Excel Challenge

Cesar Zea

June 11th, 2023

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

**Answer:**

* While Entertainment, particularly Film, Video, and Theater, represents the largest number of projects, the most successful category is technology, with a success rate of 67%. Within the technology category, web-focused projects have shown the highest success rate.
* In terms of the months that show a higher rate of success, the trends show that creating a crowdfunding project in June and July could improve the rate of success. However, the success rate varies significantly depending on the category. For entertainment categories like film, music, and theatre, June and July demonstrate an improved success rate, whereas for technology projects, the best period seems to be from April to June. It is essential to understand the project type and assess whether there is a seasonal connection to the project's specifics and the budgetary periods of potential backers.
* The goals analysis shows 2 areas of heavy activity: below 10K and more than 50K. As expected, the projects with requirements below 10K show a higher rate of success. Furthermore, when considering projects with funding goals between 10K and 49.9K, this group also exhibits a high success rate. However, due to the limited number of cases available, it is challenging to delve deeper into the funding levels that contribute to this success.

*What are some limitations of this dataset?*

**Answer:**

While the data set offers valuable insights into high-level crowdfunding results, the limited number of cases prevents a more in-depth analysis of details. For example, evaluating subcategories, month-wise trends, or success rates based on goals becomes challenging due to the insufficient sample size.

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

**Answer:**

* I created a table of proportions of success by category to understand the type of projects with a higher level of success. The same could be done by subcategory if the number of cases was larger.
* I also created box and whisker graphics for successful and failed projects to better understand what drives success in terms of the number of backers and the distribution of the data. The same could be done for levels support to see where the sweet spot is for funding for the different categories.

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

**Answer:**

The median seems to be a better score to summarize the data, because it skews to the left, for number of backers and levels of funding. There are also many outliers at the end of the curve to the right.

*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 greater variability in the number of backers for successful campaigns, as indicated by the higher standard deviation. This is expected due to the presence of numerous outliers, likely influenced by the perceived positive outcomes of these projects and recommendations from investors, which attract a larger audience to projects estimated to have a higher chance of success in the market.