Delia Hellander

September 15, 2020

Excel Homework

1. The first pivot table and chart show the relationship between Parent Category and the success of the project. This data shows how theater is the most popular parent category for the projects sampled and therefore has a high success rate but also a high failure rate to meet their initial fundraising goals. One conclusion from this analysis is that although theater has the absolute highest number of successful projects, the music parent category has the best results comparatively since it has the highest percent of successful within the category compared to the distribution of success in all the other categories

The second pivot table and chart section display the relationship between the sub-category and the success of the projects within the sub-category. These displays made a few things clear: there are some sub-categories which all projects were successful, there are some that where a mix of canceled and failed, and there’s one subcategory that surpasses all in pure count of total projects and a high number of successes. Projects under the sub-category rock, hardware and documentary contain all successful projects. In contrast, sub-categories of web, world music, and video games are just a few of the many that comprise of only failed or canceled projects. Then there is the sub-category plays which has about 5 times the total number of projects compared to all the other sub-categories. The conclusion from this analysis is that there are some sub-categories that statistically show the high likelihood of failure and others that statistically should nearly always result in success. So, if you are looking for hidden trends that essentially guarantee success if only looking at the statistics of past projects, one would want to kickstart a project with the subcategory of rock, hardware, or documentary,

The third and final pivot table and chart show the relationship between the count of the status of projects over time. What is evident from the visual is that the patterns are consistent across the success, failed, and canceled line. All this tells us is that there are certain months in which the Kickstarter app increases its total number of projects, but this does not necessarily affect the distribution of success status. However, there is one take away if looking for hidden trends which is for the month of December count of success dips below failed. Another caveat to this display is that when you look at just one year of the data there is a lot of variability.

1. There are a couple limitation that come to mind which would directly affect the results of the task at hand which is to find hidden trends in Kickstarter project data. It is probable that there would be a strong relationship between regions within the US and distribution of success by category, however this dataset does not include data on regional location. The second pivot table analysis does include a filter by country but like I said, a regional look would provide interesting information since the regions in the US greatly vary and about 75% of the data is reported as from the US. Another example of important missing data variables the large hole in data in terms of long-term success. There are no factors or data that show profitability and what happens after that meet or fail to meet the initial goals. I am curious to see the data after meeting the initial goal and getting support and funding from the app. Maybe some categories can easily be successful in their initial goal but then struggle to continue to find success after that initial goal is achieved. In order to explore hidden trends in long-term success, more data is needed. Looking at factors like this could bring further insight into how to be successful on the Kickstarter app.
2. An additional part I would add to this analysis would be to focus on the money variables. I am interested to see how average donation varies across categories as well as look at which categories of projects raised the most money. This homework assignment only really looks at the success or failure of making it to the projects initial goal. This negates certain impacts. For example, projects set their initial goals too high, while others surpass their initial goal substantially. Looking at how this factor varies based on category, country, month etc. might give unique insight on certain trends revolving around monetary success and failure.

Bonus:

1. I would conclude that the median summarizes the data more accurately. The mean is too affected by major outliers.
2. The variance and standard deviation are a lot larger for successful campaigns rather than failed. This means there is greater variability in successful campaigns. This makes sense because the successful campaigns are determined by their initial goal. Some campaigns may have had much lower initial goals and therefore less backers while other campaigns had high goals and a high count of backers. In contrast the failed campaigns all failed pretty miserably. Seems like many of them failed to get momentum and therefore there was less disparity.