**Crowdfunding Data**

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

* From the pivot created to show the outcome by the month the project was created, across all data points June and July have the highest success rate.
* The US contains the most projects created, attributed to more than 76% of the dataset. The US has a success rate of 57.14% and a failure rate of 36%, in contrast to CH which has the least number of projects created (23 or 2.3% of the dataset) with a success rate of 52.17% and a failure rate of 26%.
* Within the sub-category pivot, you can clearly see that “plays” has a significantly higher amount of projects created compared to the other sub-categories, but “web” has the highest success rate, which can be seen by the larger portion of green bar compared to red.
* From the “Crowdfunding Goal Analysis” a clear metric that can be picked out is that a project with a funding goal between 15,000 and 35,000 has a much higher chance of success than those with goals greater or less than this range.

What are some limitations of this dataset?

* **Size of the dataset** – Whilst 1000 rows is a good size, Kickstarter alone have had 600,175[[1]](#footnote-1) projects launched on their platform which means this dataset accounts for only 0.16% of those projects. A much larger dataset could support a more in-depth analysis.
* **Date range** – Having a wide dataset in terms of date can be advantageous to view historical trends throughout the years. However, in this scenario it can skew the data results. Kickstarter launched in April 2009[[2]](#footnote-2) so the data on projects which began in 2010 may not have been successful due to the lack of knowledge around the Kickstarter platform. As time has progressed and the platform has grown to be globally recognised this could contribute to a higher success rate in later years.

What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

* Other possible graphs could be simply changing the created stacked column charts to “100% Stacked column” which give a much better visual on the difference between each Series for the X Axis. The value this adds is seeing the % breakdown of each Series relative to the total number of projects for each category.
* Another PivotChart and PivotTable to bring value would be to evaluate the Outcomes by Year of creation and using a line chart. This will show how, as awareness to platforms like Kickstarter and Indiegogo has grown, the trends of successful and failed projects have increased or decreased.
* Lastly, an interesting data point to look at would be the Success, failure and canceled counts for projects by which have a “TRUE” value in either the “staff\_pick” or “spotlight” column, compared to those with a “FALSE” value. I think this could really show how a project being highlighted within specific areas of a platform’s website can increase the chances of a successful project.

**Statistical Analysis**

Use your data to determine whether the mean or the median better summarizes the data.

* In my opinion, the mean summarizes the data better as it’s much more centrally positioned (central tendency) in relation to the minimum and maximum values of both the “Successful” and “Failed” backer statistics.

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

* There’s more variability with successful campaigns over failed campaigns. This is demonstrated by the “Variance” value in the Successful category at 1,603,373 compared to Failed at 921,574, and the Standard Deviation of 1,266 for successful and 960 for failed. This makes sense for a few reasons:
  + There are 201 more campaigns in the successful column, which means there are more opportunities for a greater range of values.
  + There’s a greater difference between the Minimum and Maximum values for the backer statistics in the Successful category.
  + As a campaign gains momentum with the number of backers or the amount pledged, this would have a positive effect on gaining additional support to continually grow the count of backers.

1. <https://www.kickstarter.com/help/stats?ref=global-footer> – Kickstarter Stats (Accessed: 7/8/23) [↑](#footnote-ref-1)
2. <https://www.kickstarter.com/about?ref=global-footer> – Kickstarter Launch Date (Accessed: 7/8/23) [↑](#footnote-ref-2)