1. Given the provided data, what are three conclusions we can draw about crowdfunding campaigns?
   1. Successful campaigns peaked in Summertime
   2. Failed campaigns average 37% monthly
   3. Cancelled campaigns remain relatively low on a monthly basis
   4. Success vs failed campaigns are complimentary, exclude the first quarter where the cancelled campaigns increase
2. What are some limitations of this dataset?
   1. The dataset is drawn from different countries with different currencies; with exchange rates in mind, the data is not standardised with a comparable unit.
   2. Data is not sufficiently granular therefore, you may not draw a reliable conclusion as
      1. the geographic range is too large, and sample quantity sample is too small.
      2. the dataset is drawn for a 10-year period with only 1000 samples, this is incredibly small to compare across multiple countries.
      3. although categories listed sufficient as parent and sub-categories, the confidence level of the result is not expected to be high as the sample size is still too small (1,000 campaigns).
   3. As “most people would use the number of campaign backers to assess the success of the campaign”, this dataset was assessed based on the goal fulfilment as to whether the campaign succeeded.
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   1. Compare the parent categories with goal amounts to see an average of how much each category of campaign tends to set as goals
   2. Compare parent categories with backers to see how much attraction each category receives
   3. Compare the average goal and pledge by category, to see if the success rate is related to the category or the goal cost being expensive to fund.
   4. Staff pick and spotlight columns weren’t used, they are a Boolean data type which can be correlated against the outcome rates. Correlation of outcome with these two factors could help explain causality.   
      How having a staff picked/spotlighted campaign helped the success outcome rate?
      1. staff pick - only people on the platform picked the project
      2. spotlight – featured on the platform  
         How having a staff picked/spotlighted campaign helped increase backers? In turn potentially increasing outcome success rate.
   5. Currency standardising to accurately measure backer contribution/pledge and all other comparisons involving the goal and pledge quantity.
   6. Correlate an average pledge to backers per campaign will show backers’ average spend
   7. Consider high investing backers

Bonus Statistical Analysis

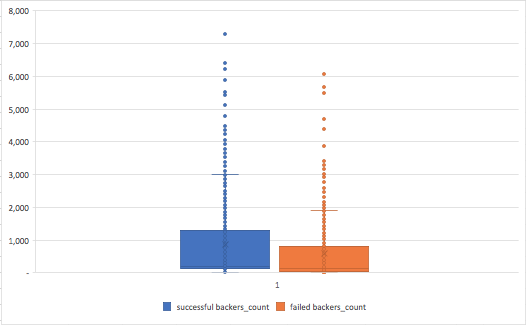
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Backers -->** | **Mean** | **Median** | **Minimum** | **Maximum** | **Variance** | **Standard Deviation** |
| Successful | 851 | 201 | 16 | 7,295 | 1,603,373.73 | 1,267.37 |
| Failed | 586 | 115 | 0 | 6,080 | 921,574.68 | 961.31 |

Use your data to determine whether the mean or the median summarises the data more meaningfully.

Considering the high variance of the minimum and maximum number of backers in the data range, the mean value of both successful and failed campaigns would seem most appropriate to represent the data meaningfully.

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

From a data sample of 1,000 campaigns, 566 succeeded whilst 364 campaigns failed. We can see that the average amount of backers in failed campaigns is 586. This could interpret as an upper boundary for the number of backers required to have a successful campaign, anything below this boundary can be seen as having a lower chance of success. However, successful campaigns could have an unlimited number of backers creating a larger variance and average, so the dispersion for successful campaigns is unbounded.



On the other hand, comparing the maximum values of successful and unsuccessful campaigns seems to tell a different tale. Successful campaigns have a maximum number of backers of 7,295 whilst failed campaigns reached 6,080, and still failed. This isn’t quite representative of the previous statement but could be caused by exceptional conditions. As this data set’s outcome is based on whether or not the campaign reaches their goal with backer investments, the extremity of the maximum backer in this specific failed campaign could either be due to the high goal the campaign wished to achieve or the contributions towards the overall pledge were insufficient despite having over 6,000 backers.