Excel Challenge: Crowdfunding Data Analysis

Data PT West December 120224

Megan Iyer

12/11/2024

Given the data provided and graphs I created; I have drawn 3 conclusions of the crowdfunding campaigns:

1. Theater, music, film and video are the most successfully funded projects, as demonstrated clearly on the “Category Stacked Pivot Chart”.
2. June and July are the best time of year to start a crowdfunding campaign, since they have the greatest number of successfully funded campaigns, based off the pivot line chart.
3. While plays have the most total number of successful campaigns in the subcategories, their success rate is not the highest. You would get a higher rate of success in a smaller niche like television despite it having fewer campaigns overall.

Some limitations of the data are:

1. It doesn’t give information into what events were happening during the different years. For example, changes in the economy, current events, and trends.
2. Doesn’t give insight to the demographics of the donors contributing to the campaigns. This would clarify why some are more successful than others.
3. Categories are not equally divided into subcategories. For example, the category Theater only has plays as the subcategory. If it was more detailed, we would have more insight into which plays are more likely to be funded versus fail.
4. You can’t tell from the data the geographic location of the donors. I know it shows different countries, but it would have been more helpful to focus on one country, then divide it into regions. Different regions of the US have different preferences and interests.
5. The data doesn’t tell you if this is from one platform or from different sites. This could affect the demographics shown if all the data is just one platform. Donors may prefer one platform than others.

Other possible tables or graphs to create are:

1. Line chart comparing the outcomes to years, to be able to reference trends or current events that might affect funding.
2. Scatter plot comparing backers count to percent funded. This might give insight if more backers equal better success rate. This could show the need for more promotion or social engagement to improve outcomes.
3. Compare the average donation amount by outcome in a bar chart to see if higher average donations correlate with a successful outcome.
4. Bar chart comparing the duration of the campaign to the success rate to help optimize campaigns for success.

**Statistical Analysis:**

For both the successful and failed campaigns, the median calculation better summarizes the data. This is true because both data sets contain outliers that skew the data. The median is less affected by the outliers than the mean, which is evident by the big difference between median and mean vaules. The high standard deviation for both data sets suggest that the data is spread out, confirming the need to use median over mean.

The data shows that there is more variability with successful campaigns versus the failed campaigns. This makes sense to me because successful campaigns can include popular, trending campaigns and normally funded campaigns. If a campaign gains popularity, it will have way more backers than usual, skewing the numbers.

I could easily tell that the successful campaigns had more variability from comparing the variance number and standard deviation to the failed campaign. Both the variance and standard deviation were higher in the successful campaigns, suggesting more variability.