Crowdfunding Data Analysis

# Peregrin Ryan

## Given the provided data, what are three conclusions we can draw about crowdfunding campaigns?

The first conclusion I believe can be made from the data is that utilizing crowdfunding as a means of financing a project is highly volatile. While most crowdfunding campaigns are in fact successful making up 56.5% of results of a crowdfunding campaign, there is still a high failure rate at 36.4%.

The second conclusion is that the most popular category of project to utilize crowd funding is ‘theatre’ making up 34.4% of all crowd funding campaigns. Almost doubling the second most popular crowd funding category of ‘film and video’, which makes up 17.7% of crowdfunding campaigns. This also causes a skew within the sub-categories as ‘plays’ make up 34.4% of all sub-categories with the next largest, ‘rock’ making up 8.5% of all sub-categories.

The third is that the most volatile time of year to start a crowdfunding project is in quarter three. Projects that begin in July statistically have the highest success rate of the year at 58%. However, August has the lowest success rate for crowdfunding projects at 41%. Which means that the most successful and least successful months are next to each other during quarter 3.

## What are some limitations of this dataset?

There are a few limitations with this data set which limit the possible analysis that can be made from this data set. The first and most clear is the limited sample size of data and where it is taken from. The overall data covers only 1000 crowdfunding campaigns with only roughly 100 per year. Making analysis comparing different trends from different years difficult. The data is also geographically limited as it is only taken from North America, Europe and Australia. With the United States of America being the largest country making up 76.3% of all data. This makes the data limited when comparing global trends or popularity.

The data is also limited in what sub-categories are listed. As the significantly largest category of campaign is ‘theatre’ making up 34.4% of all campaigns yet it only has one sub-category being ‘plays’. This creates a large outlier within the sub-categories as it takes up 34.4% of all sub-categories. In comparison the second largest category, music has 6 subcategories of “electric music”, “indie rock”, “jazz”, “metal”, “rock”, “world music”. This limits the analysis possible for trends in the largest category of projects utilizing crowdfunding campaigns.

Thus, the two largest limitations of the dataset provided are: The sample size being limited in both region and amount over time. As well as limited data of the most popular type of project to utilize crowd funding.

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

Some other possible data visualisations and analysis that could be done with this data set may be:

* A pie chart measuring the most popular Parent categories or sub-categories. As it will provide clear data visualisations of what projects utilize crowd funding.
* What categories receive the highest percent funding or highest average donation a pivot chart and visualised with a column chart. Showing what campaigns are most likely to meet or exceed their funding goal.
* A similar chart to the ‘Date Pivot Table’ provided. However, it measures the outcomes compared to the deadline date of each campaign. This would provide data on the outcomes at the end of a projects funding campaign. This could use a line chart as a data visualisation.
* A pivot table comparing the length of crowdfunding campaigns to the outcome, with a column chart to visualise the results. This could show effective lengths for crowdfunding campaigns