Steve Bonillas

Data Bootcamp: Module 1

8/6/23

**Kickstarter Campaign Analysis**

Introduction

In this module, we were presented with an Excel spreadsheet containing a dataset relating to one thousand Kickstarter campaigns from 2010 to 2020. The data included the campaign names, success or failure, monetary goals, how much was pledged, the number of backers, start and end dates, and the categories each campaign fit into. We were asked to clean the data and present it in charts and pivot tables. What follows are some of the conclusions I was able to draw from the data and my recommendations for improving upon what was requested of us.

Conclusions

1. The campaign category and sub-category most represented in our data was Theatre/Plays. Of the 1000 Kickstarter campaigns in the dataset, 344 were related to the theatre. This category had more successful campaigns and more failed campaigns than any other category provided. This is seen most effectively in the tables *Campaign Results by Category* and *Campaign Results by Sub-Category*.
2. Referring to the *Campaign Outcomes Over Time* table, when normalized over 10 years, it appears that June and July are the best months to hold a Kickstarter campaign. We see a rise in the success rate of campaigns, while the failure rate remains stable. If we combine this with the first finding that most Kickstarter campaigns are for theatre projects, this aligns with funding deadlines for some national grant programs. The Broadway Cares National Grant program requires applications for funding be submitted in May, with funds being distributed in late June.[[1]](#footnote-1)
3. If you are interested in running a Kickstarter, it looks like it is best to keep your expectations modest. Looking at the *Campaign Outcomes Compared to Initial Goal* chart, we see that most successful campaigns set a goal of $1,000 to $4,999 (191), with a very low failure percentage of 16.45% at that level. There were also a lot of successful campaigns from $5,000 to $9,999 (164) and Greater $50,000 (114), but failure rates begin to rise – 40% and 53% respectively.

Limitations and Suggestions

1. My first suggestion relates to the *Campaign Outcomes Over Time* chart. The initial request was to show the number of campaigns (successful, failed, cancelled, and live) on a line chart, showing the months that these campaigns were completed. However, the dataset contains over 10 years of data. The request is fine for making a normalization of the data and showing seasonal trends, but it doesn’t account for year-to-year differences and snapshotting data for specific dates. My solution for this, which keeps the initial ask intact, is the addition of a Timeline Slicer. With the Timelines Slicer, we can view the data at specific times of our choosing, rather than just a normalized trend. I just prefer slicers to filters in general, and I added them to each of my charts.
2. On the *Statistical Analysis* tab, we were asked to show the Median, Mean, Min, Max, Variance, and Standard Deviation for the Successful and Failed campaigns. However, a few things could have been added here to make the analysis more useful.
   1. For my Analysis page, I included the Campaign IDs for the Successful and Failed campaigns, just to ensure that the data related to something tangible. Just having an outcome next to the backer count doesn’t really mean a whole lot, unless it is rooted to an actual campaign.
   2. Each of the outcomes would have benefited from a count of the sample size, just to make it plain the numbers that were being dealt with. A simple =COUNTA formula of each campaign outcome would have told you a little bit more about the data.
   3. We were asked to do separate evaluations of the Successful and Failed campaigns, but doing a third evaluation, with the combined numbers would help to illustrate the outlying data in the set and tell us which dataset was more representative of the whole.
   4. I also found it helpful to use the =ABS formula to show the differences between the Successful and Failed Campaigns.
3. Another thing that may have been interesting to add to either the Category / Sub-Category charts or the Outcome Over Time chart might have been adding slicers for the Spotlight and/or Staff Pick columns, to see if having that extra (possibly paid) attention on your campaign would be worth it. I checked it, and it really didn’t look like the “spotlight” did much in translating to additional success numbers.
4. For the Category and Sub-Category tables, we were asked to have the chart be able to be filtered by country. If that is of interest to our hypothetical client / employer, then that information should probably also be on the *Campaign Outcomes Over Time* tables as well.

Overall, the module was edifying and gave me a better understanding of how to think about and present data. It even got me second guessing some of the ways I had been working in the past. I hope that my findings and suggestions are what was expected.

1. Broadway Cares, National Grants: https://broadwaycares.org/national-grants/ [↑](#footnote-ref-1)