Anastasia Skrypnyk

Data Analytics Boot Camp

Week 1 Challenge Assignment

05/02/2023

1. Three conclusions that we can draw about crowdfunding campaigns based on the data are: crowdfunding campaigns in parent categories of journalism, technology, as well as photography were generally most successful by overall percentage of successful outcomes; the first, second and third quarters of the year were generally most profitable for project crowdfunding; and count of successful outcomes in each parent category varies distinctly across each country, as can be seen when filtering through the data by country code.
   1. When looking at Sheet 1 Pivot Table & Count of Outcome by Parent Category Graph, we can see a pattern of greater successful outcomes versus failed outcomes in categories of journalism, photography, and technology, wherein successful outcomes over grand total are highest by percentage & number. This pattern is also visible in the Count of Outcome by Parent Category Graph, wherein the greater portions of the bars are green, representing crowdfunding projects that were successfully funded per parent category. Journalism has a 100% overall success rate as 4/4 crowdfunding projects in that category were successful; technology comes up second with a 67% overall success rate since 64/96 projects were successful; and photography is third with a 62% overall success rate with 26/42 projects being successfully funded.
   2. When examining Sheet 3, we can see that the first 3 quarters of the year (Jan-Sept) saw the greatest number of crowdfunding projects, with a high of 257 by the end of the first quarter, and a low of 240 by the end of the fourth quarter. Looking directly at the Year & Category-Based Outcome Visualization graph, the peaks of the “successful” line on the graph are reached in March, June & July, that is, between the first and third quarters, suggesting a seasonal relationship between pledges made and time of year. Simultaneously, February sees the lowest number of successful and failed outcomes, suggesting a slower period for crowdfunding. September sees a sharp decrease in failed crowdfunding projects, while the number of successful projects plateus at this point, remaining relatively static and unchaning for the rest of the fourth quarter. As the crowdunding period draws to a close at the end of the fourth quarter, the number of canceled projects increases, suggesting they have come to the end of their crowdfunding period. As a result, we can expect a spike in the number of failed projects at the end of the fourth quarter, which is true as represented by the proportional climb of the light blue line in the graph as it increases almost at the same angle as the dark blue line. Overall, the relationship between the number of successful & failed projects is logically epxlicable when looking at the end of the second quarter and the beginning of the third: as the number of failed projects plunges around June, the number of succesful projects peaks right after, specifically in the beginning of the third quarter during the month of July.
   3. While overall the most succesful categories of crowdfunding projects are journalism, technology and photography, both parent and sub-categories of most-supported crowdfunding projects vary particuarly depending on country. When filterning the Count of Outcome by Parent Category & Count of Outcome by Sub-Category charts by country, we can see that categories like theatre, which is predominatly not successfully funded in Australia and Canada, is highly successful in being funded in European countries like Italy and Denmark. While technology is highly successful category in its crowdfinding amongst Australia, Canada, China, Great Britain and Italy, it is not so successful in Denmark and the U.S. While projects in the publishing category are largely well-funded in countries like Canada, China, and Denmark, they are utterly unsuccessful in Great Britain. Projects in the category of games, however, are thriving on the Kickstarter crowfunding platform in Denmark and Great Britain, where they are absolutely funded. This reveals particular relationships between certain preferred categories and specific cultural settings, while highlighting a clear common denominator in technology. This preference points to a global interest in supporting technological advancement -- via crowdfunding of such projects as belonging to this parent category.
2. Some limitations of this dataset include country list, which is not representative of a global scale to be globally relevant. Another one is its wide range of years, which compares a variety of projects across almost a decade (2010-2020), not taking into account external economic and political factors that could be impacting the success of crowdfunding at a particular moment in time. The lack of an actual qualitative report of the nature of each project rather than simply the quantitative values of crowdfunding campaigns: I.e, what kind of game is being crowdfunded, the author & volume of the written projects, the style or kitchen of the restaurants/eateries that are being crowfunded under the food category, the nature of journalism that is being publicly supported (in communist China, the nature of press & publishing is often politically restrictive and/or propagandisttic, unlike the nature of journalism in democratic Canada).
3. Other possible tables and/or graphs that we can create are pie charts that examine the percentage of successful, failed, live and canceled projects filtered by year & category. These charts would provide additional value by allowing us to determine success rate in percent-terms, enabling us to compare & analyze the success of various projects in a uniform way that can be easily presented to any potential investors or future backers interested in which category of projects to support depending on previous success, time and cultural setting.
4. Statistical Analysis:
5. The median better summarizes the data because the data is not symmetrically distributed and varies depending on a variety of factors, such as outcome, category, year, country, currency, etc.
6. According to the standard deviation calculations on Sheet 5, there is more variability with the successful campaigns. This could be the result of higher mean, median, minimum, maximum, and variance values, which can be attributed to the wide range of data and its large sample size.