Natalie Ericson

Module 1 Challenge

Crowdfunding Data Analysis Report

When presented with a large dataset such as this, there are many places we can look for potential indicators of crowdfunding campaign successes or failures. There are several conclusions to be drawn and factors to take into consideration when picking out the trends worth noting. Ultimately, this all comes back to our initial question – what is ‘the trick’ to successfully meeting or exceeding an initial funding goal? I’ve picked out what I believe to be the top three important conclusions we can draw that will help to answer this question. First, the highest rate of successful campaigns fell inside the funding goal range of $1,500 to $35,000. 92% of the campaigns inside this goal range were successful. This indicates there may be a ‘sweet spot’ in how much money the campaign aims to fundraise in order to meet or exceed those funding goals. Second, 70% of all crowdfunding campaigns fell under one of three categories: music, theater, or film and video. This suggests there is a large market for entertainment with many entities willing and ready to invest in this category. If the project is more desirable to potential backers, then there is a higher chance the project will see a successful donation campaign. 69% of the total campaigns within these three categories found success in their crowdfunding venture. Third, there does not appear to be a correlation between the likelihood of a campaign being cancelled and the total investment goal set by the project. The highest percentage of canceled projects fell inside three very different goal ranges: (1) $5,000-9,999, (2) $35,000 to 39,999, and (3) $50,000 or above. To a potential project coordinator, this could suggest that there may not be any benefit to setting a certain investment goal aimed at preventing the future cancelation of the campaign. This insight does not provide us with any external factors that may have also played into the final status of these canceled campaigns.

While the crowdfunding data set is large and appears to be very detailed, there are some ways in which it is limited toward helping answer our overarching question. When seeking to find ‘the trick’ to successful crowdfunding, it would be helpful to know how each campaign landed their backers. What platform was used by the project coordinators? There are many options out there, so how does the performance of Kickstarter members compare to Indiegogo or Crowdrise? In addition to the platform used to collect funding, did the project coordinators use any other kind of marketing strategy? We all know how influential social media these days is. Did successful campaigns use additional marketing methods to encourage donations while failed campaigns did not? Additionally, what kind of price points were offered to investors? Did backers have the option to make small investments, such as increments of $1, $5, or $10? Were incentives provided to backers that made significantly larger investments? We were able to calculate the average donation per backer from the original dataset, but we don’t see the method behind how this was achieved. Perhaps the is a correlation between successful campaigns and providing a wide range of donation options, as well as incentives for larger donations.

Pie charts are a classic and easy to read chart that’s helpful when looking at data that makes up percentages of a whole. Generating a pie chart could provide insight if looking to specifically evaluate what percentage of campaigns were successful, failed, cancelled, or live within a certain category or subcategory. This visualization would be useful for a potential campaign if they are trying to determine which category or subcategory they might have the most success in. We could also create a scatterplot for certain scenarios. A scatterplot would be most helpful if we’re evaluating a correlation, or lack thereof, between two variables. This would provide the most value if trying to determine a correlation between total amount pledged and average donation, or between funding goal and number of backers, for example. A potential campaign may find importance in determining if there’s a correlation between the funding goal and number of backers. Data such as this could help determine if the campaign is on track to be successful or not.

Statistical Analysis Summary

The mean best summarizes the data as it is more inclusive of the full range of the values rather than identifying the middle value of the dataset which is displayed as the median.

There is more variability in the number of backers per campaign with successful crowdfunding campaigns than with unsuccessful. The variance of the dataset for successful campaigns is 1,606,216.6 compared to the variance of the unsuccessful campaigns calculated at 92,4113.5. There are 201 more successful campaigns than unsuccessful, so it does make sense that there would be a higher variance in the number of backers for successful campaigns. With more campaigns falling under the successful category, there are more chances for the campaign to see a varying number of backers (from the mean) invest money into the project.