

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

1. The category that benefited the most from Kickstarter campaigns was the Theatre category. It had the highest success rate (839 successful campaigns) in comparison to other categories.
2. Journalism did not benefit from Kickstarter campaigns as they had 0 successful campaigns. The Journalism category was the least active and had very little participation. It had 24 campaigns total and all 24 were cancelled.
3. By analysing the data on all sub-categories, it is very clear that Plays are massively successful as they have the most amount of campaigns (a total of 1066) and 694 successful campaigns. That is a little more than 65% success rate!

What are some limitations of this dataset?

This dataset does a good job at listing the state of the campaign. It tracks if the campaign was successful, cancelled, if it failed or if it is still live. To further analyse the reason behind the success or failure of a campaign to find commonalities or correlations it would have been helpful if this dataset included the reason for *why* a campaign was cancelled. There are many reasons why campaigns were cancelled but this dataset did not collect this data, therefore we are unable to draw conclusions as to the reason behind cancellations. Knowing the most common reasons as to why a campaign was cancelled can impact future campaigns because Kickstarter can make changes to increase the success rate of each campaign.

Another limitation of this dataset is that it doesn't provide other helpful geographical data such as states/cities per country. Kickstarter might want to analyse what geographical area per country has higher support for campaigns but is unable to because this data was not collected. This limitation affects how the company can analyse the areas that possibly need better marketing strategies to increase the success rate per campaign.

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

A scatter plot can be created to see the correlation between the launched date and the outcome of each campaign. This will help us understand if the success of campaigns is affected by seasonality.

A bar graph that displays the percent funded per category and sub-category would be beneficial to understand and clearly see the success. By adding a line that determines the average percent funded to this bar graph we can understand which category was above and below average in terms of percent funded.