Juan Sebastian Galindo  
Kickstart my chart

5/9/2020

# Overview

## Project Background and Description

|  |  |
| --- | --- |
|  | 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 to discover some trick for finding success. |

## Purpose

|  |  |
| --- | --- |
|  | We were asked to organize and analyze a database of 4,000 past projects in order to uncover any hidden trends. |

## Limitations

|  |  |
| --- | --- |
|  | What are some of the limitations of this dataset? |

The first limitation I faced when examining the data were the outliers. From this database, I was able to measure the success of the campaigns in 3 ways:

* Ratio of money pledged to goal which was called percent funded (% funded).
* Number of backers.
* Count of outcome.

Before starting my analysis, I realized that the data was originated from different countries and as a result, contained different currencies. In order to convert the currency data to a single currency, the “country” data was converted to geographical data and a web table was imported from [www.floatingrates.com](http://www.floatingrates.com). Using the geographical information and the web table, I was able to convert all the data into US Dollars and then populate that data in two new columns called “Goal in Dollars” and “Pledged in Dollars”.

.

A statistical analysis was performed from the % funded and number of backers’ data. The live projects were excluded from the analysis and projects classified as cancelled or failed were reclassified as unsuccessful. The results from the statistical analysis showed that the data for successful and unsuccessful projects was leptokurtic and positively skewed thus indicating that the data was heavily influenced by outliers; results are included in the appendix Table 1. In order to reduce the influence of the outliers in these data sets; a statistical analysis was performed again to each individual category and the upper and lower values were determined (Appendix Table 3). Two new columns called “Capped % Funded” and “Capped Backers” were created and in these two new columns, the outlier values were capped with the upper bound values for each individual category.

The second challenge I faced was the definition of success. In order to analyze the data, I decided to use two different approaches. The first approach focused on the countries, categories and subcategories with a probability equal or higher than 50%. The second approach compared the “Capped % Funded” and analyzed the trends of the projects that were able to achieve 100% of funding or more.

The third limitation I faced was the bias of the database towards a few countries. When I filtered and compared the geographical location of the projects, I discovered that 74% of the projects were held in the United States (see Appendix Table 2). The next two countries were the United Kingdom with 15% and Canada with 4% of the total projects. When combining these three countries, they accounted for 92% of the total data. It is important to note that the findings from the analysis of this data were heavily influenced by the demographics and socio-economic conditions of these three countries.

## Observations

|  |  |
| --- | --- |
|  | What are some other possible tables and/or graphs that we could create? |

Campaign Duration Time:



Figure 1.

Observations:

* Most of the projects with successful outcomes had a campaign duration time between 15 to 40 days and peaked during the 30 to 34 days interval.
* The campaigns’ duration times should be no less than 10 days or no more than 60 days.
* The campaigns with duration times higher than 60 days are rarely successful.

Target Funding:



Figure 2

Observations:

* 70% of the projects with a funding goal less than $1,000 had a successful outcome.
* 65% of the projects with a funding goal between $1,000 and $5,000 had successful outcomes.
* The probability of success was 50% for projects with a funding goal of $10,000.
* Projects with a funding goal between $10,000 and $40,000 had an estimated average success of 45%.
* The probability of a successful outcome decreased considerably for projects with a funding goal of $45,000 or more.

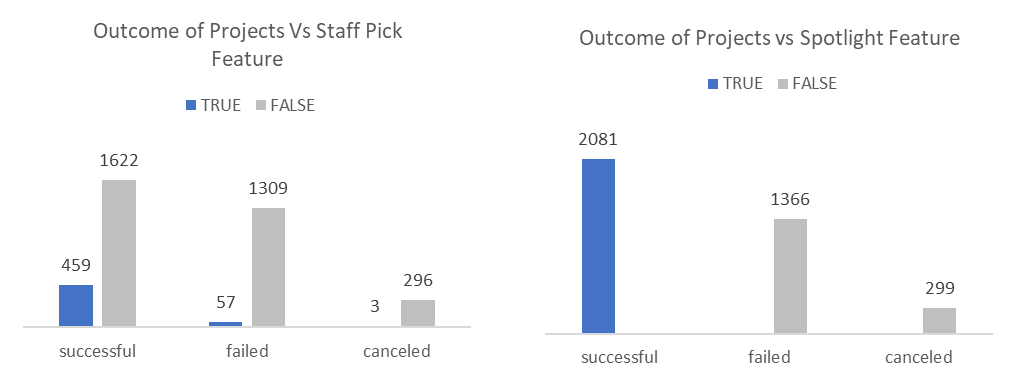
Visibility options:

Figure 3

Observations:

* Staff Pick:
  + Staff pick had no substantial effect on the outcome of the project.
* Spotlight:
  + 100% of successful projects had a spotlight feature.

Outcome Distribution of the top 3 countries:

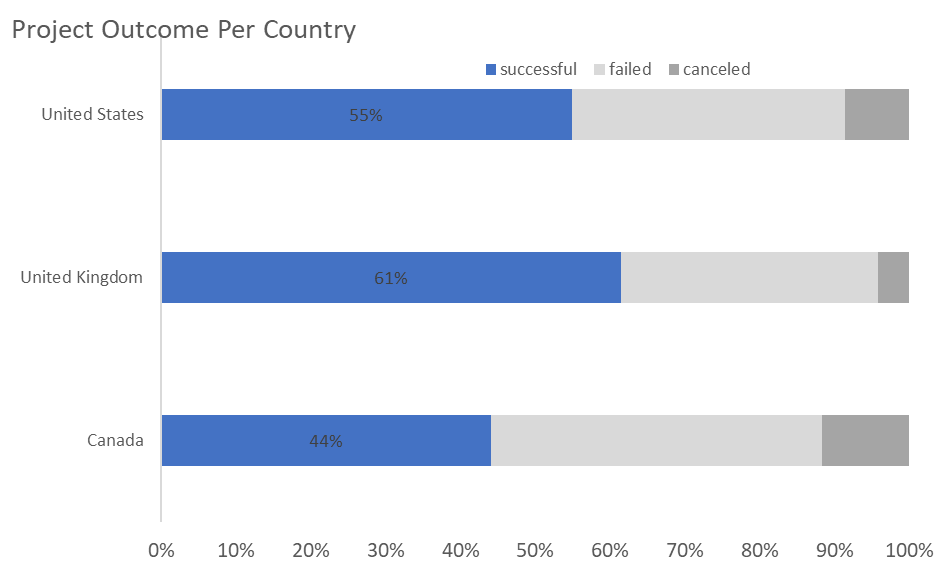


Figure 4

Observations:

* Projects held in the United States had a probability of success around 55%.
* Projects held in the United Kingdom had the highest probability of success (61%).

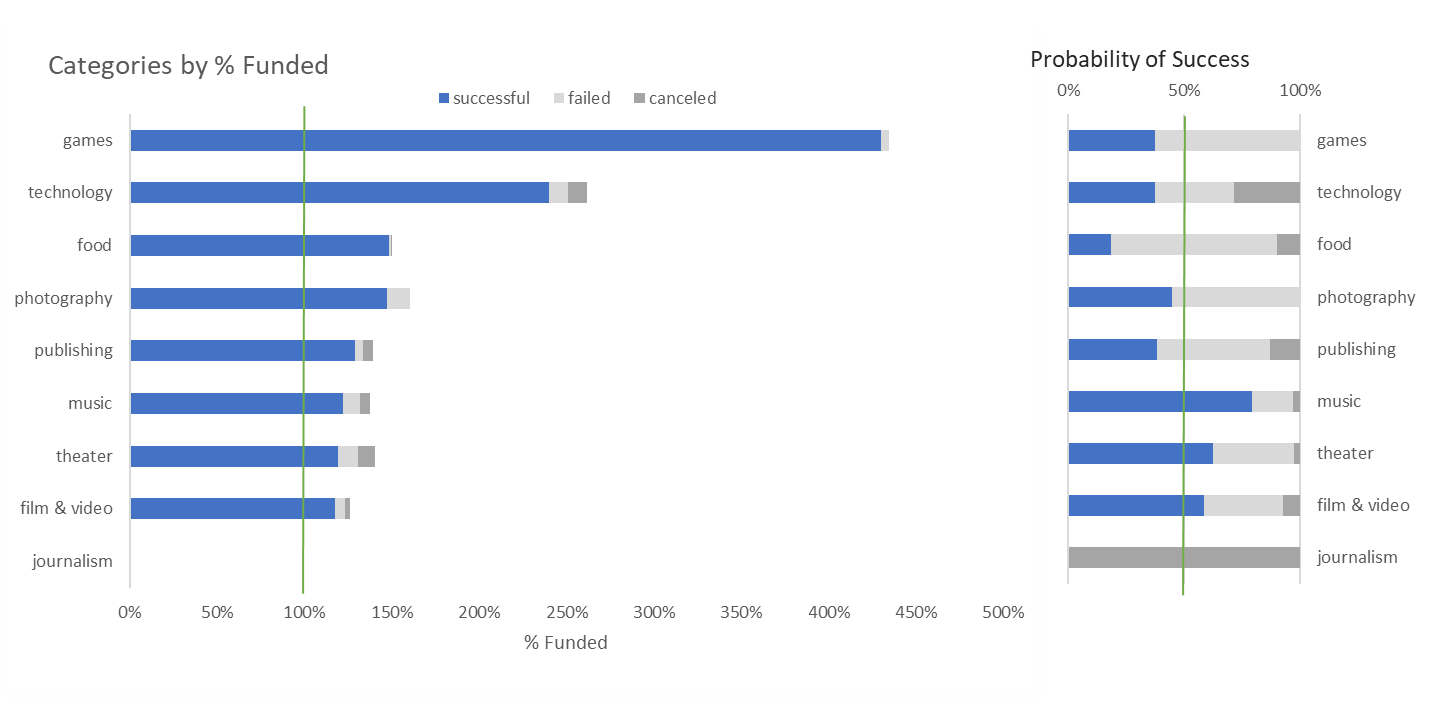


Figure 5

Observations:

* 8 out 9 categories were able to achieve an average of 100% or more of the funding goal.

1. Games
2. Technology
3. Food
4. Photography
   1. Food and photography had similar average % funded but photography had a higher probability of success.
5. Publishing
6. Music
7. Theater
8. Film & Video

* Only 3 out of 9 categories had a probability of success higher than 50%.

1. Music
2. Theater
3. Film & Video

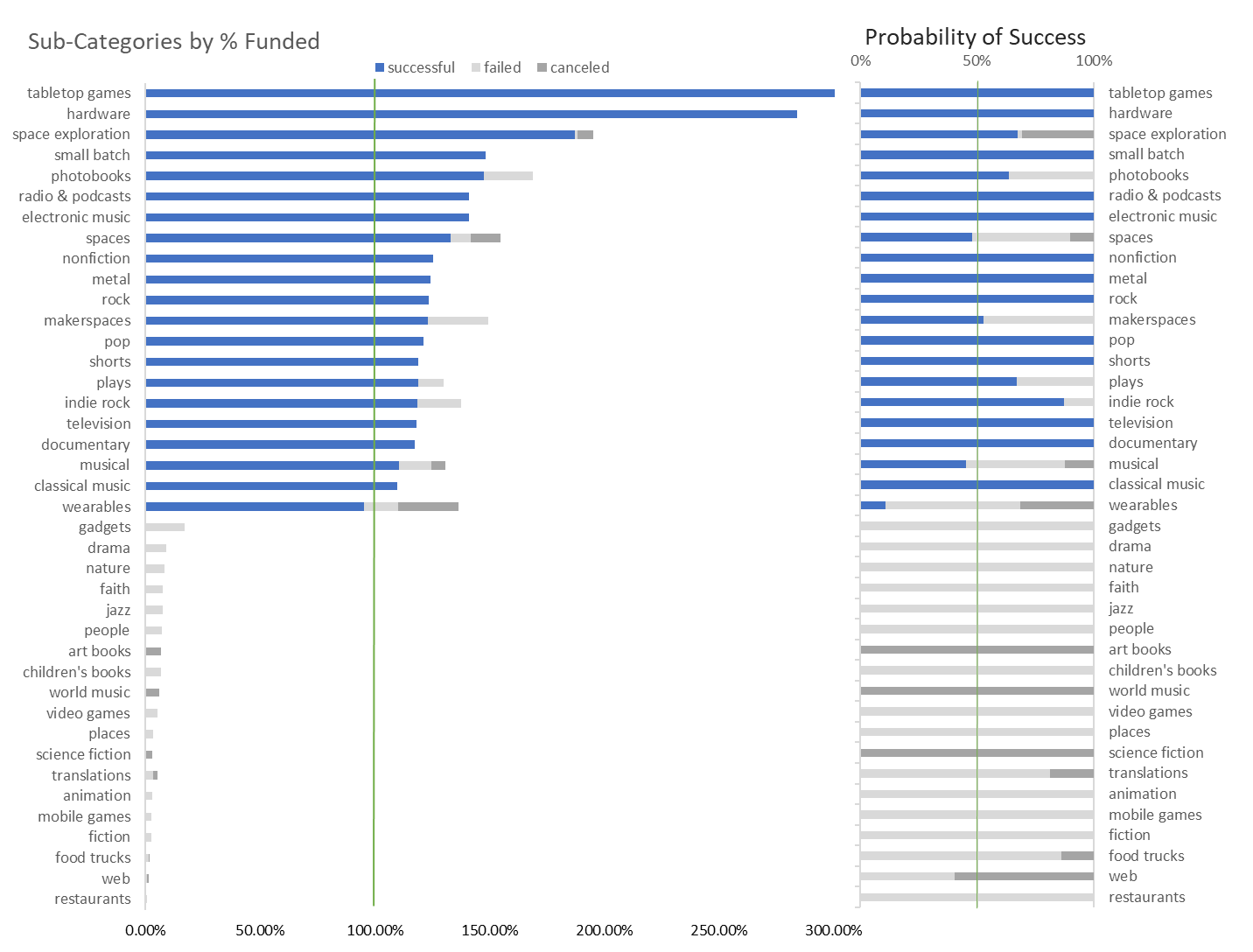


Figure 6

Observations:

* Only 20 out of 41 subcategories were able to achieve an average of 100% or more of the funding goal.

1. Games
   1. Tabletop games
2. Technology
   1. Hardware
   2. Space exploration
   3. Markerspaces
3. Food
   1. Small Batch
4. Photography
   1. Photobooks
5. Publishing
   1. Radio & podcasts
   2. Nonfiction
6. Music
   1. Electronic music
   2. Metal
   3. Rock
   4. Pop
   5. Indie rock
   6. Classical music
7. Theater
   1. Spaces
   2. Plays
   3. Musical
8. Film & Video
   1. Shorts
   2. Television
   3. Documentary

* 18 out of the 20 subcategories had a probability of success higher than 50%.

1. Games
   1. Tabletop games
2. Technology
   1. Hardware
   2. Space exploration
   3. Markerspaces
3. Food
   1. Small Batch
4. Photography
   1. Photobooks
5. Publishing
   1. Radio & podcasts
   2. Nonfiction
6. Music
   1. Electronic music
   2. Metal
   3. Rock
   4. Pop
   5. Indie rock
   6. Classical music
7. Theater
   1. Plays
8. Film & Video
   1. Shorts
   2. Television
   3. Documentary



Figure 7

Observations:

* The top 3 categories by % funded:
  1. Games
     + Games were the most popular among backers. The average project had 293 backers.
     + Games had the lowest average pledge. The average pledge per project was $68.
  2. Technology
     + Technology was the second most popular category with an average of 254 backers per project. The average backer for this category was only 13% lower than the average for games.
     + Technology was the category where backers were willing to spend the most with an average pledge of $153. The average pledge for technology was 125% higher than games.
  3. Photography
     + Photography was the least popular of the top three categories with an average of 145 backers per project.
     + Photography had an average pledge of $100 which was relatively high when compared to other categories.
* Categories with 50% or more probability of a successful outcome:
  1. Theater
     + The average backer per project was the lowest among the other 3 categories with an average of 52 backers per project.
     + Theater’s donation average was $86 per project.
  2. Film & Video.
     + Film & video was the most popular of the 3 categories with an average backer count per project of 126.
     + Film & video had the highest pledge average of the 3 categories with $107 per project.
  3. Music.
     + Music was a little bit more popular than theater with an average backer count of 68.
     + Music’s lowest donation average of the 3 categories with an average of $73.

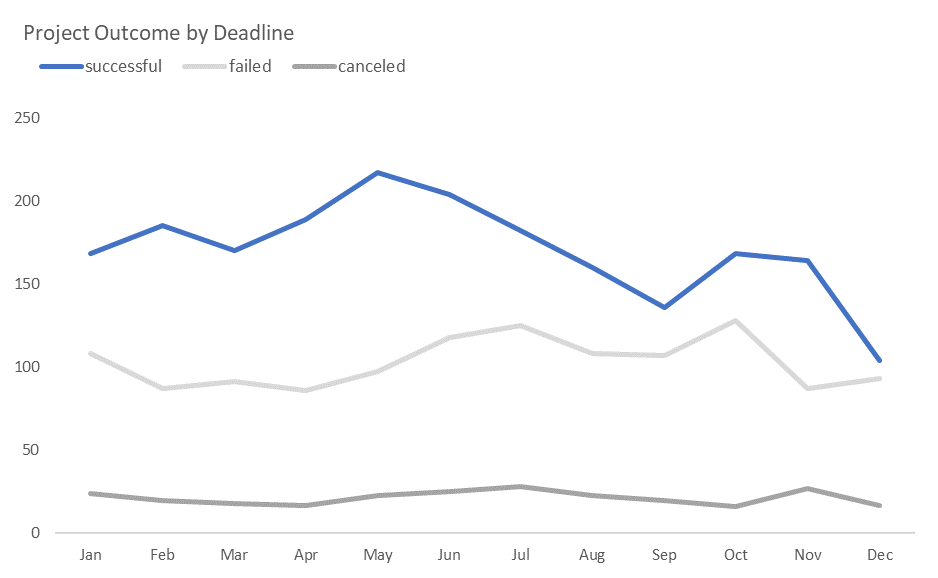


Figure 8

Observations:

* Most of the successful projects had a deadline set for May.
* There was a declining trend in successful projects from May to September.
* There was a second peak in October, but this peak was short lived, and it fell again right before the holidays.

## Conclusions

|  |  |
| --- | --- |
|  | Given the provided data, what are three conclusions we can draw about Kickstarter campaigns? |

1. My first recommendation regards the geographical location of the project. Based on the data, I would recommend the project be in one of the next 3 countries listed in order of importance.
   * + 1. United States
          1. From 2009 to 2017 the average probability of a successful outcome was 55% (fig 4).
          2. In 2019, the U.S population was 328.2 million.
       2. United Kingdom
          1. From 2012 to 2017, the average probability of a successful outcome was 61% (fig 4).
          2. In 2019, the U.K population was 66.65 million.
       3. Canada
          1. From 2013 to 2017, the average probability of a successful outcome was 44% (fig 4).

Even though the United Kingdom had the advantage due to the highest probability of a successful outcome, this advantage was limited by the smaller population when compared to the population of the United States.

1. The second recommendation I would like to make is regarding the scope of the project. Based on my analysis performed in fig 5 and 6, I determined that in order to secure at least 100% of the funding and have a probability of success higher than 50%, the project should fall in one these sub-categories:
2. Games - Most popular by category (fig 7).
   1. Tabletop games
3. Technology – Highest average pledge by category (fig 7).
   1. Hardware
   2. Space exploration
   3. Markerspaces
4. Food
   1. Small Batch
5. Photography
   1. Photobooks
6. Publishing
   1. Radio & podcasts
   2. Nonfiction
7. Music – Highest probability of success by category (fig 5).
   1. Electronic music
   2. Metal
   3. Rock
   4. Pop
   5. Indie rock
   6. Classical music
8. Theater
   1. Plays
9. Film & Video
   1. Shorts
   2. Television
   3. Documentary
10. My final recommendation is in regard to the goals, the duration and the features of the campaign. Based on my analysis, the duration of the campaign should be between 30 to 34 days (fig 1). The project should be posted in April and have a deadline sometime in May (fig 8) in order to have the highest probability of success.

Based on fig 2, I can make the following recommendations in regards to the funding goal of the project:

* 1. Funding goal should be less than $10,000 in order to have a probability of success higher than 50%.
     1. The average probability of success for campaigns with funding goals between $1 - $5,000 is 67.5%.
  2. If the project requires a funding goal higher than $10,000, the campaign goal should not exceed over $43,000 in order to achieve an average probability of success of 45%.
  3. The project should not have a funding goal higher than $43,000 because the probability of success decreases significantly after this point.

The project must have a spotlight landing page and it should be kept up to date as the project progresses (fig 3).

*“Spotlight is a permanent landing page where viewers and investors can stay up-to-date on your project, even after completion. It emphasizes the results of the project and showcases how you got to where you are today. The page is highly customizable so you can design it to match your brand and image. Additionally, you can use it to direct traffic to sources outside of Kickstarter – such as to your website or social media accounts.”* “How Campaign Creators Can Use Kickstarter’s Spotlight Feature. *“Art of the Kickstart,* <https://artofthekickstart.com/campaign-creators-can-use-kickstarters-spotlight-feature/>. Accessed on May 3, 2020.

The staff pick badges has been discontinued and it was replaced with “Projects We Love.” Based on the data, the staff pick had no influence in the outcome of the project. Unfortunately, I do not have the appropriate data to make any recommendations for the “Projects We Love.”

## Appendix

|  |  |
| --- | --- |
|  | 1. Use your data to determine whether the mean or the median summarizes the data more meaningfully. 2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not? |

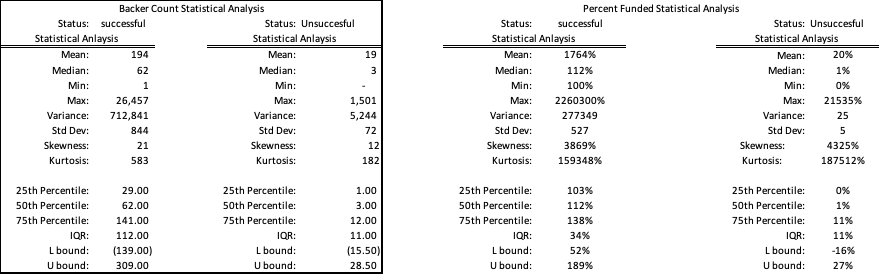


Table 1

Observations:

1. Based on the statistical analysis of the backer count data, the median summarizes the data more meaningfully.
2. Based on the backer count data, there was more variability with successful campaigns.
   * The high variability in the successful backers data makes sense because some projects were able to attract thousands of backers most likely due to social media exposure but these projects were outliers and they do not represent all the other projects in the database.



Table 2



Table 3