns is that the Summer season, at least for the months of May, June and July, have the most amount of campaigns compared to earlier months and later months. By checking the “Date Pivot” sheet and the line graph, there is also a huge uptick of successful campaigns as well but a dip for the month of August. This uptick comes with several factors, with some of them being weather and some possible more vacation time for people, although these are all assumptions. The total for the month of January could be an outlier for the Winter season due to the holidays. May had 84 campaigns while June had 86 and July had 93, by far the highest total from any other month. In other words, May, June and July seem to be very popular months for these campaigns to start regardless of their status if the campaign failed, was cancelled or succeeded.

The third and last conclusion that can be drawn from these crowdfunding campaigns is that the United States dominates each parent category in relation to the total amount of crowdfunding campaigns, regardless of its status. Some examples of this can be shown how the United States completely dominates in the “Theater” category with a total of 149 successful campaigns. The country with the greatest number of successful campaigns behind the United States was Italy with only 10 successful campaigns. Another category that is telling with the United States’ dominance was “Film & Video” where they had 76 successful campaigns and the country behind the US at this point was the United Kingdom with only 7.

Some limitations with this dataset are that there may be a chance not all of the crowdfunding campaigns have been recorded. For example, how this data was recorded can come into question and what was the criteria in order for these campaigns to be recorded? I think you can also say that while the United States has the most campaigns compared to the other countries, you may question the fact the United States just favors in every category. That could have been a goal or not but it depends on how this dataset was recorded, especially with how advanced the United States is compared to other countries. I also think whoever created the campaigns may matter as well, as some companies or organizations may have a better chance at succeeding than others.

Another possible table or graph that we may create is the box and whisker plot, or box plot. We can determine any sort of outliers and not just with the United States being the one outlier as we know from the conclusions that the United States dominates in every parent category. We would also determine and compare campaigns from each country, if possible, to see how tight other countries can compare to one another. In other words, see if other countries have one advantage to another as their data pulls close without considering the United States.

Statistical Analysis

Looking at the data of the successful outcomes, I think the mean summarizes that data more than the median. This is mostly due to the fact that by looking at the dataset itself and seeing how many outcomes are over the median of 201 speaks for itself. It is better to look at the mean here and see that a number like 851 will see more success than 201, which is the median, which is a number at least in this sense, not as prevalent to 851, which is the mean. Additionally, the dataset is huge and this specific set skews toward all successful outcomes, which makes the mean more compelling for a success rate. There will also be more backers for a successful outcome than a failed outcome as well. For the failed outcome, the median will be a better. Additionally, failed outcomes result in less interest so seeing a big mean of 585 would not make sense despite how big the dataset is. The median of 114.5 would make sense especially as it is lower than the successful outcome’s median. For this, there would be a less backer count than more, which makes more sense that the median works for the failed outcomes in terms of summarizing.

Regarding if there is more variability with successful or unsuccessful campaigns, you would get a clearer result with the successful campaigns. This is because the successful campaigns have a more better number than some failed campaigns and this is because they were successful and not failed. For failed campaigns, the number would vary and therefore, not be clearer.