Kickstarter Project Campaign -Outcome Analysis

Report

Assignment #1

Submission date: 07/09/2020

Student:

Cicily George

Instructor:

Juan Luong

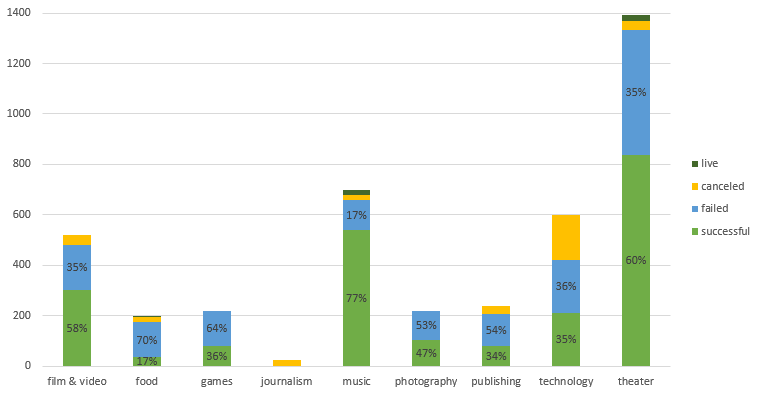
Course:

Bootcamp - Data Analysis

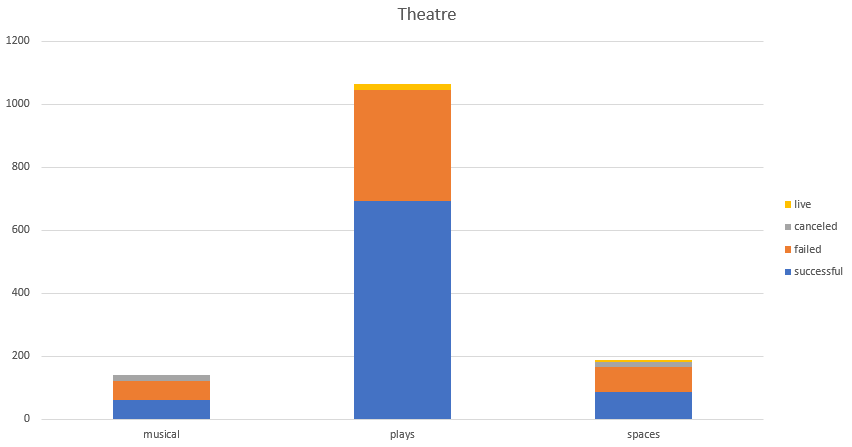
1. **What are the conclusions we can draw about Kickstarter campaigns?**

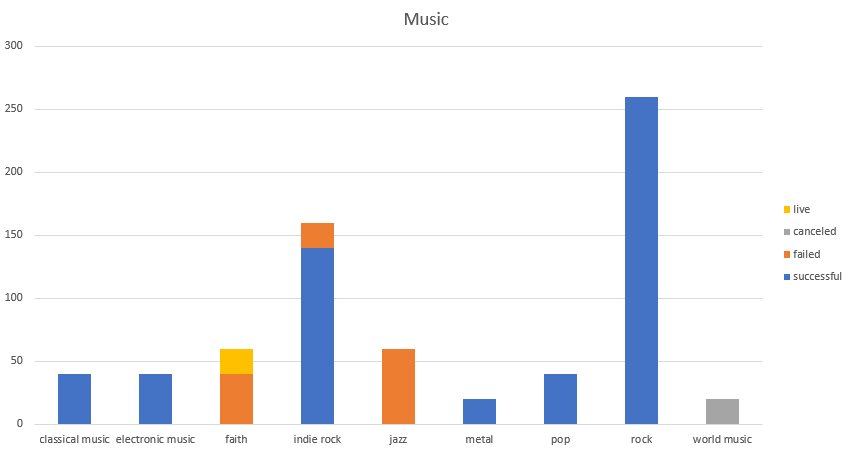
Analyzing a database of 4,000 past projects, the inferences are:

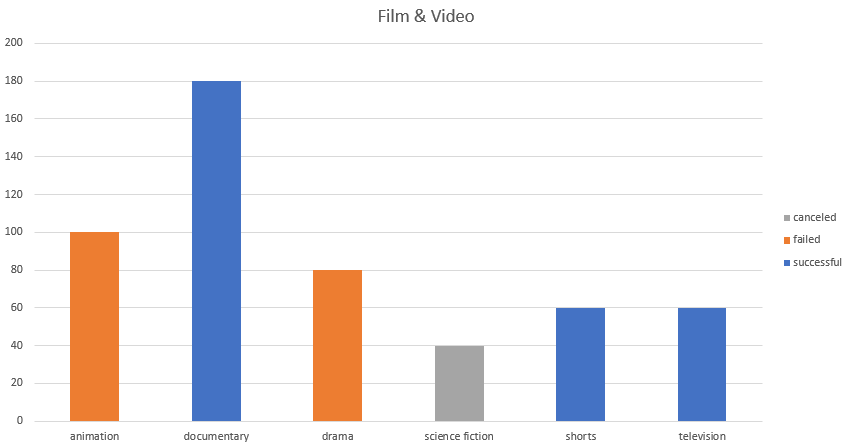
* Theatres category has the greatest number of successful projects but the percentage of successful projects is highest in the category Music.



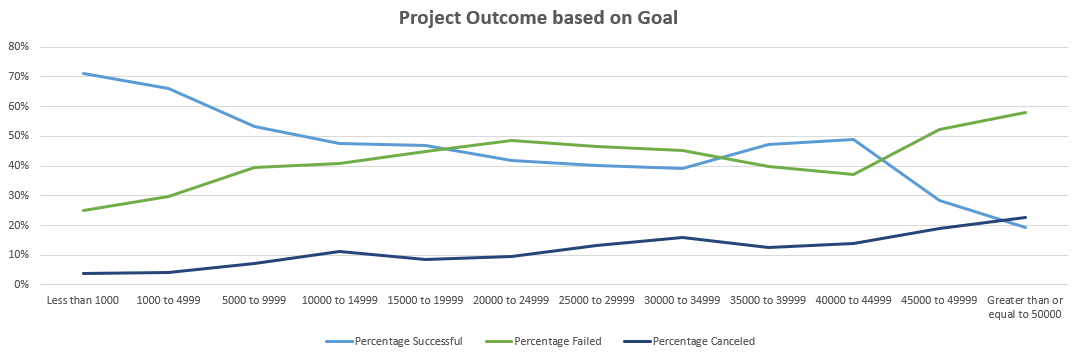
* Studying the Theatre, Music and Film & video categories where the success percentage is higher, there are some sub categories where all the projects were successful under Music and Film & video categories. They are
  + Music: **Rock, Classical music, Electronic music, Pop and Metal** and
  + Film & video: **Documentary, Short, and Television films**







* Lower the Goal amount, higher the success rate. As the Goal amount increases, the percentage of projects failed and canceled increases and the percentage of Successful projects decreases.



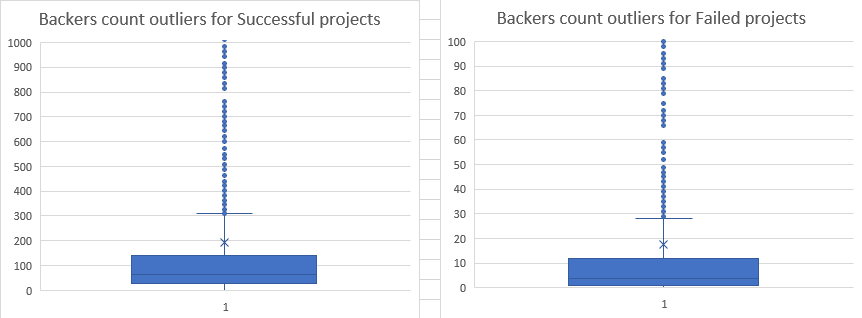
1. **What are some limitations of this dataset?**

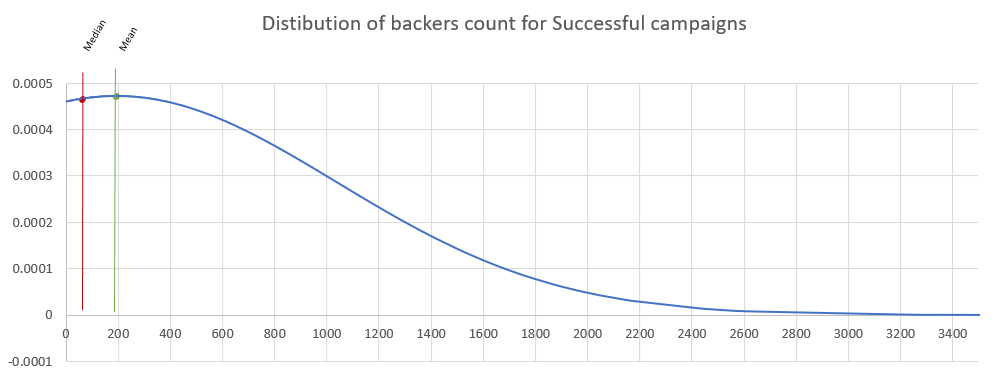
There are limitations on making decisions based on this dataset alone. There are external factors too affecting the outcome of a project campaign like:

* Trend- whether or not each industry or category or subcategory does well or poorly during the relevant time periods.
* Kickstarter’s popularity and consumer trust- certain categories are more successful during Kickstarter’s initial years while some others kick in later.
* Most of all, success of a campaign depends on the project idea and its implementation

The data set has not been normalized; it contains outliers (extreme values) which affects the analysis. The variance of the Backers count is high especially for successful campaigns. Its lowest (minimum) value is 1 and the highest (maximum) value is 26457.

1. **What are some other possible tables and/or graphs that we could create?**
2. Tables and Graphs that analyse the project duration for Successful and Failed project campaigns.
3. Graphs to visualize the significant funding Goal outliers.
4. Graphs and tables that exclude significant funding Goal outliers.
5. Graphs to visualize the distribution of Backers count for Successful and Failed project campaigns and check for outliers



****

****

Backers count above 309 for successful and above 29 for failed campaigns are outliers.

**Bonus Questions:**

**1.Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

Median summarizes the data more meaningfully. Because there are outliers in the data, the Variance is high.

**2.Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

The Backers count for Successful campaigns is more variable than for Failed campaigns. Which makes sense because if the project idea is good, there will be a greater number of Backers or more Pledged amount even if there are a smaller number of backers, which leads to success of the campaign.

And if the idea is not good there will be a smaller number of Backers and Lesser Pledged Amount. Which implies Failure. When the values are small, the variance will be less.