Module 1 Challenge Report

Alyssa Hondrade

# Analysis

**Question 1** – Given the provided data, what are the three conclusions that we can draw about crowdfunding campaigns?

1. The top three categories with the most campaigns, in order of highest to lowest, are: theatre, film & video, and music (Figure 1). Although “journalism” had the least number of campaigns, with a total of 4, all “journalism” campaigns were successful.

A graph with green and red bars

Description automatically generated

Figure : Pivot Table 1 - Outcomes based on Category.

1. When the campaigns are considered by sub-category, “plays” had the highest number (total of 344), making up almost a third of all campaigns (Figure 2). “Rock” followed with 85 campaigns, then “documentary” with 60, and “web” with 51. The remaining sub-categories all had less than 50 campaigns each.
2. Overall, there are more successful campaigns than failures and cancellations over the course of 2010 to 2020 (Figure 3). The month of July had the highest number of successful campaigns – that is, campaigns started in July. April had the least cancellations, of 1 campaign, compared to a peak of 8 in August (note this is a total over the decade). In the month of August, there were nearly as many successful campaigns, 41, as there were failed, 35.

A graph with red and green squares

Description automatically generated

Figure : Pivot Table 2 - Outcomes based on Sub-Category.

A graph of a number of outcomes

Description automatically generated

Figure : Pivot Table 3 - Outcomes based on Month.

**Question 2** – What are some limitations of this dataset?

To receive funding, the project must meet or exceed an initial goal, so many organisations dedicate considerable resources looking through old projects in an attempt to discover “the trick” to finding success.

1. The definition of “success” in the given problem is whether the campaign met their “goal”. This is reasonable, as backers’ donations are often returned if the goal is not met. This assumes that the campaigns have chosen the correct amount for their goal. There is a possibility that campaigns incorrectly gauge their goal and fail as a result of not acquiring enough backers, even if their project might succeed given a lower goal amount.
2. The dataset is a sample of 1,000 projects. For the data to be valid, there is an assumption that this sample is representative of the population. It is also not clear whether the data was collected from a single crowdfunding platform or combined from multiple sources.
3. The categories, and subsequent sub-categories, available in the dataset is relatively limited compared to the variety available on crowdfunding platforms. A quick search on the internet identifies the following categories not included, as an example: Kickstarter (arts, crafts, comics, dance, design, fashion), Indiegogo (community projects, transportation, travel). This can be problematic if an analysis is conducted on which category or sub-category might be most successful, given historic data.
4. The data in the spotlight column can have inconsistent meaning, as a “spotlight” is defined differently on different platforms. For example, in Kickstarter, the spotlight page is only available for “successfully funded creators” [1] whereas Crowdfundr requires an opt-in to their feature (with unclear requirements) [2], and Indiegogo is at editorial discretion [3].
5. No data on other efforts or strategy employed by each campaign. For example, whether they advertised on social media, pitched videos, utilised professional graphics to improve campaign and marketing materials, etc [4]

In conclusion, given limitations on data availability and collection, the dataset to answer the question of determining “the trick” should identify data where the campaign might have influence or control of the outcome.

**Question 3** – What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

1. Ratio of success to failure, by sub-category, and plotted as a descending bar chart. This would give a ratio of success depending on the chosen sub-category. However, care should be taken for sub-categories with a small sample size, for example with “journalism” having a 100% success rate over a total of 4 campaigns.
2. A plot of success vs “staff pick” to answer the question: “Does making the staff pick list increase the likelihood of a campaign’s success?”. This would be useful as the campaign has a reasonable level of influence, depending on the time and effort exerted in providing a quality submission for staff pick. Applying for the staff pick could be a worthwhile activity if there is evidence picked campaigns are more likely to be successful.

# Bonus Statistical Analysis

**Question 1** – Use your data to determine whether the mean or the median better summarises the data.

The mean better summarises the data as it accounts for the outliers better than the median. The statement “successful campaigns had a mean of 851 backers while unsuccessful campaigns had a mean of 585” is more insightful as it demonstrates two points:

1. Successful campaigns have more backers than unsuccessful campaigns.
2. Campaigns can still be unsuccessful even with approximately 585 backers.

If median is used, the second point would not be as impactful as the values are lower and closer (201 for successful versus 114 for unsuccessful).

**Question 2** – Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variability in successful campaigns that unsuccessful campaigns.

This makes sense as the definition of success in the given problem is whether the project “met or exceeded the initial goal”, hence there can be more variations of what “success” looks like compared to failure. The more backers, the higher the likelihood of success, meaning there is no upper limit per se, resulting in greater variability. Unsuccessful campaigns on the other hand, tend to have less backers, hence less variability as the data is “constrained” to relatively low values.

# References

|  |  |
| --- | --- |
| [1] | Kickstarter, "Introducing Spotlight," Kickstarter, 28 Mar 2015. [Online]. Available: https://www.kickstarter.com/spotlight. [Accessed 25 Jul 2023]. |
| [2] | Crowdfundr, "Crowdfunder Spotlights," Crowdfundr, 24 Jun 2022. [Online]. Available: https://crowdfundr.com/spotlight/. [Accessed 25 Jul 2023]. |
| [3] | Indiegogo, "Will Indiegogo feature my campaign?," Indiegogo, 5 Mar 2014. [Online]. Available: https://support.indiegogo.com/hc/en-us/articles/527466-Will-Indiegogo-feature-my-campaign-. [Accessed 25 Jul 2023]. |
| [4] | Indiegogo, "7 Free Hacks for Any Crowdfunding Campaign," Indiegogo, 3 Nov 2015. [Online]. Available: https://go.indiegogo.com/blog/2015/11/free-crowdfunding-hacks.html. [Accessed 25 Jul 2023]. |