Michael Naffziger

Week 1, Excel Challenge Questions

2023 March 23

* Question 1: Given the provided data, what are three conclusions we can draw about crowdfunding campaigns?
  + There is more data originating from the USA (76%) than from the other countries (24%). The remaining six countries represent ca. 4% of the data. This could potentially misrepresent the summary. However, in this case, the overall percentage of successful/failed/cancelled parent category campaigns are similar between the US and the other six countries.
  + Campaigns focused on the performance arts (film/video, music, theater) were consistently in the top 3 for most popular campaign parent categories started, regardless of outcome. From these parent categories, “Plays” remained the most popular sub-category.
  + Between 2010 and 2019, an average of 98 campaigns started. The campaigns from 2020 only included the month of January and were not included in the average since it did not completely represent the data comparison. This is probably due to COVID-19’s global impact. Campaigns started in the summer months had a greater success rate than those started in the winter. An interesting outlier here is the “Photography” parent category, which had a much higher success rate in January and March, perhaps due to taking pictures of snow (Jan) and the arrival of spring (Mar).
* Question 2: What are some limitations of this datasheet?
  + As mentioned above, the data originating from the US outweighs all the other countries combined.
  + Only six countries are reported. Why were these countries chosen for this data set? Though these countries share similar demographics (mostly white, European heritage) it would be a good idea to include a more diverse set of sample countries with similar spending habits as well as a more active user group with crowd funding sites.
  + All the countries are in the northern hemisphere, except Australia. If a start/end date comparison is used, more southern hemisphere countries should be included. This was the pivot tables can be sliced by hemisphere to compare success rates and verify/note if/how seasonality is a valid metric.
  + All currencies are treated empirically. Each campaign funding goals and pledges are considered on an equal basis, i.e., there is no conversion to a universal/common currency.
* Question 3: What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + There are some fields that were not analyzed. It would be interesting to compare the impact of a “Staff Pick” or “Spotlight” on a campaign’s success rate.
  + It would be interesting to compare the origin of the backer to the geographic location of the campaign. If the IP of the backer is not available, it might be possible to check financial records to determine if the backer’s payment currency had to be exchanged before the payment went through. This way it could measure how important local backers are versus international backers. However, large countries such as AU and US would benefit from IP location trackers due to the population being spread across a larger area.

Bonus Sections

* Question 1: Does the mean or median summarize the data more meaningfully?
  + There is a large difference between the mean and median across each of the countries: the mean and median tell different stories.
    - The median is consistently lower than the average, regardless of the outcome. This suggests that there are many campaigns with a low number of backers. When comparing the success rate, the median is roughly the same. Which might provide some insight into the overall view traffic campaigns receive.
    - The average is skewed higher probably due to the popularity of a campaign. This might be due to the description, public perception/appeal, and novelty of the product to the users that browse crowd funding sites. With a successful marketing campaign, a product receives more traffic leading to more users to pledge.
  + In this case, I would lean towards the average to summarize the data more meaningfully. Since the campaign’s organizers will want a metric of how many backers are needed to pledge to be successful, they can use the average of successful campaigns and an early benchmark/target. This is also beneficial since the average of successful campaign backers is much larger than the average of failed campaign backers. This choice is also supported by the median and average of the Average Donation (regardless of outcome) are very similar.
* Question 2: Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or why not?
  + There is a larger variability with successful campaigns.
  + Yes, this makes sense to me because it likely represents how a campaign’s funding goal is defined and how it is described to potential backers. The larger the variance suggests a successful campaign can get the interest of a larger group of people, i.e., the larger the net the more fish you could catch.
  + In the wine industry, it is often said that the first bottle of wine is easy to sell; the second bottle is the hardest. This is due to a consumer’s ability to connect with a particular bottle of wine, remember that connection, and overcome all other market influences to buy the same bottle again. There are a vast number of wines available on the market- similar to the vast number of funding campaigns on a crowdfunding site. The campaign must be able to reach and resonate with a large diversity of people, AND be memorable/important enough to tell a friend to donate as well. It is a numbers game- the more $10 donations you have, the less you need a $1000 donation.