Kickstarter Data Analysis

Pratyush Painuly

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

**Conclusion 1 : Smaller goals have a higher success rate (Goals 15k and below are most likely to succeed)-**

A screenshot of a map

Description automatically generated

From the chart, we can see that goal limits over 15,000 mostly stay below the 50% mark as opposed to the ones below, which mostly stay above it. The average success rate for goals below 15,000 is 77% whereas the average success rate for amount over 15000 is 34%. Hence, the campaigns with lower goals have a higher rate of success. Most successful campaigns raise less than 15,000 dollars for their goal amount.

A screenshot of a cell phone

Description automatically generated

**Conclusion 2: Theatre, music and film & video are the most successful categories in the dataset.**

A screenshot of a cell phone

Description automatically generated

Before diving into the dataset, I had this misconception that tech would be a popular category because most famous campaigns that I had heard about were tech related. Turns out that there are far lower successful campaigns for tech and the dataset is dominated by theatre, which is the most successful, music and film and video categories. Its interesting that Kickstarter founders were artists and a huge fan of Burning Man music festival and their company has organically catered to projects that relate to artists one way or the other.

A screenshot of a cell phone

Description automatically generated

Comparing the total campaigns in the dataset(regardless of the outcome), there are far more campaigns for theatre and music, but technology becomes the third highest category for the total projects in kickstarter data.

Looking at the sub-categogry, Theatre Plays far outcrosses any other sub-categories, with rock music and documentaries being second and third respectively. So we can ask ourselves a question, is kickstarter a great vehicle for theatre/plays?

A screenshot of a cell phone

Description automatically generated

The graph below gives some insight into that question. The following is all the unsuccessful campaigns by category. (All projects that were cancelled or failed to reach their goal are defined as unsuccessful)

A screenshot of a cell phone

Description automatically generated

Here, we see that the sub category Plays also dominates the dataset. So Sub-cateogry Plays is not only the most successful, but also the least successful. Same goes for Rock music and Documentary. This gives us an indication that on average, Kickstarter is very popular and is a huge attraction for people who want to create theatre plays, rock music and documentary. Just that its not a great indication that these fields will yield a successful campaign, even though the success rate for them is the largest.

**Conclusion 3: Late Spring to Early Summer is a good time for Kickstarter campaigns**

A close up of a map

Description automatically generated

Month of May has the most successful campaigns, followed by July, Feb and April. Early Summer seams to bring in most campaigns and of those, the number of successful ones dominate this time period. Whereas, winter(December has the least successful campaigns on average), has the worst number of successful campaigns, but at the same time, it does not have the lowest number of failed ones either, hence not a very productive time for kickstarter campaigns.

Comparing May(most successful with high failed number), April(4th most successful) and December(least successful)

A screenshot of a computer

Description automatically generated

April has the 4th highest successful campaigns and also the lest failed campaigns. Clearly a great month for kickstarter campaigns for all potential entrepreneurs. So maybe April stands out as the best month even though it has the 4th most successful campaigns.

**What are some limitations of this dataset?**

Location : The dataset gives the Origin country but it would be interesting to look at the city that the campaign started from. It would give us an indication of how spread out this data set is for location. Do most campaigns originate in Big cities? We can then look at which category is popular in which city.

Social Interaction : The data set does not give details about the comments, updates for each campaign. Is social interaction with backers an indicator of the popularity of the idea? We can then look into what does the engagement with backers tell us if a campaign will be successful. Do creators who give regular update see an increase in backers?

Backer count by time : The data set does not provide any insight into how did the backer count increase or decrease over time until the deadline. We can use that data, along with social interactions, date of the year and location to get some idea as to how does one correctly formulate the approach for a successful campaign.

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

* Duration of each project for all campaigns. How does duration impact success or failure? Do shorter durations result in more success?
* We can create tables and charts for the top 10 successful campaigns with the highest goal amount and analyze by category, sub-category. Technology for example, has a lower number of successful campaigns, but 8 of the 10 highest goals are campaigns related to technology/hardware.
* We can create chars for which category brings in the most money, and then see how many were successful, and how many just fell short of their goal despite generating large amount of revenue.