Conclusion report on the charts of Crowdfunding data

Mason Seifaddini

Challenge 1

UWA Data Analysis bootcamp

March 2024

Crowdfunding data is showing 1000 records of successful and failed events along with the dates, goals, category, subcategory and many more. Our goal is to analyze what factors affected the events to be successful or failed.

**What are three conclusions that we can draw about crowdfunding campaigns?**

**Conclusion 1:** According to the bar chart above we can see that campaigns held under Publishing has 100% chance of success following with Technology that has 80% success rate. While we have the most number of campaigns held in Theater, the success rate is well below 50%. Also, we can see that Food and Photography had zero chance of success for the events taken under them. Music with 65% success rate should also be considered as a good category for fund raising events after Technology.

**Conclusion** **2:** Considering the subcategories we can easily realize that 34% of the fund raising events are happening under Plays with a 55% rate of success which is really important. Othe subcategories with a high rate of success can be seen as Rock, Documentary, Animation, Web, Food trucks and Drama.

**Conclusion 3**: While we can see a slight rise in the number of successful campaigns in Jun, there can’t be seen any significant rise or dive on other months of the year. So, we can say that the flat chart is showing no significant relation between the month of the event and the success rate of them.

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

Although the dataset provides us with some useful data about the campaigns, it lacks some important information that could potentially lead to better analyzing and conclusions about it. We need to know how they campaigns were advertised. Different methods of advertisement is very important to encourage people to participate. Also, the cities that the campaigns were held in can play a role which is missing from the dataset. In addition to that, the key persons who were supporting the campaigns in the process of advertisements and holding them can be effective in success rate and we don’t have any information about them in the current dataset.

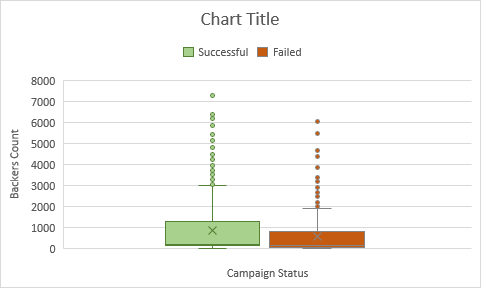
**What are some other possible tables or graphs that we could create and what additional value would they provide?**

Firstly, we can create a pivot table and categorize our data based on the country that the campaigns are held in. This can tell us about which country had the most success rate in campaigns. Also, we could categorize the date based on Spotlight and Staff\_pick to see how they influence the success rate and see if we could find any pattern there.

#### Bonus Statistical Analysis

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

|  |  |  |
| --- | --- | --- |
| **Statistics** | **Successful** | **Failed** |
| Mean | 851 | 586 |
| Median | 201 | 115 |
| Minimum | 16 | 0 |
| Mmaximum | 7295 | 6080 |
| Variance | 1603374 | 921575 |
| Deviation | 1266 | 960 |



The chart above clearly shows the data has a large number of outliers and they have significantly skewed the mean to the right. So, I would consider median to describe the tendency. . According to the median, the number of backers have a positive effect on the success of a campaign. We can see that the median for successful campaigns is nearly double comparing with the median for failed event.

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

According to variance and deviation, it is clear that successful events have more variability. I think the reason is that we have more outlier numbers in successful campaigns and they can cause the spread out from mean and median.