University Of Minnesota

Data Visualization and Analytics

Wanderson Oliveira

Date: 12/12/2023

Excel Module 1 Challenge:

Question:

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

Answer:

* Looking at the data collected, category of Theater was the most successful, but also presented the most failure rate amongst other campaign projects. Let alone Plays sub-category, which also had higher rate of failure and success.
* When totalizing Failure, Live and canceled, 43.5% didn’t met their goal for Crowdfunding and 56.5% succeeded on their goal.

* Overall, crowdfund campaign are more Successful than not successful based on the results.

Question:

**What are some limitations of this dataset?**

Answer:

* One of the limitations is that the data collected can’t specifically tell which factors contributed for failure or success for each category and sub- category.
* Another limitation is that data can’t show if there is any correlation between rates of failure or success over the months. Example: Overall, July was the month with the most successful crowdfund campaigns. August and May in General had the highest Crowdfund failure and August was the highest for Crowdfund cancelation.

Question:

**What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

Answer:

* Another Graphic that we could utilize is the Pie chart and show the proportion of the crowdfund campaign on how much percentage was Canceled, failed, live or successful per country or in General. Or per category.
* For each category there was a Date started and a date that ended. Does this timeframe contribute to failure or Success? Could create a graphic with correlation between time frame per each category and results, that way, overtime, based on the trend observed expectations could be created when planning ahead for future crowdfund campaigns.
* What Improvements in the crowdfund campaign could have been done moving forward in order to achieve more success rates in other categories such as ‘’mobile games’’ for example, where it had more failure than Success? That being said a Pivot table and graphic over the years to set goal Expectations.

Question:

**Use your data to determine whether the mean or the median better summarizes the data.**

Answer:

* In both situations (Successful and Unsuccessful) the MEDIAN better summarizes the data. When looking at the median, it’s value is very far (lower than the Mean) this shows us an asymmetrical distribution, it happens because there are extremely high backer counts that are bringing the average up, creating a positively skewed graphic (Mean >Median> Mode). One example of that is the category Music/Electric music, where it had a high volume of backers count as we can observe in the maximum Value.

Question:

**Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

Answer:

* There is more variability in the successful campaign, this is reflected on the results of variance calculated for successful backers count which is higher than the variance in unsuccessful campaigns. This does make sense because the dataset in the successful campaign is bigger, it presents a higher variance and more deviation in the dataset.