Report Kickstart my chart

Adriana Avalos Vargas

03/08/2020

Contents

Introduction

Over $2 billion has been raised using the massively successful crowdfunding service, Kickstarter, but not every project has found success. Of the more than 300,000 projects launched on Kickstarter, only a third have made it through the funding process with a positive outcome.

Getting funded on Kickstarter requires meeting or exceeding the project's initial goal, so many organizations spend months looking through past projects in an attempt to discover some trick for finding success.

In this report the analysis of the database of 4,000 past projects is done in order to uncover any hidden trends that might help to identify the characteristics of successful campaigns.

The dataset

The dataset is composed of 4114 projects with a very wide range of categories and subcategories. The data base is distributed between four categories of successful and unsuccessful campaigns, as shown in Table 1.

Succes of the campaign	Number of campaigns
canceled	349
failed	1530
live	50

successful	2185
Total	**4114**

Table 1. Distribution of campaigns according to their success.

Therefore, it might be useful to analyze trends of successful campaigns. It is also noticeable that the database is disaggregated by country as shown in Figure 1.
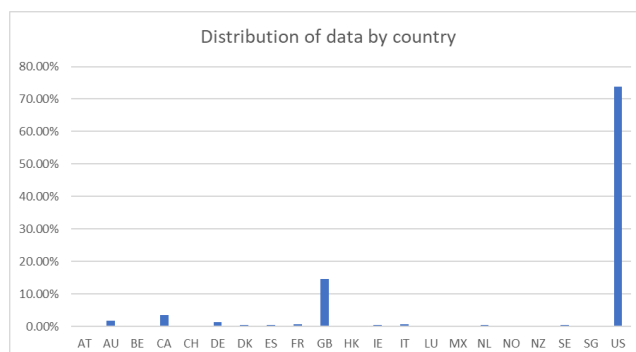

Figure 1. Distribution of data by country

Most of the data came from English-speaking countries, and to a less extent from other countries, like Mexico. Therefore, some caution must be taken when generalizing conclusions to all the countries. Also, the data is partitioned by project category and subcategory. Particularly, by category the campaigns are distributed as shown in Table 2.

Category	Number of campaigns
film & video	520
food	200
games	220
journalism	24
music	700
photography	220
publishing	237
technology	600
theater	1393
Total	**4114**

Table 2. Campaign categories

As we can observe most of the campaigns belong to the theater category, followed by music and technology. It is also possible to divide de campaigns into subcategories. The subcategories are divided as shown in Table 3.

Category	Number of subcategories	Most common subcategory
theater	3	plays
music	9	indie rock
technology	6	wearables

film & video	6	documentary
publishing	6	nonfiction
games	3	video games
photography	4	photobooks
food	3	food trucks
journalism	1	audio

Table 3. Number of subcategories per category.

Other data that is analyzed is the number of backers that funded the campaign, the goal fixed in each campaign, the pledge gathered by each campaign, the date when the campaign was launched, and the date when the campaign finished.

First conclusions from data

In global, more than half of the studied campaigns where successful, where the highest rate of success is for campaigns that are focused in music, theater, and film & video. On the other hand, the highest rates of fail occur in the food, games and publishing categories, while all the journalism campaigns were canceled. Also, most of the successful campaigns fulfil their goal and more.

On the other hand, in a subcategory level the higher probability of success is located in campaigns of rock (music), and documentary (film and arts), while the less probability of success is located in food trucks (food), wearables (technology), and animation (film and video). The play subcategory from the theater category dominates de probabilities of success and also of fail, but this could be due to the fact that most of the observations are from this subcategory.

There are few live campaigns, and are distributed among theater, music and food. These campaigns are located in the United States (33), Great Britain (8), Mexico (3), France (2) and one in Canada, Austria, Ireland and Netherlands. This could be due to the fact that the data base only contains data of the first months in 2017.

Also, in average if a project started in February, May or June has big chances to succeed, while if it starts in January, July or October it has big chances to fail.

Limitations of the dataset

There are some limitations than has to be considered when using the conclusions above, and are:

1. Even the data set is global, it contains more information about campaigns in the US in comparison with other countries, therefore the conclusions might apply only to the US
2. The number of projects is not evenly distributed in time. The years with mor projects are 2014 (976), 2015 (1225) and 2016 (950). Also, there are year with only one kind of campaign.

3. In order to make a fair comparison between the amount of the goal and the pledge by country, the amounts should be standardized to a given currency and base year, for example dollars PPP (purchasing power parity).
4. There is a big range between the number of backers, the goal and the pledge.
5. There might be some bias since most of the campaigns are in the theater category. However, this could be a natural trend and more investigation about the structure and rules of the Kickstarter crowdfunding service is required.

Future work for better understanding data

It could be interesting to analyze which are the most expensive campaigns according to their success, their category and subcategory. This could be achieved through a stacked bar graph of different ranges of goal for category, subcategory and also success. The same could be done for pledge.

Another important fact could be to understand the campaigns with more backers by category. In this way we could try to figure out which projects are more interesting to people in a given country. This could be achieved through a pivot table than has in its rows the category, in the column the number of backers (as a sum) and filter by country.

Finally, learning more about backers might help to decide where is appropriate to launch a given kind of campaign in order to get a success. For this, it might be useful to make a pivot table that in its rows has the subcategories, in its columns the mean of the average donation, and filtered it by country.