Analysis Report

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

The US is the country with the most crowdfunding campaigns (>75%). The theatre is the category with the most campaigns overall and all the countries except for China where music is the category with the most campaigns. Film & video and music are in second and third with a very close number between themselves but still far from the number of theatre campaigns. The percentages of successful campaigns between the three categories are very similar in the range of 54-57%. Theather campaign's performance in most of the months is positive but in may and august is possible to visualize that the number of failed campaigns is higher than the successful ones.

Journalism is the category with the least campaigns but with a 100% success rate and the campaigns were made in the US, with 2 campaigns in 2014 and 2 in 2019 being September the month with the most campaigns.

Food and games are the 2 categories with the highest number of failed campaigns, with games being the worst overall. In the games category, we have mobile and video games. Video games have more campaigns and 43% of failed projects, mobile games have 62% of failed projects. And the peak with most projects was the month of August.

1. What are some limitations of this dataset?

The biggest limitation is that we cannot see the reason for the failure or success of the campaigns. If the reason is that the campaign has a high target goal; and in categories with a small number of projects and a high number of success (ex. journalism with a 100% success rate) if the category is, in reality, popular or if was just that those projects were doomed to success.

It’s also impossible to know which categories and sub-categories exist since the beginning of crowdfunding and wich ones were added later. This can affect how the data presents itself for analysis.

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

I believe we could have entered one more pivot table and chart like we did on the bonus task but making a correlation between target goal, number of brackets, campaigns categories and sub categories. That would show the if the success of the campaign were due to low targets or popularity.

Bonus Statistical Analysis

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

Mean being the most common value in a collection of number and subject to every change in the data sheet is more meaningfully for data analysis than median. Being median only applicable to separate the higher half from the lower half on a data sheet and relate to only one or two number in the middle of a data sheet.

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

The standard deviation and variance are higher between the successful campaigns. I believe it makes sense based on the fact that we have a higher number of successful campaigns and that way the number will have a higher range of variance and standard variance how far each value lies from the [mean](https://www.scribbr.com/statistics/mean/).