**BUSINESS MEMO**

Evaluating the dataset available on Kickstarter reporting 4000 past and current projects, Kickstarters can submit their proposals to get financed by backers in 9 categories. Each category has many sub-categories inside.

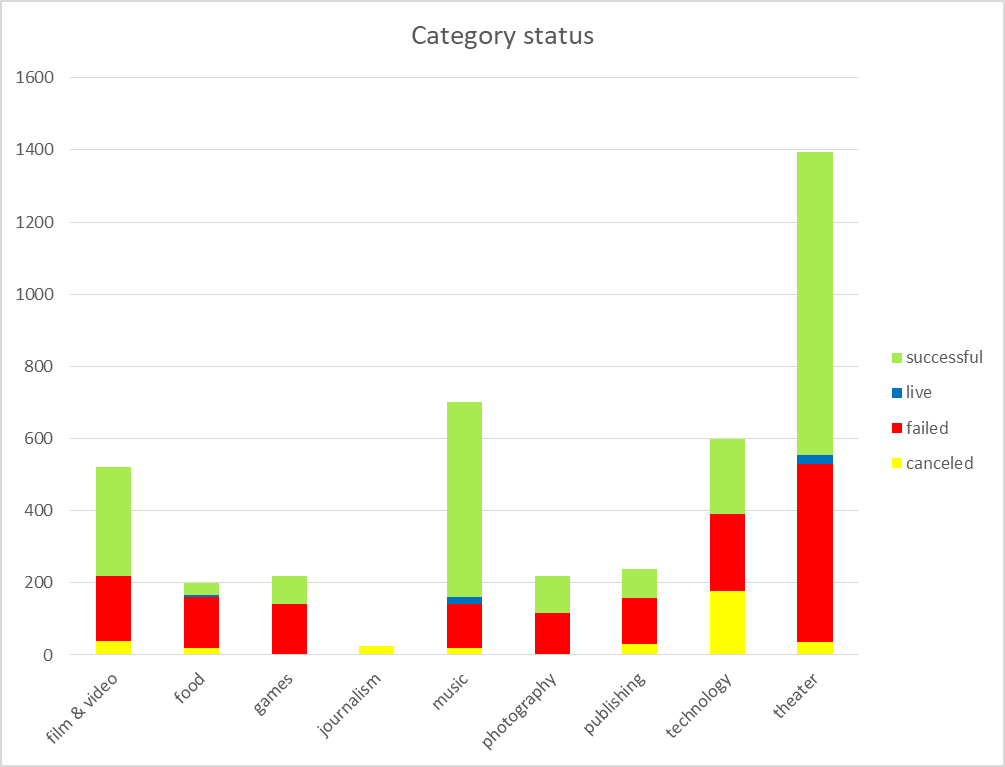
At a glance, we can say that Kickstarter seems an efficient mean to retrieve capital based on the total number of campaigns:

54% positive chance of getting funded

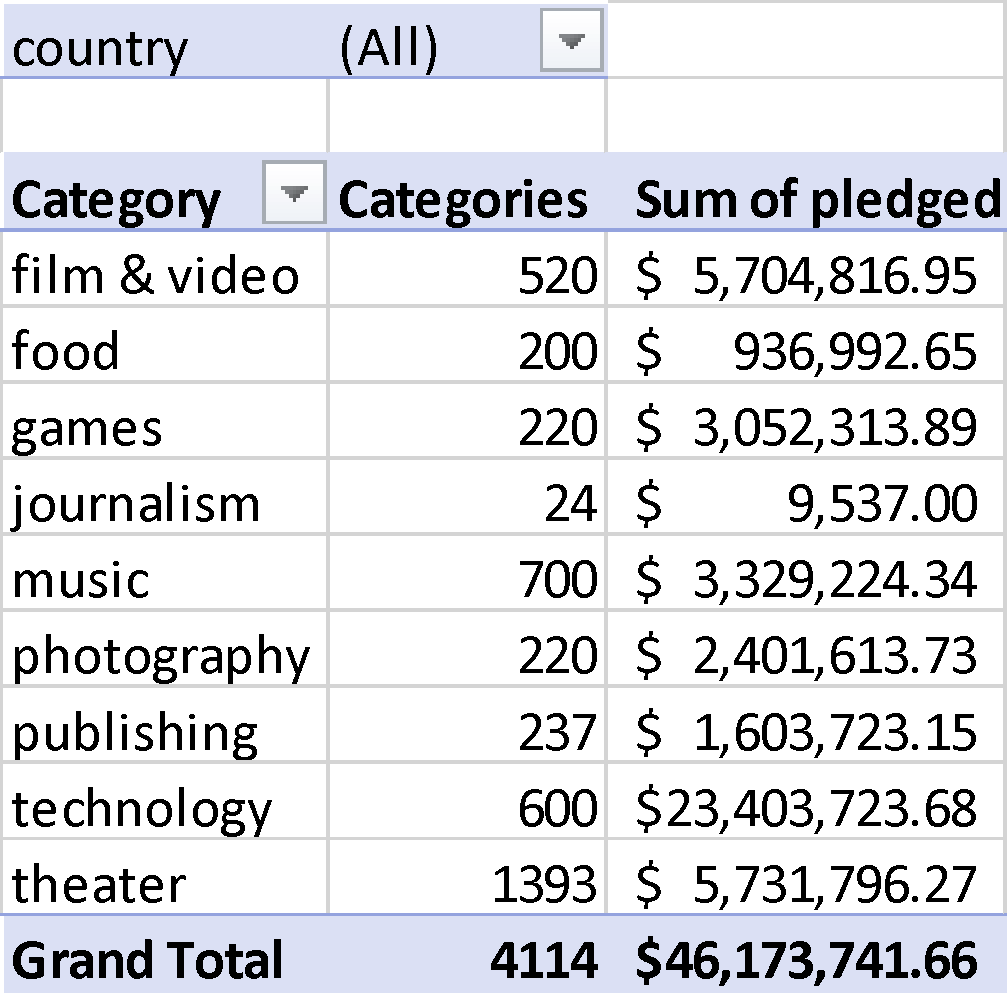
46% of not getting funded (between canceled and failed campaign).

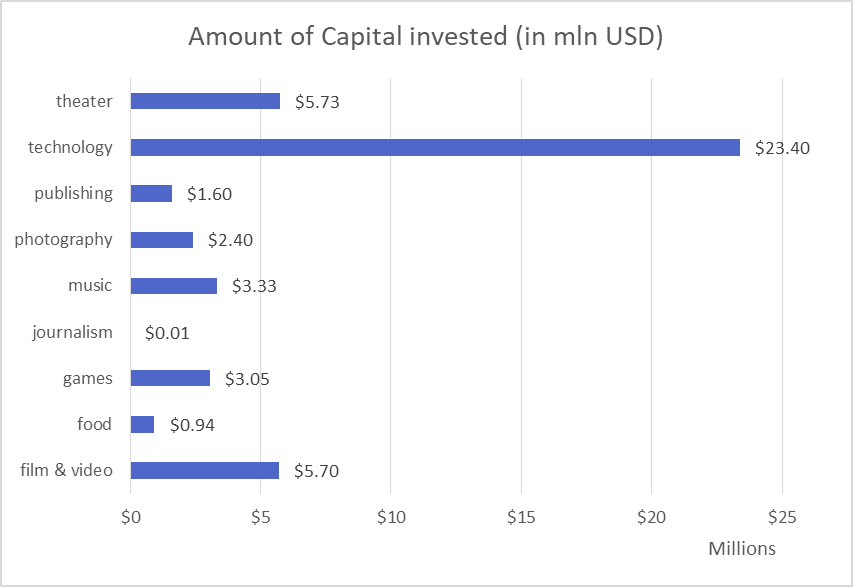
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| country | (All) |  |  |  |  |
|  |  |  |  |  |  |
| **Categories** | **state** |  |  |  |  |
| **Category** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| film & video | 40 | 180 |  | 300 | 520 |
| food | 20 | 140 | 6 | 34 | 200 |
| games |  | 140 |  | 80 | 220 |
| journalism | 24 |  |  |  | 24 |
| music | 20 | 120 | 20 | 540 | 700 |
| photography |  | 117 |  | 103 | 220 |
| publishing | 30 | 127 |  | 80 | 237 |
| technology | 178 | 213 |  | 209 | 600 |
| theater | 37 | 493 | 24 | 839 | 1393 |
| **Grand Total** | **349** | **1530** | **50** | **2185** | **4114** |

Also, the category with the biggest number of campaigns is ‘theater’, and it is also the category with the most ‘successful’ campaigns (60% of the campaign turned positive), followed by ‘music’. However, it is also the category that has the most numerous ‘failed’ campaigns.



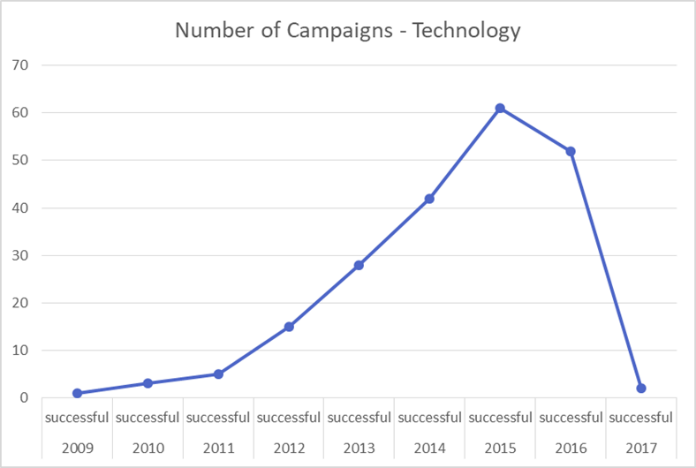
Analyzing closer this result, we should ask ourselves if it is really the number of the campaign a standard of quality? It doesn’t look like this. Data tells us that the category able to get more financing capital was by far ‘technology’ (here below in millions of dollars).

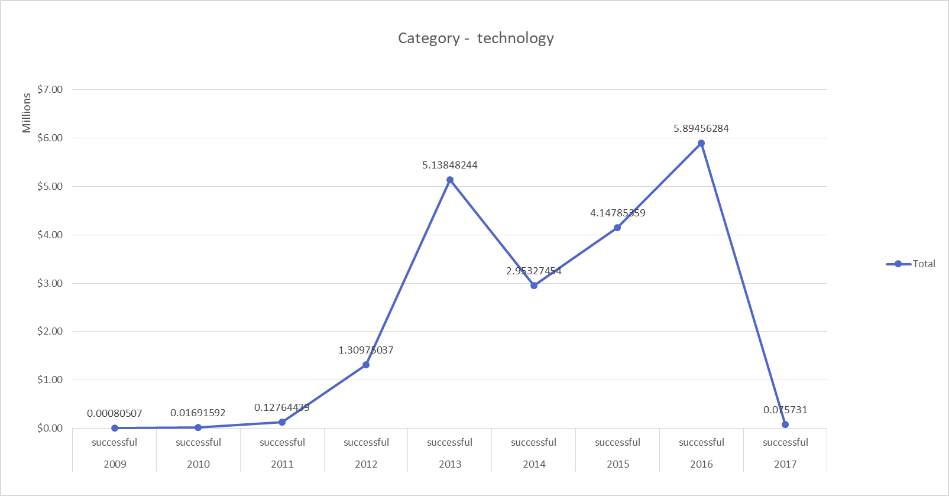




Also, taking into consideration technology, we can definitely say that the mere number of campaigns is not a quality factory if we analyze the trend pattern over the years.

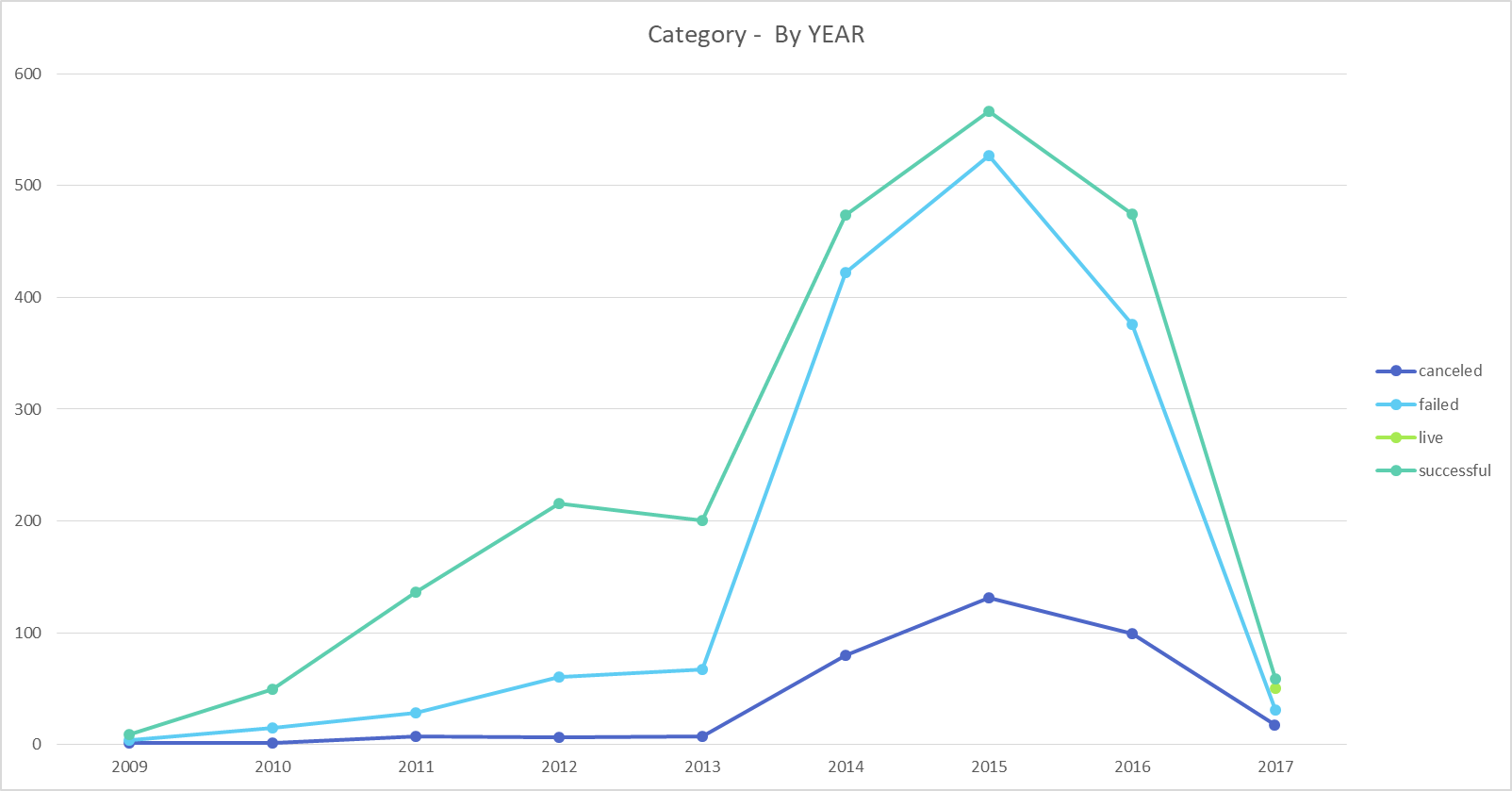
2015 and 2016 were the years where technology had more campaigns in terms of amount, but if we consider the invested capital in this category the best years were 2013 and 2016.

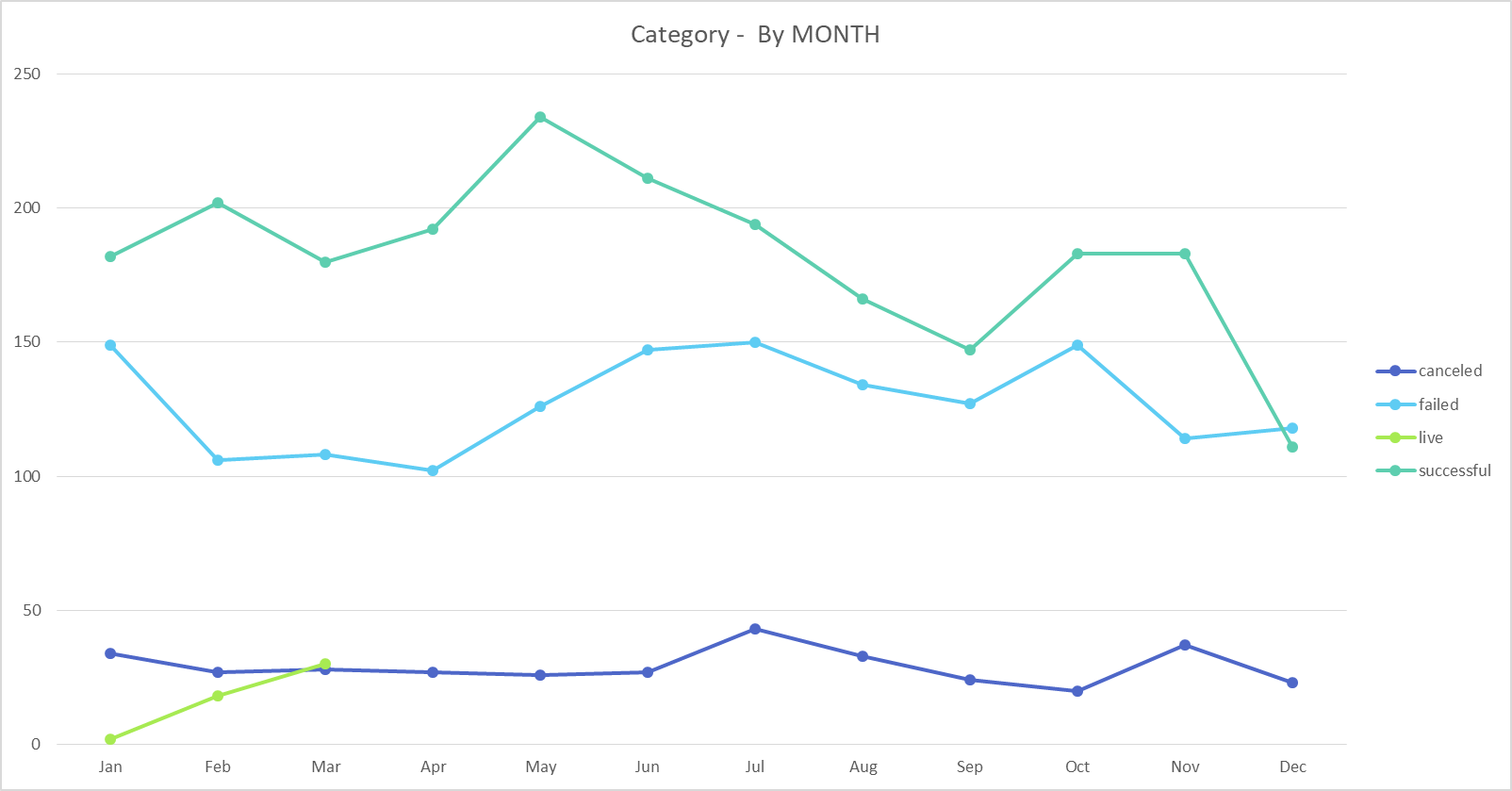




Again, analyzing paths over the years, we can see how all the categories picked in volume in 2015 (successful, failed, and canceled), and slowed down after peaking in 2015.

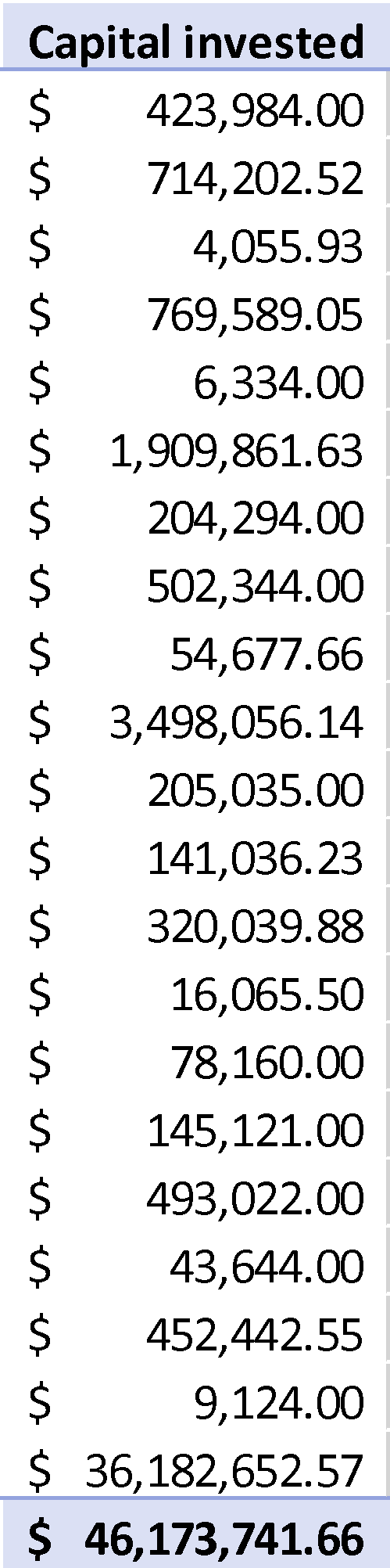
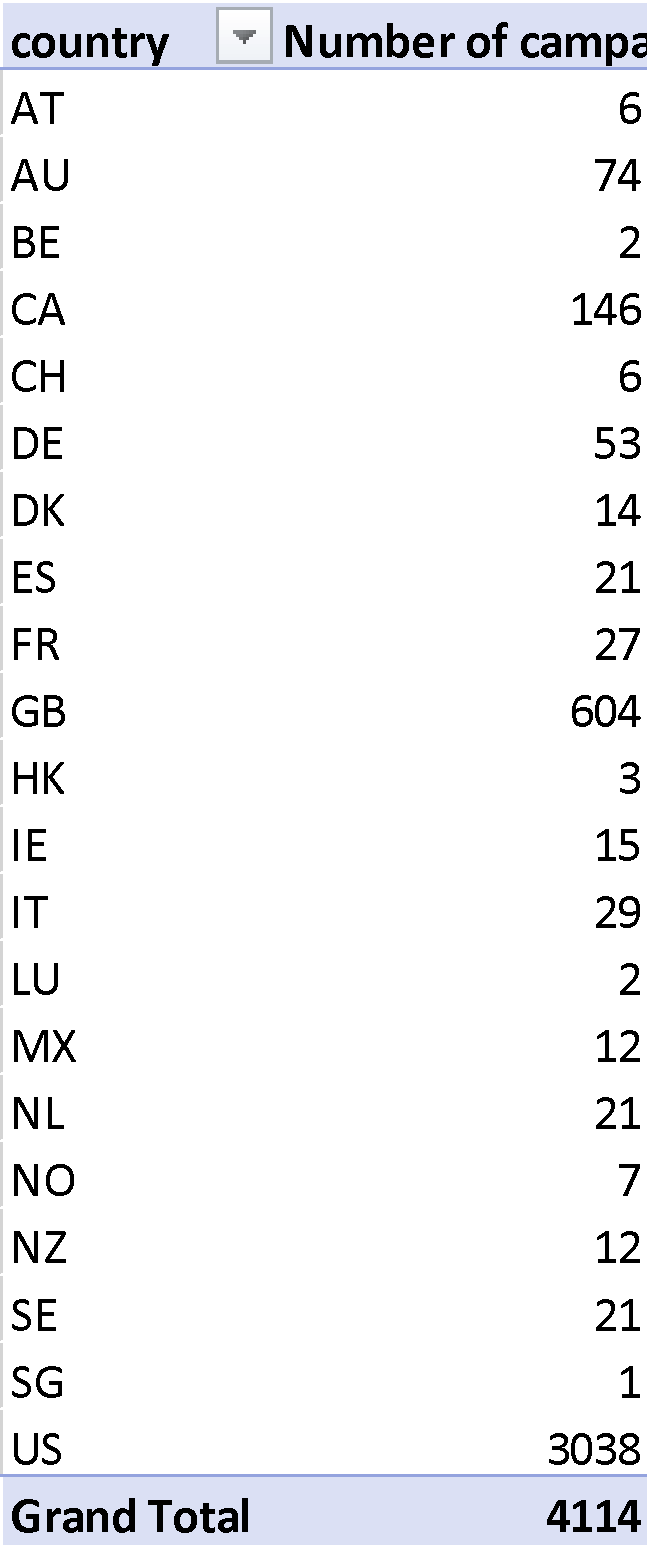
If we consider the month as a parameter of when it is more effective for the business to ask for funding, better submit in Q1 and Q2 (more consistent successful campaigns).

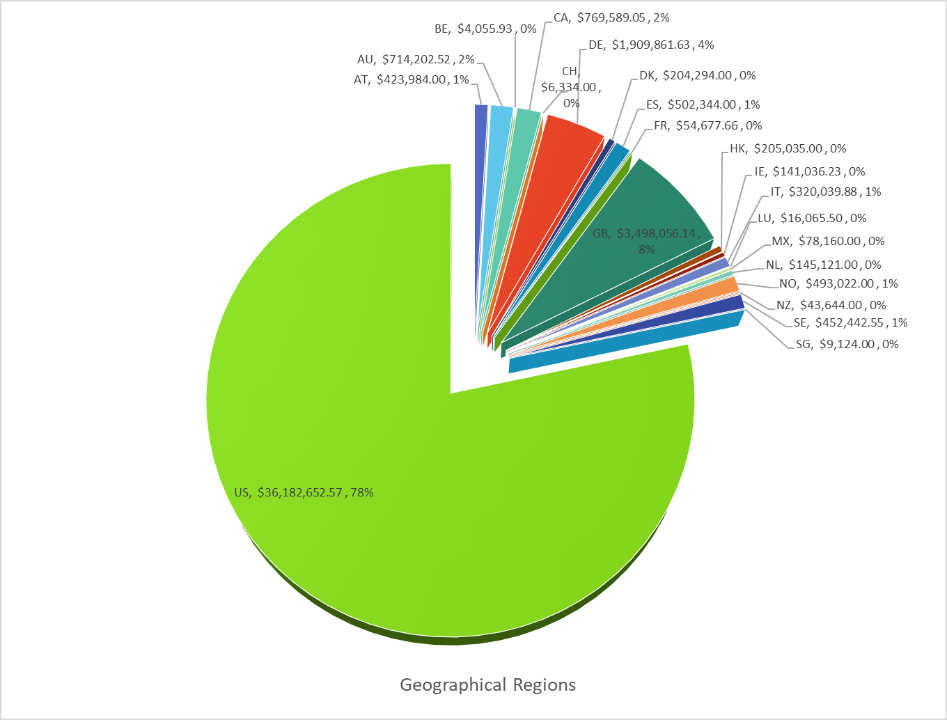


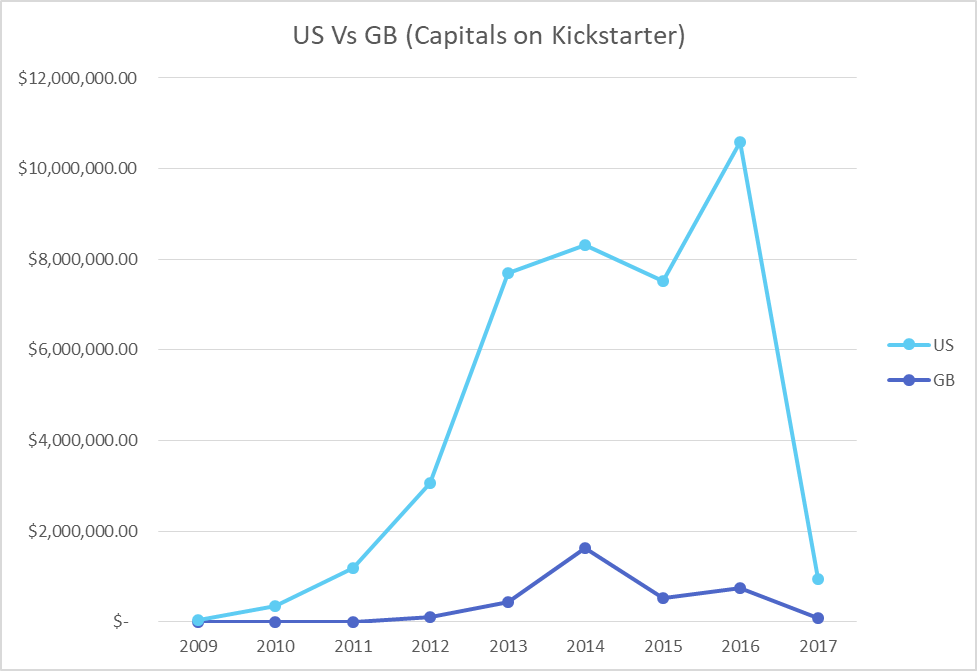


Another important consideration is related to the geographical region: the US alone not only is the country where most of the campaigns have been launched but also where almost 80% of the capital was raised in the examinated years.

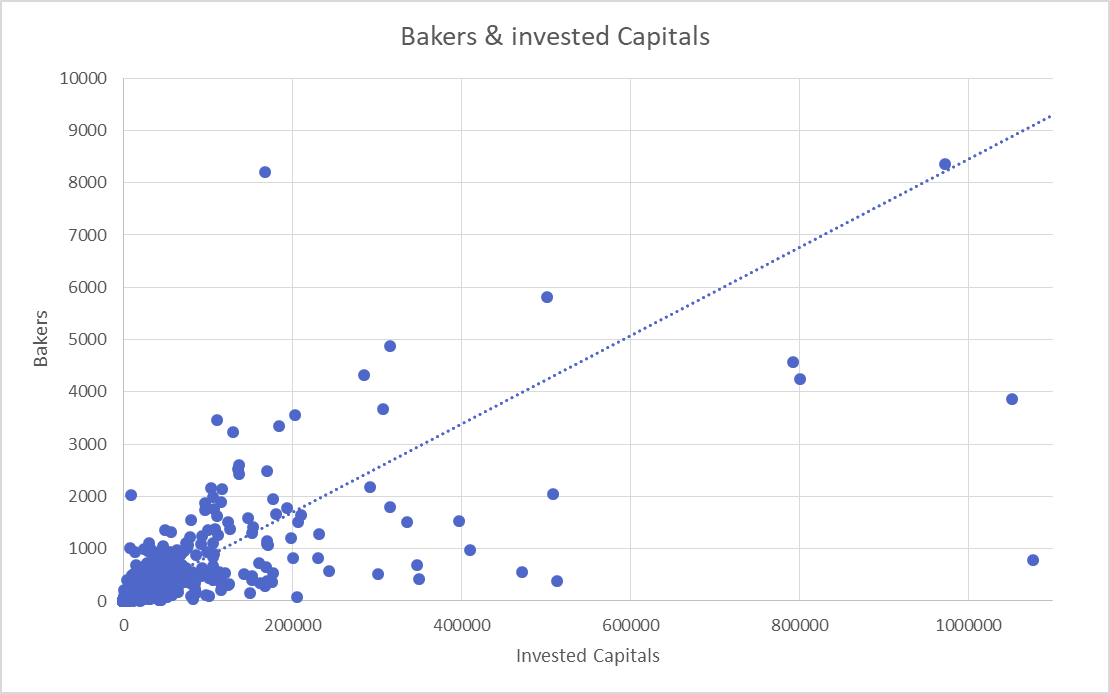
The second is GB. Analyzing further the years, we discover that the US has seen a substantial increase in the use of this mean from 2013 and 2016, while GB stayed almost stable (only with a peak in 2014).







Moving on, we can infer that the number of backers can be seen as a metric of success of the campaign itself for a crowdfunding campaign. A scatterplot can prove it. If the number of backers grows, also the invested capitals seem to be growing. The more backers invest, the more chances the kickstarters have their campaign to be successful.



As of last, continuing to evaluate the backers’ data and considering the outcome of successful campaigns and failed campaigns, we can infer that the MEAN does not summarize the data in the most meaningful way.

That is done considering the MEDIAN and the MODE. The median and mode values are in both datasets lower than the mean values (being the mode the most frequent value in a dataset).

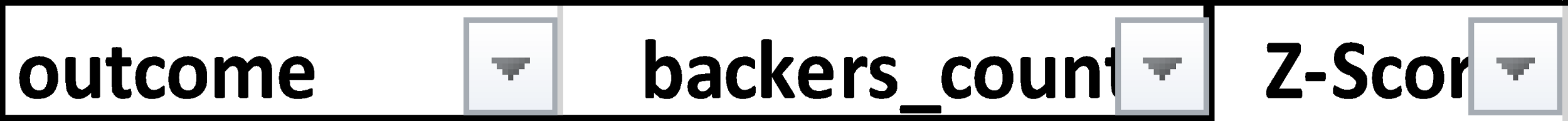
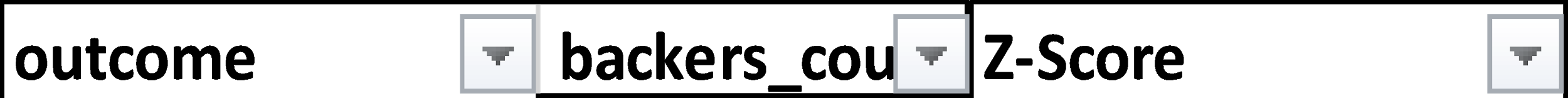
The maximum values are respectively high in both datasets as well.

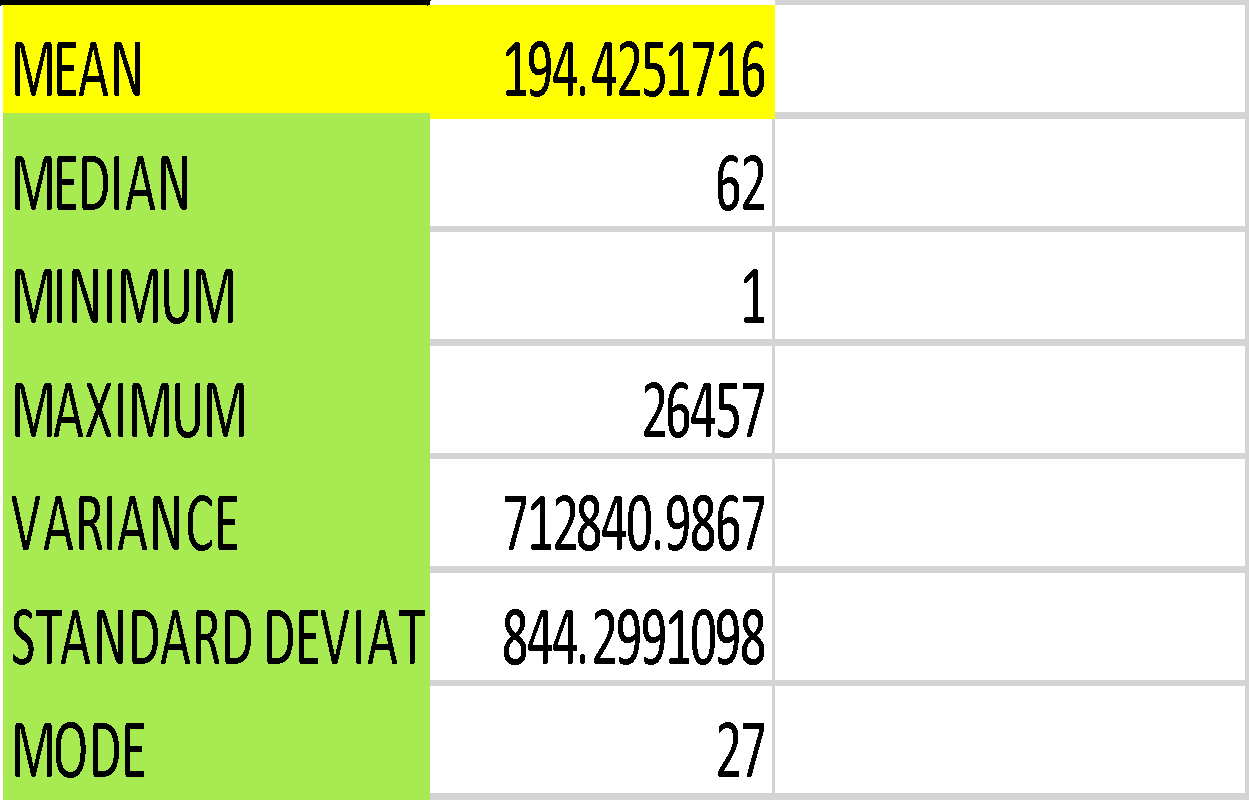
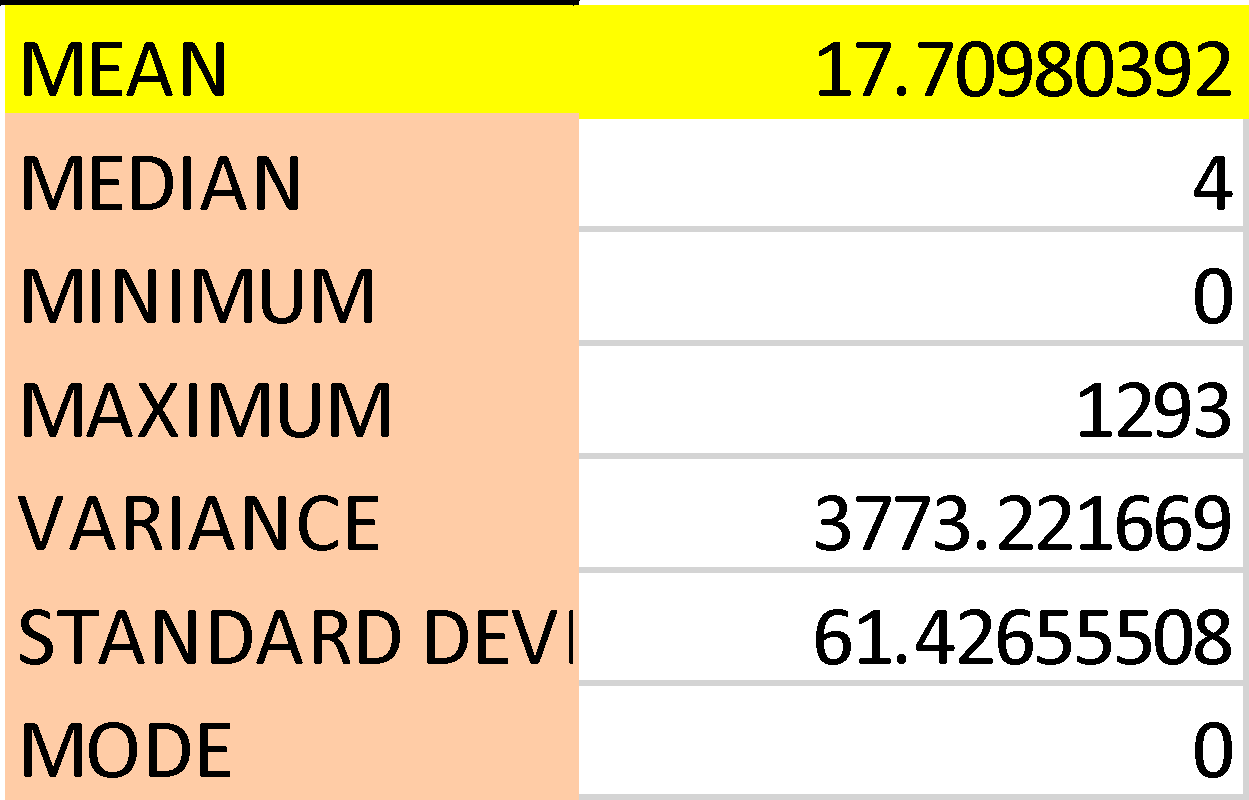
Going to analyze the variance and the standard deviation values, we also see that there is a high value of dispersion from the mean in both the datasets. We can conclude that there are a number of high values in both datasets that bring our average to be high as well and skew our analysis unless considering median and mode.

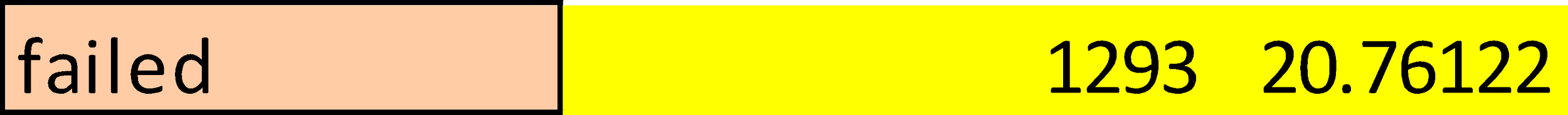
Looking at the data, there is more variability in the number of successful campaigns. This dataset is more spread out compared to the failed campaign dataset.

I would infer that this makes sense as a measurement of ‘reassurance-reinforcement’. People tend to believe more in things (ideas and businesses) if they see other people believe in those. If a campaign gets many investors in the first period of being launched, it is more likely to get more backers as proof of quality that the product/service will be successful. If the campaign doesn’t get traction in the first days of the launch, I would say that it is likely to not work at all (which is why the number of backers doesn’t get into the thousands).

Being Median ¡= Mode ¡= Mean and after doing two quick histograms, we can say that the 2 datasets are not normally distributed (the columns don’t form a Gaussian curve).





There are three limitations to this dataset.

* The first one resides in the inconsistency in currencies and lacking information of exchange rates. This chart reports different currencies without either giving me exchange rates to uniform the data or telling me which exchange rate to adopt for each of them.
  + Ways to solve the impasse: I could have looked online and transformed all entries using the available rates, but which ones and from which entity? Then, I could have taken today’s value or the value of the last 3 months (on average). This approach is known as “Constant exchange rates”, which be calculated in numerous ways. One approach is to convert current numbers using the prior period’s average exchange rate (the 3 month-value). The other one is to adjust previous numbers to reflect the current year’s exchange rate. In all cases, the set of figures that investors look at to see how trading has improved relative to the [comparative](https://www.investopedia.com/terms/c/comparative-statement.asp) period will no longer be distorted by foreign currency swings. Working in corporate for some years - especially in companies after IPOs -, I know these rates should be decided at first or given by an authority (to avoid audit issues). For the sake of time, I arbitrary considered all entries as dollars as all decisions I would make would have likely been inconsistent.
* The second issue and third issues are more on the qualitative side: there are campaigns lasting a month, some last lasting some days. Is it just a decision due to the campaign not reaching the goal? Is it on the kickstarter to decide to close the campaign or decision is made by the website itself? This dataset seems to trigger more questions instead of solving them.
* What is the true difference between canceled and failed campaigns? Just withdrawal made by the kickstarter or there is an internal kickstarter authority that cancel campaigns not meeting requirements? If so, what these requirements would be? This information would be necessary in case of a startup seeking investors and funds. We could not afford to go for assumptions. The end of a business evaluation should be always “Is there an added value in this?” This data set doesn’t exstensively answer my question and triggers more research.