Excel-Challenge: Data Report

By Oliver Zagorin

Crowdfunding sites like Kickstarter have been a growing trend in recent years, with individuals funding everything from inventions to events, to health care costs. As crowdfunding is becoming a significant part of the social fabric of the modern age, they become an unavoidable aspect of society and thus draw the attention of analysts hoping to glean insight about how these projects succeed or fail. In this analysis, we will look at a dataset of 999 crowdfunding efforts, to see what are the keys to success or the foreshadowing of failure, we will examine the limitations of our analysis, and we will suggest a direction for future analyses. This will provide insight to individuals who are attempting to engage with crowdfunding either as a donor, or as someone who aspires to fund a project through these crowds.

This dataset provided several key insights that will be invaluable to those engaging with crowdfunding. First, we can see that the most popular parent category for projects is theater followed by film & video and music. These parent categories have similar proportions of success and failure, which, combined with their large number means that people are putting a lot of their funding towards these projects. The most popular subcategory is plays, which have about 7 times more projects than the average program. The vast majority of these plays take place in the US. We can see that between the months of May and July there is a decrease in the number of failed projects and an increase in the number of successful ones, suggesting that this is the optimal time to launch a project. And finally, we saw an apparent relationship between success rate and goal amount. With projects with a goal between $5000 and $9999 having the lowest success rate and projects with a goal between $10,000 and $24,999 having 100% success rate as well as the interval of $30000 to $34999. This suggests that this goal range of $5000 to $9999 should be avoided, while more projects should try to have goals of $10,000 to $24,999 or $30000 to $34999.

When we examined the content of this dataset, we did notice several limitations which limits the ability of these conclusions to be generalized across all settings. First, we noted that this dataset includes entries exclusively from 2010-2020, excluding 2021 and 2022 during which significant economic and geopolitical changes occurred, such as the war in Ukraine and the Tripledemic, which both impacted the crowdfunding community specifically. We also noted that a disproportionate amount of data was gathered from the US (over 70%). This means that the insights gleaned are much more likely to be true of US crowdfunding projects, and less applicable to non-US projects. Finally, we noted that there was a significant excess of theater and play related data compared to other categories. This could mask programs that have a very high success rate, but which are not represented in great number in this data set. We specifically noted that Audio projects and World Music projects had a 100% success rate, but that they only occurred in this dataset 4 and 3 times respectively.

If we were to run further analysis on this dataset, we would recommend several analyses which were not previously indicated. First, we would analyze the outcome of campaigns by the country where the campaign takes place and include an analysis of percentages of each outcome. This would allow for comparisons between the over-represented US projects and those originating from other countries. Next, we would compare the category where the campaign takes place in relation to the percentages of each outcome. And finally, we would compare the sub-category where the campaign takes place in relation to the percentages of each outcome. These last two analyses would limit the effect the different number of projects in each category, to create a more even comparison of outcome rates.