**Crowdfunding Trend Analysis**

Crowdfunding platform *Background*: Crowdfunding platforms are becoming popular for funding new projects for upcoming businesses in various fields. With the help of Excel functions, my conclusions on the crowdfunding campaigns are as below.

Crowdfunding main data outlook:

A screenshot of a computer

Description automatically generated

Here with the help of formula (=column E/Column D \* 100) I have found out that the percent funded to projects and added the column to get the outcome to see if a project is successful, failed, canceled or live. These have been conditionally formatted to color code visual presentation of the data in the worksheet with information readily. With the backers count column I was able to find the average donation that was made.

Analysis by Launch date outcome:

* *The first conclusion* I am making by looking at the graph made from launch date is, projects found success that were started in the month of June and July as compared to other months.

A graph of a graph showing the growth of a company

Description automatically generated with medium confidence

* By saying that, I looked at the failure rates as well. The numbers are relatively low in June and July when compared to the other months.
* launched\_at column used Unix timestamps that were converted with the formula provided to format them to standard date format.
* Pivot table was used to draw information from main data to help analyze the outcome better.

A screenshot of a spreadsheet

Description automatically generated

Analysis by Category and Sub-Category:

* *The second conclusion* that I am making is by analyzing the data from Category and sub-category. We can see here; the tables show number of projects that have been successful are from Theater/ Play.

A screenshot of a spreadsheet

Description automatically generated

* The graph says the category Theater/Plays which has seen more success has also seen more failure.
* Here I am utilizing stacked-column pivot chart. The color coding in the bar is to make it easier to show how many successful category and sub-categories are present.

A screen shot of a graph

Description automatically generated

A screenshot of a graph

Description automatically generated

### Analysis of Outcomes based on Goals:

* My third conclusion comes from analyzing the goal outcome from the main data. To do so, I have used CountIfs function on the goal column.
* 100% success rate was found in the range 15000 to 34,999.

A table with numbers and symbols

Description automatically generated

* We can see failure rate is higher in the range 5000 to 15000.
* The projects that saw the highest success rate had goals between 15000 and 25000.

A graph showing the results of a successful business

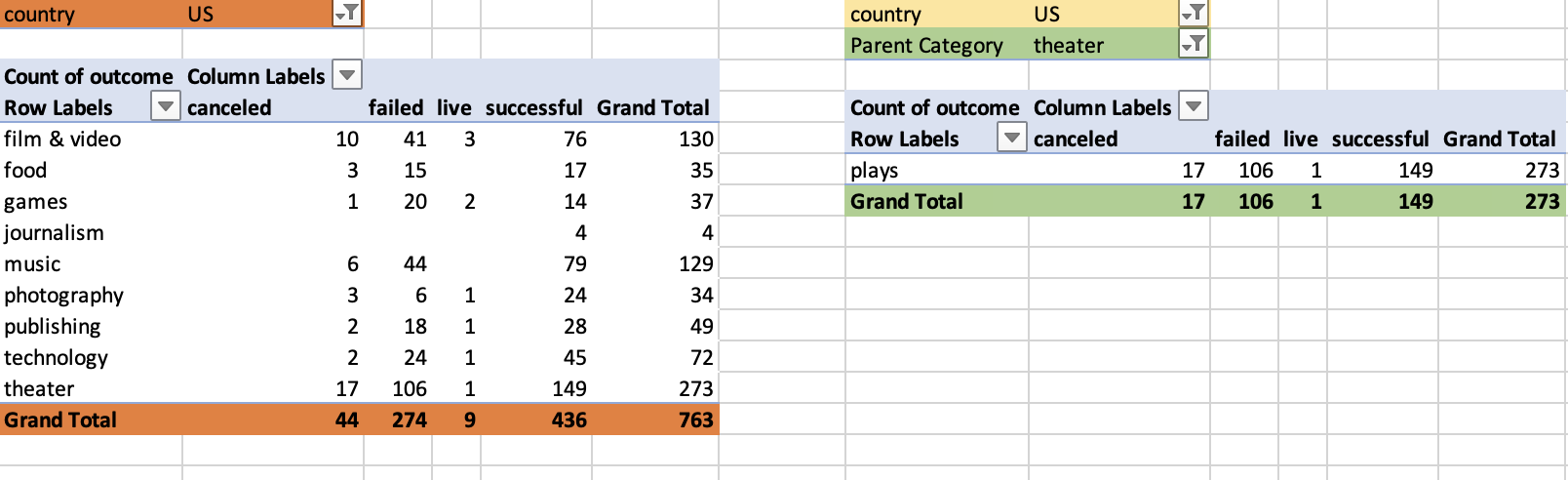
Description automatically generated with medium confidence

Limitations of this data:

* Insufficient data of the backers’ country of origin. With the country of origin, we can analyze the correlation between the country of origin and the interest in the category.
* Number of times Project initiators have reached the goal successfully before. This will tell us if they are first time or experienced.
* The sample data set size may not be representative of the raw data. The sample set may be too big or too small.
* How much each backer contributed to each project and did they contribute to multiple projects.
* Was there any marketing campaign done? To identity the correlation between marketing and success rate.
* There is no clarity how the staff pick and spotlight columns influences the project.

Additional values and graphs:

* We can filter the category and sub-category pivot tables to obtain data only for the selected country.



* Years column in the Launch date outcome tab can be filtered to particular year and watch the trend.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parent Category | theater |  |  |  |
| Years | 2018 |  |  |  |
|  |  |  |  |  |
| **Count of outcome** | **Column Labels** |  |  |  |
| **Row Labels** | **canceled** | **failed** | **successful** | **Grand Total** |
| Jan |  | 1 | 2 | 3 |
| Feb | 1 | 1 | 1 | 3 |
| Mar |  | 1 | 1 | 2 |
| Apr |  |  | 2 | 2 |
| Jun |  | 3 | 3 | 6 |
| Jul | 1 |  | 2 | 3 |
| Aug |  | 1 | 1 | 2 |
| Sep | 1 | 1 | 5 | 7 |
| Oct |  |  | 2 | 2 |
| Nov |  | 1 | 1 | 2 |
| Dec |  | 2 |  | 2 |
| **Grand Total** | **3** | **11** | **20** | **34** |

#### Statistical Analysis:

* In the given data of 1000 crowdfunding example, mean is higher than the median. Data is skewed towards right. Median is the correct measurement to consider when the values are extremely high.
* When considering variance for successful and failure campaigns, we are getting to see the number being very high. Therefore, confidently arriving at conclusions will be hard.