
Hidden Trends on **KICKSTARTER**



Suman Raj, Jan. 2019

Introduction

Kickstarter is a crowd funding service that helps creators and startups find the resources and support they need to make their projects come to life. Creators make campaigns on Kickstarter where they include the description of the project and funding goal. Backers financially support campaigns by purchasing different items that the campaign offers. The Kickstarter community comes together and helps many musicians, artists, designers and other creators achieve their missions. Since their launch on April 28th, 2009, 16 million people have backed a project launched on Kickstarter, \$4.06 billion has been pledged, and 156,852 projects have been successfully funded. This report analyzes 4,114 projects using Microsoft Excel visualization tools and outlines some key hidden trends found in the data.

Data

The data was collected in May 2017 and contains 4,114 Kickstarted campaigns that took place between May 2009 and May 2017. The following attributes were collected for each campaign:

name and description of the campaign

goal amount	<ul style="list-style-type: none">• how much the campaign was hoping to raise
state	<ul style="list-style-type: none">• “successful”, if the campaign was able to meet its goal• “failed”, if the campaign did not meet its goal before deadline• “canceled”, if the campaign was canceled• “live”, if the campaign was still going when data was collected
pledged amount	<ul style="list-style-type: none">• the total amount that the campaign was able to raise
backers count	<ul style="list-style-type: none">• the number of backers that supported the campaign
launch date	<ul style="list-style-type: none">• the date when the campaign started
category & sub-category	<ul style="list-style-type: none">• based on the its description, each campaign was categorized in one of the following categories and sub-categories (Figure 1).

Categories & Sub-Categories

film & video					food		games			music							photography				publishing				technology				theater											
animation	documentary	drama	science fiction	shorts	television	food trucks	restaurants	small branches	mobile games	table top games	video games	audio	classical music	electronic music	faith	indie rock	jazz	metal	pop	rock	world music	nature	people	photos	places	art books	children's books	fiction	non-fiction	radio & podcasts	translations	gadgets	hardware	makerspaces	space exploration	wearables	web	musical	plays	spaces
deadline																																						the date when the campaign ended		

Figure 1 Categories & Sub-Categories

Tools & Methods

In this project, Microsoft Excel was used to calculate new attributes from the given data and to create different visualizations that offer a better understanding of the data. The percentage-funded and average donation were calculated. The campaigns were divided into categories and sub-categories. Pivot tables and pivot charts were created to find hidden patterns in the data set.

Category pivot table

A pivot table was created to count the total number of “successful”, “failed”, “canceled” and “live” campaigns for each campaign category. (Table 1).

A stacked column pivot chart (Figure 2) was obtained from the pivot table and the following observations were made:

- The highest number of campaigns are in the theatre category.
- The theatre category contains the highest number of successful campaigns, but also the of failed campaigns.

Count of state	Column Labels				
Row Labels	canceled	failed	live	successful	Grand Total
film & video	40	180		300	520
food	20	140	6	34	200
games		140		80	220
journalism	24				24
music	20	120	20	540	700
photography		117		103	220
publishing	30	127		80	237
technology	178	213		209	600
theater	37	493	24	839	1393
Grand Total	349	1530	50	2185	4114

Table 1. Count of State for each Category

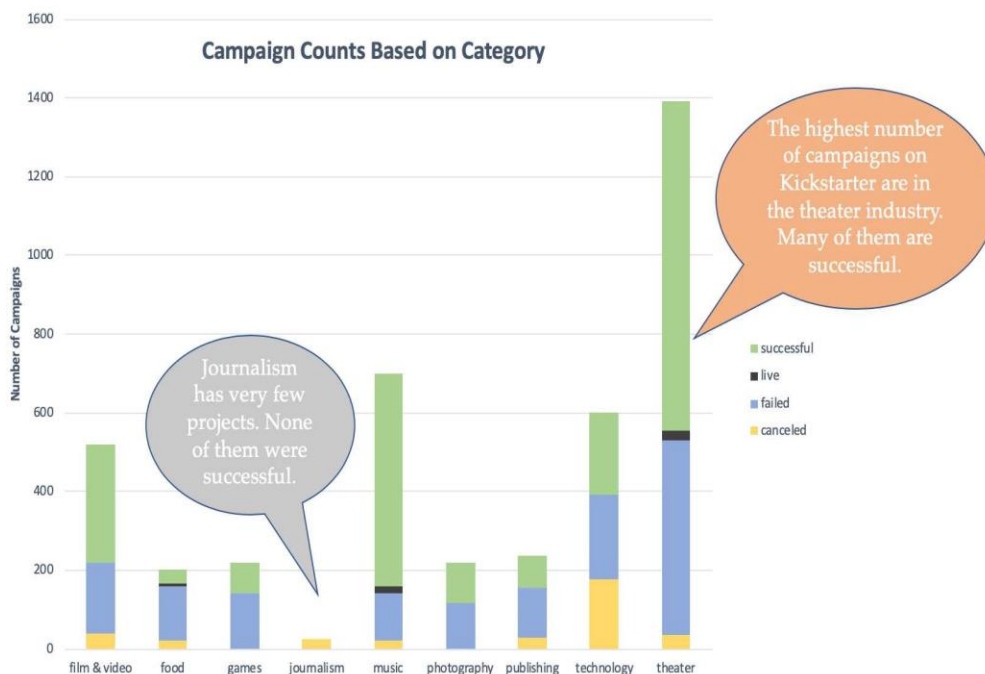


Figure 2 Campaign Counts Based on Category

- Journalism shows the lowest number of campaigns. All 24 of them were canceled.
- There are a lot more successful campaigns in Music than failed ones.

Sub-Category pivot table

A second pivot table was created to count the total number of “successful”, “failed”, “canceled” and “live” campaigns for each campaign sub-category.

The previous chart determined that theatre was the industry with the highest number of projects. From this pivot chart (Figure 4) we can see that plays, by far has the highest number of campaigns.

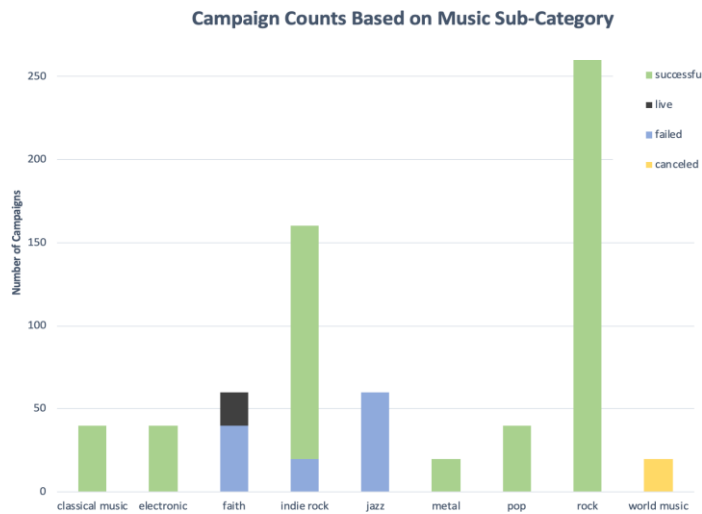


Figure 3 Campaign Counts in Music Category

Since the pivot chart can be filtered based on

Count of state	Column Labels			
Row Labels	canceled	failed	live	successful
animation		100		100
art books	20			20
audio	24			24
children's books		40		40
classical music				40
documentary				180
drama		80		80
electronic music				40
faith		40	20	60
fiction		40		40
food trucks	20	120		140
gadgets		20		20
hardware				140
indie rock		20		140
jazz		60		60
makerspaces		11		9
metal				20
mobile games		40		40
musical	20	60		60
nature		20		20
nonfiction				60
people		20		20
photobooks		57		103
places		20		20
plays		353	19	694
pop				40
radio & podcasts				20
restaurants		20		20
rock				260
science fiction		40		40
shorts				60
small batch			6	34
space exploration		18	2	40
spaces		17	80	5
tabletop games				80
television				60
translations		10	47	57
video games			100	100
wearables		60	120	20
web		100	60	160
world music		20		20
Grand Total	349	1530	50	2185

Table 2 Count of state for each Sub-Category

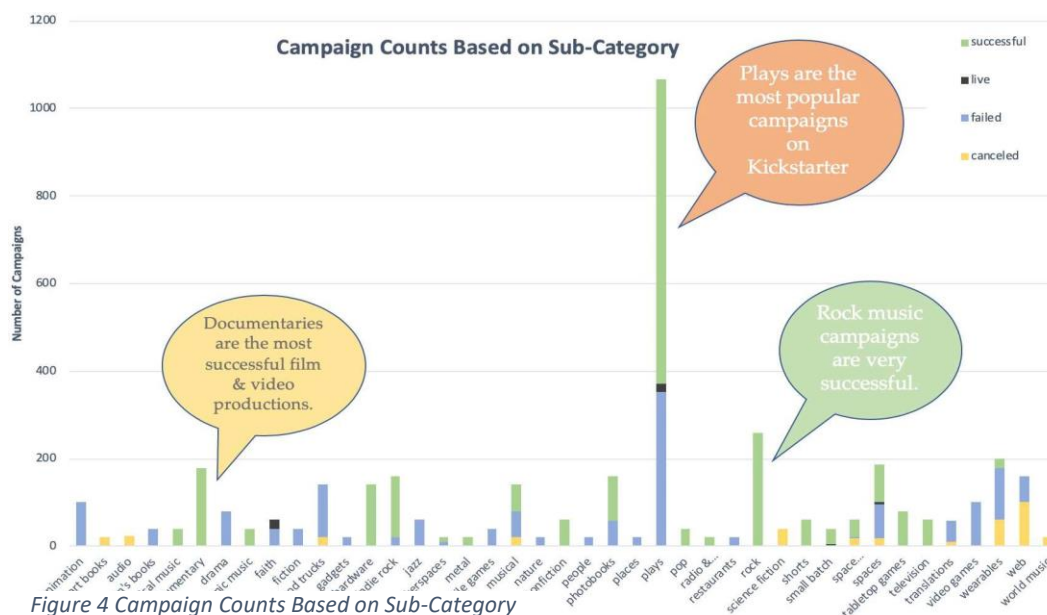


Figure 4 Campaign Counts Based on Sub-Category

the parent category, the most popular sub-category in each category can be seen. One example is the music category in Figure 3.

Start Date Line chart

A line chart (Figure 5) was created based on the starting dates of each of the campaign. This graph clearly shows that overall there are more successful campaigns on Kickstarter than the ones that failed or are canceled.

It also shows that more campaigns are started in January than in December. The data suggests that December is not a good time to start a new campaign. More than 50% of the projects that started in December failed. According to the data, a better time to start a campaign would be the month of May where there is a peak in the number of successful campaigns.

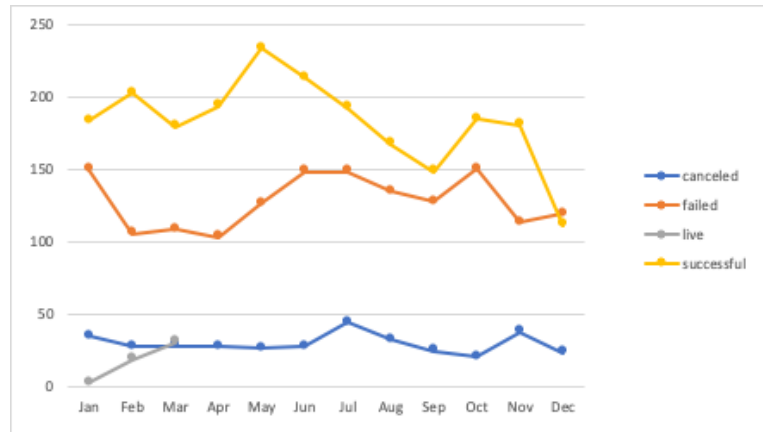


Figure 5 Campaign Counts Based on Start Campaign Date

Conclusions Made

1. The *theatre* industry category has the highest number of campaigns, out of which the subcategory *plays* has the highest number.
2. Journalism is not an industry that seems to be successful on Kickstarter. There were only 24 campaigns out of 4114 in journalism, and all of them were canceled.
3. All *rock music* campaigns were successful.
4. Overall, more campaigns are successful on Kickstarter than the failed ones, showing that Kickstarter is a great place to feature a new business.

Limitations of the Data Set

There are a few limitations to the data set and the tools that were used. Microsoft Excel can be quite slow when processing even this little amount of data depending on the operations that were taking place.









The data reveals correlation which suggest that, for example, May is the best month to start a Kickstarter campaign while December is the worst month to start a Kickstarter campaign. However, the data does not necessarily reveal why this is true. The cause for this correlation cannot be found in this data. It is important to find trends in the outcome of the different types of campaigns but finding out why these trends are happening would be of a higher interest to someone who wants to start a new campaign.

It is unknown if the chosen data for this analysis accurately represents all the 156,852 projects. Since we analyzed only 4,114 campaign, some bias might have been introduced into the data based on which campaigns were selected for analysis.

Other Possible Tables/Graphs That Can Be Created

1. What are the top 5 countries with the greatest number of successfully funded projects? What are the top 5 countries with the greatest fraction of successfully funded projects?

In order to calculate how many successful campaigns were funded in each country, a pivot table (Figure 6) was created with *country* field as rows and *state* field as columns. The value in the pivot table was *count of state*. To calculate the fraction of the successfully funded projects, the data was shown as % of row total. The table was sorted from the largest value to the smallest value based on the *successful* column. The results were the following:

Top 5 countries with the greatest number of successful campaigns		Top 5 countries with the greatest fraction of successful campaigns	
 United States	1651	 Luxembourg	100%
 Great Britain	366	 Singapore	100%
 Canada	64	 Great Britain	61%
 Germany	23	 United States	54%
 Australia	19	 Ireland	53%

Row Labels	cancel	failed	live	successful	Grand Total
AT		2	1	3	6
AU	14	41		19	74
BE		1		1	2
CA	17	64	1	64	146
CH	3	2		1	6
DE	3	27		23	53
DK	4	6		4	14
ES	1	9		11	21
FR	5	10	2	10	27
GB	25	205	8	366	604
HK	1	2			3
IE	2	4	1	8	15
IT	3	19		7	29
LU				2	2
MX	1	8	3		12
NL	4	14	1	2	21
NO		5		2	7
NZ	4	5		3	12
SE	5	9		7	21
SG				1	1
US	257	1097	33	1651	3038
Grand Total	349	1530	50	2185	4114

Figure 6 Outcome counts based on country

2. How does the goal amount influence the outcome?

In order to answer this question, the data was distributed in 8 different categories based on the goal amount the project was hoping to get. The categories were the following:

- Goal is less than \$500.
- Goal is between \$501-\$1,000.
- Goal is between \$1,001-\$2,500.
- Goal is between \$2,501-\$5,000.
- Goal is between \$5,001-\$10,000.
- Goal is between \$10,001-\$25,000.
- Goal is between \$25,001-\$50,000.
- Goal is more than \$50,000.

A pivot chart was created (Figure 7) and the following observations were made:

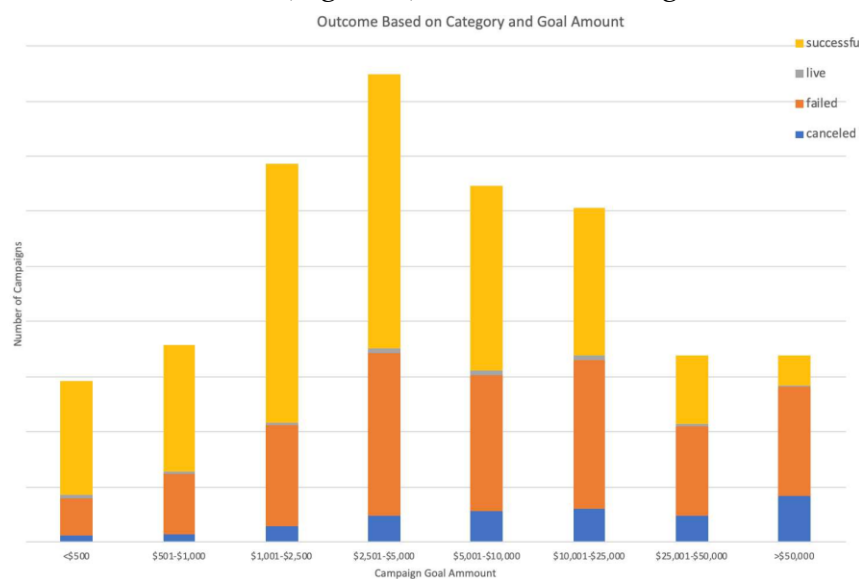


Figure 7 Outcome based on Category and Goal Amount

- Most campaigns chose a goal between \$2,501-\$5,000.
- The number of failed campaigns is much smaller than the number of successful campaigns when the goal is less than \$500.
- The number of successful campaigns that have a goal bigger than \$50,000 is very small.
- As the goal amount increases, the number of successful campaigns tends to decrease.

3. Given a category and a goal range, what is the probability a new campaign reaching its goal budget?

In order to answer this question, the *Category* field was added to the pivot table created for the previous question and the data was displayed as percentage of the row total. The successful outcome was plotted in a bar chart (Figure 8).

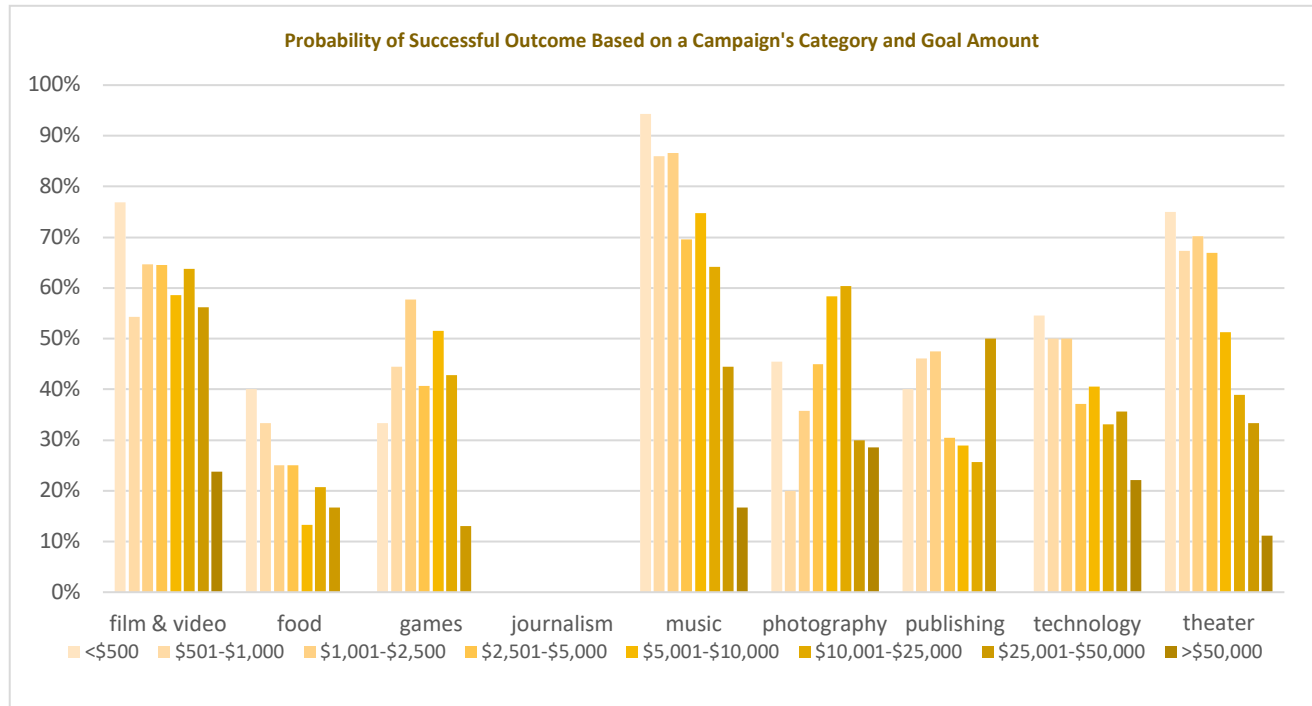


Figure 8 Probability of Successful Outcome based on Goal and Category

This chart in Figure 8 can be used for new campaigns to make informed decisions. For example, it can clearly be seen that a new campaign in music is more likely to succeed if the goal amount is less than \$500. This however does not apply for the publishing business. In the publishing business, more campaigns succeeded that had a goal amount higher than \$50,000.

Despite the data limitations used for this assignment, the dataset could still be used to drive interesting conclusions. These are just a few examples of analysis that can be done with this data. There are many other methods that can be used to drive insights from the data.