KickStart My Chart

Unit 1 Assignment

Data Analytics Bootcamp

Sandra Mejia Avendaño

# Background

Popular crowdsourcing platform Kickstarter has successfully raised 4.1 billion dollars to fund 158,240 projects, according to its own published statistics. Receiving the raised funds is dependent of meeting or exceeding the project’s initial goal, which has led to a success rate of 36.69%. From a total of 15,795,195 backers, almost a third – 5,155,329 – contribute at least once more, which leads to the reasonable assumption that recurrent funders might tend to back similar projects. Given the low success rate, it is worth analyzing past projects to try and find trends. Four thousand past projects are analyzed to this end.

# Highlights and manipulations on dataset

* In the attached Excel file, sheet called ‘**Project list**’, the state of each project is highlighted in different color: **successful**, **failed**, **cancelled**, or currently **live** (such colors will be consistent throughout the analysis).
* Column O was added to calculate the percent raised towards the project. The percentage is highlighted in a gradient that goes from red at **0%**, to green at **100%**, and to blue at **200%**.
* Column P was added to calculate the average donation of each backer to the project – left blank if there were zero contributors.
* Column Q was added to describe of the category of the project, and column R to describe the subcategory of the project.
* The launch date and deadline dates were converted to YYYY-MM-DD format in columns S and T respectively.

# Status of each campaign per category

The sheet called ‘**Status per category**’ contains a pivot table and chart, showing the number of successful, failed, cancelled or live campaigns per category, which can be filtered by country.

Figure 1 is a summary of the all the campaigns in all the countries, clearly showing theater as the most popular category. Figure 2 presents a summary of the campaigns and their status in four different countries. US, Great Britain and Canada account for 92% of all the projects worldwide. Mexico account for only 0.29% and is shown only for its geographical relevance.

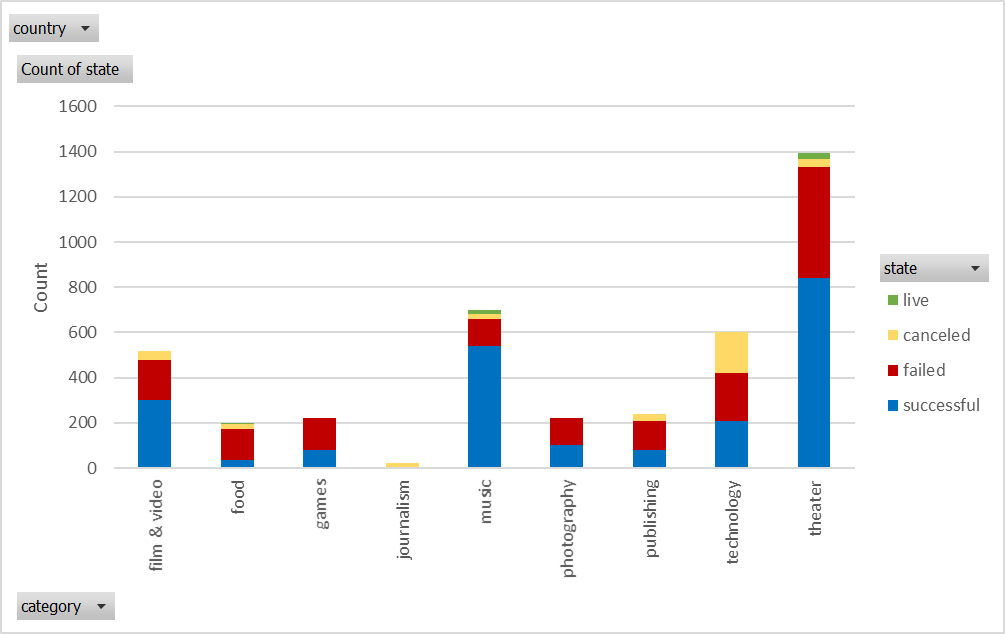


Figure 1. Count of successful, failed, cancelled and live projects per **category**, all countries pooled (4114 projects).

|  |  |
| --- | --- |
| **United States (3038 projects)** | **Great Britain (604 projects)** |
| **Canada (146 projects)** | **Mexico (12 projects)** |

Figure 2. Count of successful, failed, cancelled and live projects per **category**, in the three **countries** with more projects (US, GB, CA), and Mexico.

## Observations

* US is the country that dominates the number of projects with 3038 (73.8% of total), distantly followed by Great Britain (14.7% of projects) and Canada (3.5% of projects).
* Theater is consistently the category with more campaigns.
* Overall, theater is the category with the highest number of successful campaigns. However, the success rate of other categories (mainly music) is higher.
* In Great Britain, theater category clearly dominates all the postulated projects, and therefore has the highest number of successful campaigns.
* In the US and Canada, theater still dominates the number of campaigns, but there is also a significant presence of other categories such as music, film & video, and technology.
* The category with an apparent highest success rate is music.
* Mexico is a wildly non-representative country, since it accounts for only 0.29% of the projects, and none has been successful.

# Status of each campaign per sub-category

The sheet called ‘**Status per sub-category**’ contains a pivot table and chart, showing the number of successful, failed, cancelled or live campaigns per sub-category, which can be filtered by country and parent category. Figure 3 is a summary of the campaigns in all countries, divided by subcategory. It is evident that the most popular subcategory is plays.

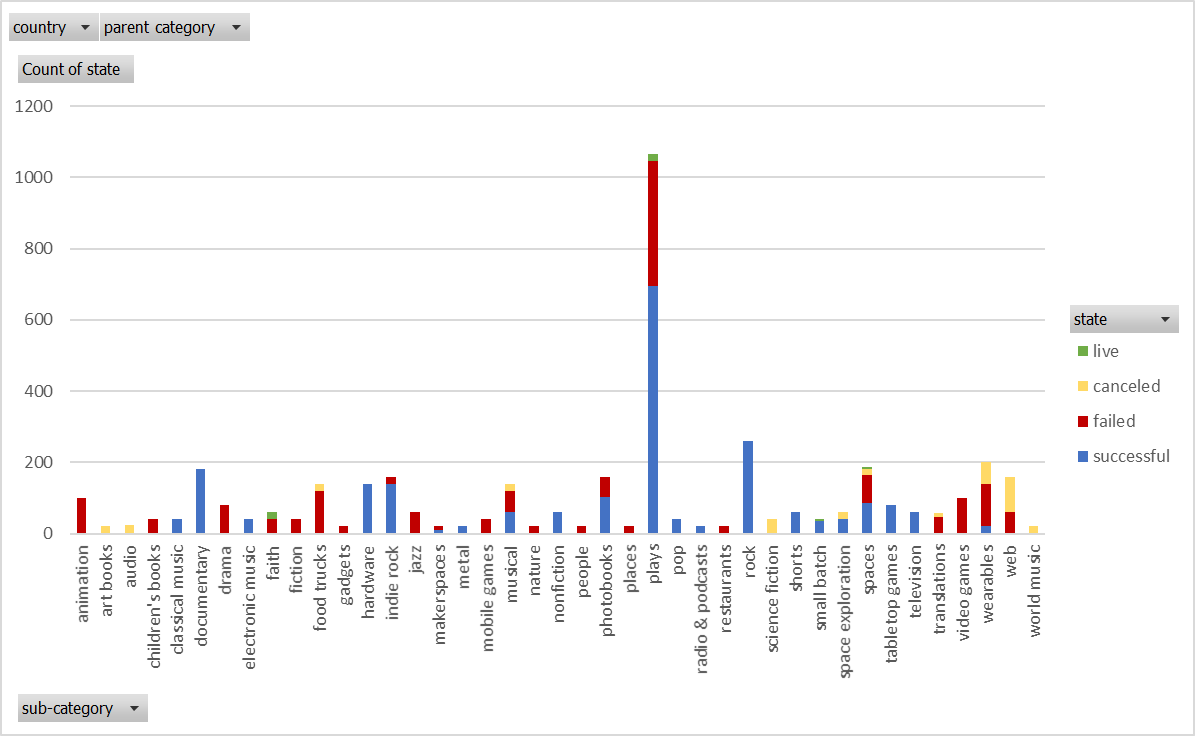


Figure 3. Count of successful, failed, cancelled and live projects per **sub-category**, all countries pooled (4114 projects).

It is shown in Figure 4 a summary of the campaigns and their status in the four most popular categories: theater (33.9% of projects), music (17.0%), film & video (12.6%), and technology (14.6%).

|  |  |
| --- | --- |
| Theater (1393 projects) | Music (700 projects) |
| Film & Video (520 projects) | Technology (600 projects) |

Figure 4. Count of successful, failed cancelled and live projects per **sub-category**, in the general **categories** of theater, music, film & video, and technology.

It is observed how most subcategories have either very high or very low success rate, which can be better observed in Figure 5, which shows the percentage of successful, failed, canceled or live campaigns in each sub-category, rather than their count.

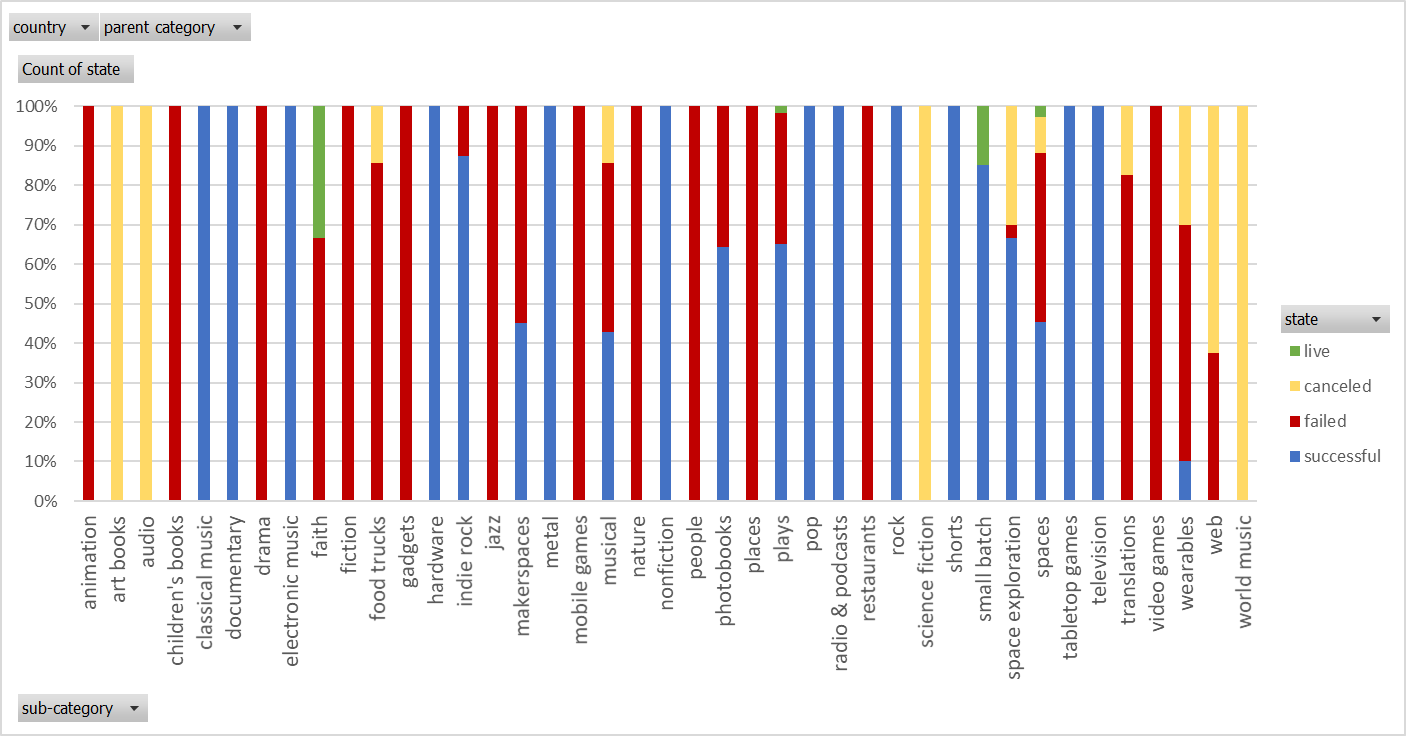


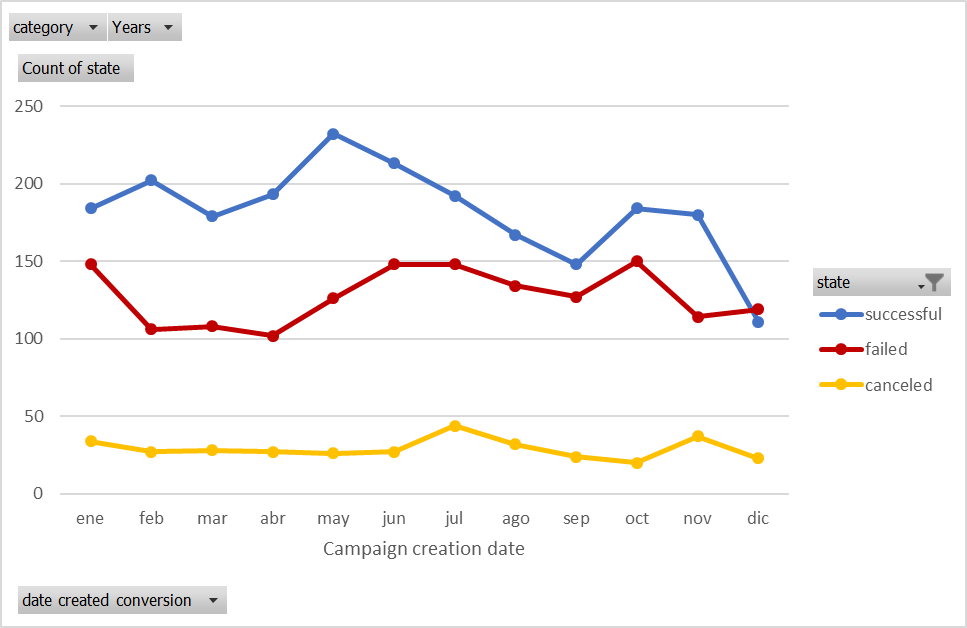
Figure 5. **Percentage** of successful, failed cancelled and live projects per **sub-category**.

## Observations

* Plays is shown to be the most popular sub-category, with more the one quarter of all projects presented.
* All sub-categories in theater are very competitive, with success rates from 43% to 65%.
* Most sub-categories in music, technology, and film & video, on the other hand, have more definitive success rates, meaning they are very high or very low:
  + 100% success rate: classical music, electronic music, metal, pop, rock, documentary, shorts, television, hardware.
  + High success: indie rock.
  + Low success rate: wearables.
  + 0% success rate: jazz, faith, world music, animation, drama, gadgets, web.

# Status of each campaign per launch date

The sheet called ‘**Status per launch date** contains a pivot table and chart, showing the number of successful, failed, and cancelled campaigns per launch date, which can be filtered by year and parent category. Figure 6 shows the count of successful, failed and cancelled per the date in which they were launched.



­ Figure 6. Count of successful, failed, and cancelled projects per **launch date (month)**, in all years and categories.

The previous chart shows a compilation of the status of the campaigns launched in each month from 2009 until the first trimester of 2017. There seems to be a slight trend of increased number of successful projects in May, and lower success rate in December.

However, over the course of eight years, it is likely that the trend might change. The same graph was explored filtered per year, and Figure 7 shows a summary of the same graph every two years. It is very clear that the number of projects has increased significantly since the first years of Kickstarter. An interesting observation is that in previous years, there were far more successful than failed projects. With an increase in the number of campaigns, there seems to be a decrease in the success rate.

|  |  |
| --- | --- |
| 2016 (949 campaigns) | 2014 (977 campaigns) |
| 2012 (282 campaigns) | 2010 (65 campaigns) |

Figure 7. Count of successful, failed, and cancelled projects per **launch month**, in 2016, 2014, 2012 and 2010.

When exploring the temporal trend on each category, there was not a clear trend for most of them, except for one. For theater, the most populated category, there is a clear increase in the number of successful projects in the month of May, and a decrease in the month of December, which is showed on Figure 8.

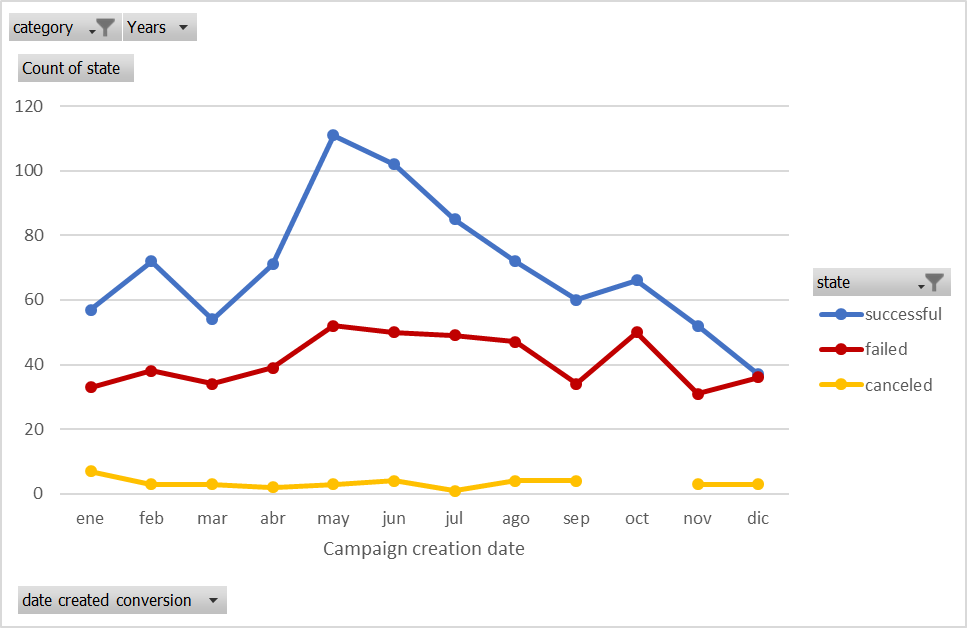


Figure 8. Count of successful, failed, and cancelled projects per **launch date (month)**, in Theater.

## Observations

* There is an increase in the number of successful projects in May, and a decrease in December.
* The number of projects has increased significantly over the years. This increase has been accompanied by a decrease in the success rate.
* For most categories, there was no discernible temporal trend, but for theater there is a clear increase in success in May with a decrease in December.

# Extra: Outcome based on goal

The sheet called ‘**Outcome based on goal’** contains a table on success rate depending on the goal category. Figure 8 show a summary of such table.

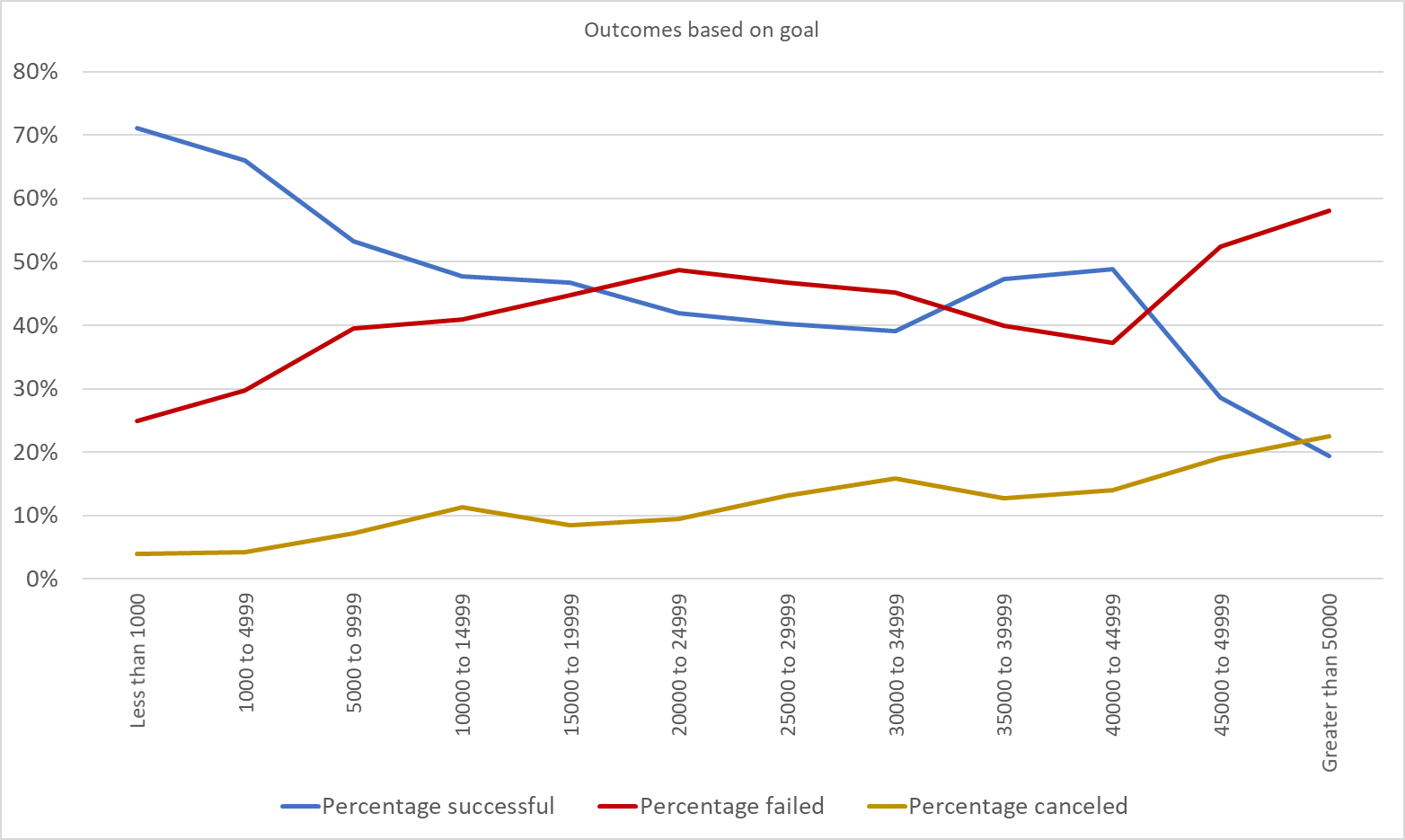


Figure 8. Success rate of campaigns according the amount pledged.

## Observations

* There is a decrease of success with increasing goal. Such observation is intuitive, since the more ambitions the goal, the more difficult it is to pledge such amount of money.
* There are, however, several successful campaigns that pledged amounts much larger than &50,000 – the most successful project pledged over 2 million dollars. The most ambitious projects need to be much more engaging to meet their goal.
* There is an increase in failed projects with increasing goal. Once more, this is intuitive, since the largest the goal, the harder it is to attain.
* There is also an increase in cancelled projects with increasing goal. Kickstarter cancels those projects considered to be a fraud. It is likely that there are more fraudulent projects with higher goals, or that the incentive to catch those frauds is larger and therefore Kickstarter focuses on finding them.

# Conclusions

Based on previously listed observations, some relevant conclusions can be drawn from this exercise:

* As Kickstarter has become more popular, an increasing number of projects has meant a decrease in success rate. Therefore, an analysis of influencing factors has become more valuable.
* Sub-category is a very relevant aspect to analyze, since most of them have either a low or high success rate, providing a reasonable estimate of future success (Do not start a campaign in Jazz!).
* The more ambitious the goal of a campaign, the more difficult it is to attain. Therefore, care must be taken to set an appropriate goal of each project.

# Limitations of dataset

* The current dataset has only 4000 points, much less than the total number of Kickstart campaigns, which could lead to distortions in the observed trends. In fact, the success rate of this dataset is roughly 50%, while the overall success rate is 36%.
* Although the initial and final date of campaigns is given, it would also be relevant to know how fast successful projects accomplished their goal.
* Most data come from three countries, the limited information given for other countries means that any conclusion drawn for them is likely very weak.
* The number of campaigns grows every year, and with an increasing number of projects, the success rate decreases. The trends could be changing quickly from one year to the next one.

# Other possible analysis

* Average number of backers and average donation in every category (My hypothesis is that campaigns with high number of backers are more successful than those with high average donation).
* Time series of success rate (instead of count of successful projects) in every category of sub-category.
* Success rate of every sub-category depending on the range of money they asked for, in order to find out if there is more likely to succeed with in certain range.