Letha Varughese

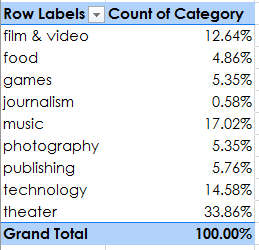
HW 1

01/04/2021 **Analysis of Data**

Analyzed the data using the Excel table to detect any patterns within the data to predict success or failure. A pivot table was used to create the tables and charts.

There have been 4113 projects included in this dataset. Among these projects, 34% are

Theater. Other major projects are music (17%), technology (15%), and film & video (13%).



Among all projects, about 53% succeed in meeting their goal, 37% failed, and 10% are cancelled or still live.

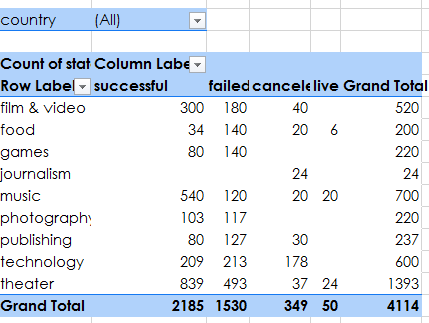
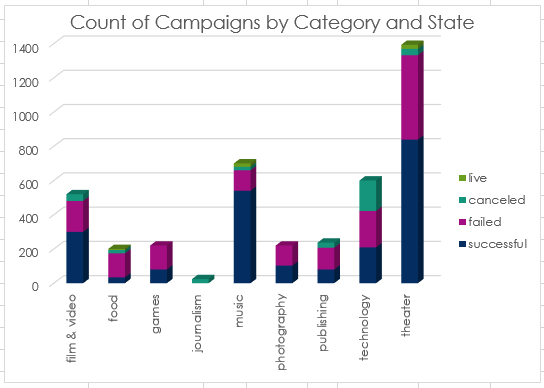


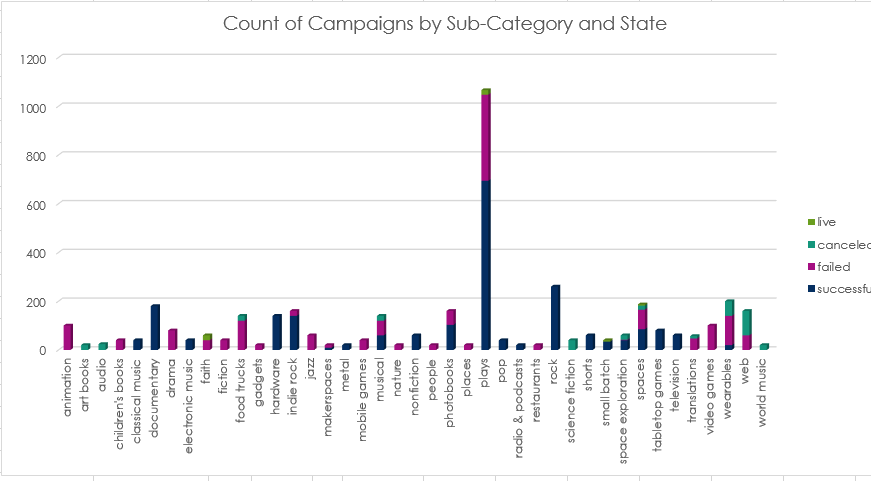
Chart A

Shows all campaigns and state by category. This counts the number of campaigns that were successful, failed, canceled, and those currently live.



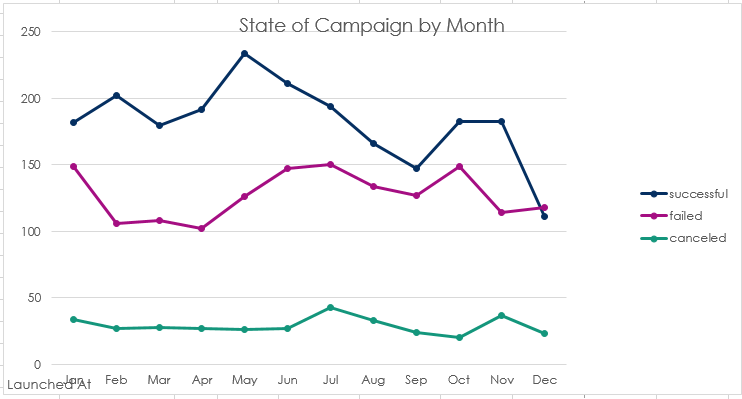
**The most successful campaigns start at theater, then music, followed by film & video.**

Chart B



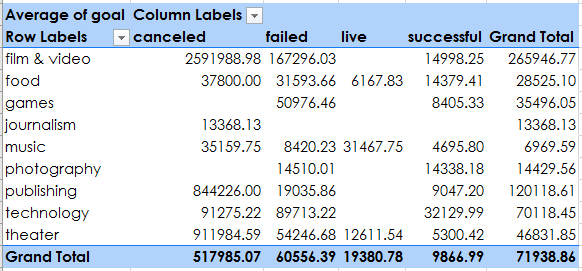
In Chart B, sorted by all counties the sub-category **plays** from category theater top the chart. The music sub-categories **rock** and **indie rock** are the most successful, and **documentaries** are the most successful in the sub-category of film & video.

Chart C



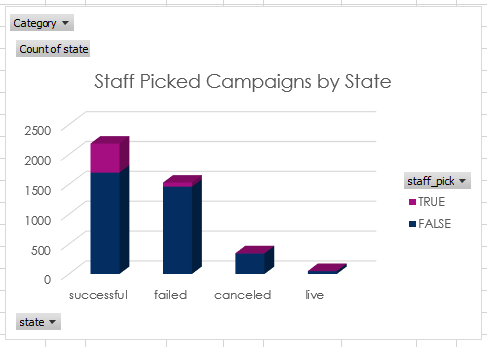
The most successful campaigns begin in May, whereas the least successful campaigns seen in December.

Average Goal Amount Sought by Successful and Unsuccessful



Overall, the average successful project has a goal that is about 50% lower than the average unsuccessful project.

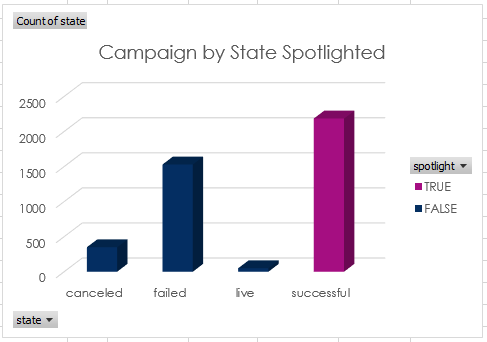
Chart D



|  |  |  |  |
| --- | --- | --- | --- |
| **State** | **FALSE** | **TRUE** | **Total** |
| successful | 1699 | 486 | 2185 |
| failed | 1464 | 66 | 1530 |
| canceled | 345 | 4 | 349 |
| live | 49 | 1 | 50 |
| **Grand Total** | **3557** | **557** | **4114** |

Campaigns picked by staff members had no influence in success.

Chart E



|  |  |  |  |
| --- | --- | --- | --- |
| State | FALSE | TRUE | Total |
| canceled | 349 |  | 349 |
| failed | 1530 |  | 1530 |
| live | 50 |  | 50 |
| successful |  | 2185 | 2185 |
| Grand Total | 1929 | 2185 | 4114 |
|  |  |  |  |

There is 100% success in campaigns that are spotlighted.

Conclusion

Among all projects, about 53% succeed in meeting their goal, 37% failed, and 10% are cancelled or still live.

There are more project for category theater; and within theater, sub-category ‘Plays’ has the highest number of successful projects.

This data suggests that smaller goals are often better. There are, however, some simple things can do to increase the success rate:

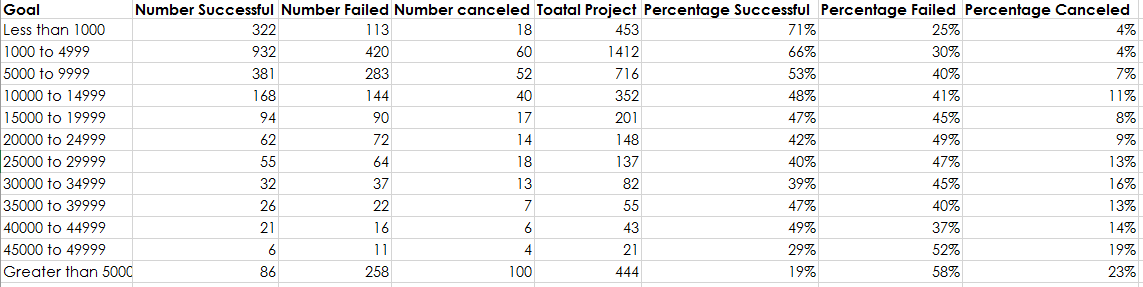
The spotlighted campaigns were 100% success, but campaigns picked by staff members had no influence in success. The most successful campaigns are happened in May. The success rate declines towards the end of the year.

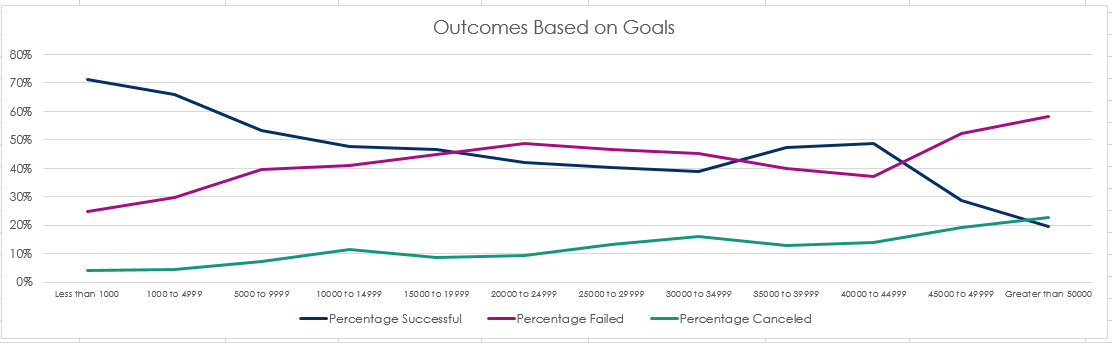
As a limitation in the dataset, I observed that most of the entries are from US. There is not enough data from other countries to compare the success rate in US. There are no details about the backers in the dataset, so it is hard to analyze the influence of backers in the success rate. Other than this, there is no limitations noticed in this dataset.

**Other possible tables and/or graphs that we could create**

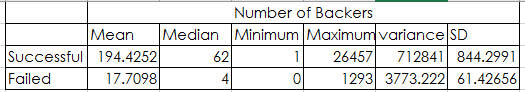
* We could do a graph showing the relationship between success of a project and the number of backers.
* We could do a table and graph for the count of the state based on the length between the date created and the date ended.

**Bonus**





**Bonus Statistical Analysis**



The mean of the number of backers in successful campaign is 194 with a median of 62, whereas, the mean in failed caapaigns is 18 with a median number of 4.

In both cases, the mean and median are considerably different, this indicates that the the data is not normally distributed (the data are skewed) and the median gives a more appropriate idea of the data.

The high standard deviation indicates that the data points are spread out over a large range of values.