Sam Robinson

7/18/2021

Excel Challenge: Kickstarter Data Analysis

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

The highest number of successful campaigns lie in the theater category (839) while the highest percentage of successful campaigns lie in the music field (540/700 = 77%).

The highest count of subcategory successful campaigns within the theater category are plays at 694 (65% of total plays), while the highest count of subcategory successful campaigns within the music category is rock at 260 (100% of total rock, meaning that all of the rock campaigns were successful).

The highest number of successful campaigns for theater were created in May at 111 count while the highest number of successful campaigns for music were created in May at 54 count.

1. What are some limitations of this dataset?

On the whole, the amount of data in the data set is very low, however in some categories such as Journalism, it is far too low to make any strong conviction conclusions. Ideally, we would want much more data to ensure that our conclusions are backed by statistical accuracy.

This causes even more issues when we break the data down by Category, Sub Category, etc. because in each smaller segment that we take, our data set becomes even smaller making any conclusions that we draw from the smaller segments even more error prone.

1. What are some other possible tables and/or graphs that we could create?

We could take a look at States vs. Country to get a better understanding of which countries have the most successful campaigns.

We can also take a look at States vs. Country and sort by Total Amount pledged to see where the biggest amount of pledge dollars is being allocated to which Category and Sub Categories.

We can also take a look at the average donation size to the pledged amount by Category and Sub Category to get a better understanding of how many people are supporting which campaigns and in what amounts are they pledging. For example, some campaigns may have many small donors while others have very few but large donor support. For the purposes of marketing your campaign, this is helpful as it helps to clarify and focus marketing efforts.

Bonus Statistical Analysis

(these answers are also provided next to the summary table in the Excel file)

Question: Does the mean or the median summarize the data more meaningfully?

Answer: The median summarizes the data more meaningfully in this case because the standard deviation is much higher than both the successful and failed campaign means. (For successful campaigns, the Std Dev / Mean = 4.3 while for failed campaigns Std Dev / Mean = 3.5)

If the Std Dev were less than the mean, then it would give us an indication that most of the data points lie around the mean. In this case, there is a lot of variability in the data and there are some large numbers which are skewing the mean data to the upside.

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 with successful campaigns. This makes sense when we consider the effect of successful marketing campaigns that take hold.

It comes down to mind share, once a campaign starts to take root, it usually spreads by word of mouth

particularly in events among the arts like theater and music where people will spread the word of the events and more and more people will want to participate in those events.

It's a cascading social phenomena effect where people want to be with other people and be part of the fun and popular event.