Hunter Craig

2/25/2023

Module 1 Analysis

Conclusions

I think one of the largest conclusions we can draw from this data set is by looking at it by categories, specifically parent categories. When looked at it shows the majority of successful campaigns occur in what would be considered entertainment, specifically in theater, film, and music. The only exception being games which received what could be deemed a “normal” amount of projects and successful campaigns. When looking at the aforementioned three theater comes out at nearly 350 total campaigns with one-hundred and eighty-seven of them being successful while film and music sit at about one-hundred and seventy-five campaigns and roughly a hundred of those being successes. I think this shows that while not everyone is interested in photography or technology, everyone requires entertainment of some kind, not very surprising. I do think it is interesting that theater was by far the most successful and had the most campaigns, I believe it shows that the demographic backing these campaigns are far more willing to give out money to projects they would like to see themselves. It also shows the dedication from the backer base to the medium despite how much more popular film and music is in everyday life.

After seeing how popular plays were in the data, I did some more digging into why there are so many successful campaigns. When looking at the average donation per supporter for each category plays once again came out on top at over twenty thousand dollars donated on average. This could show several different things. It could possibly show that more affluent people are willing to donate exorbitant amounts of money to theater projects because they are more likely to go and see a production. It could also show a small niche of people donating large sums of money to very specific projects in their area. For instance New York’s Broadway is known all over the world as a premier setting for theater, it could be possible that places such as this are attracting lots of attention and lots of revenue from affluent members of their nearby communities.

When looking at which countries are putting out the most campaigns the United States is leading the pack by a very large margin with the most successful campaigns at four-hundred and thirty-six. For reference the next largest contributor of successful campaigns is the United Kingdom at only twenty-eight. From this we can discern one of two things. Either the U.S is a country that has a much larger entrepreneurial spirit compared to the rest of the countries in the data set or that these crowdfunding platforms are almost exclusively used within the United States. I believe it is the latter. While the U.S has been known as a great place to start projects or businesses, it is not the only one on this data set that can say that. However speaking broadly, the majority of crowdfunding sites are not only based out of the U.S but they are also assumedly marketed the most in the U.S thus creating a massive disparity between it and its next largest contributor.

Limitations and Additions

In my opinion one of the largest limitations of this data set is not exploring where exactly in the U.S these projects were located. If that data could be collected I believe it would show a much clearer picture of who is starting these projects and possibly even where they are finding so much support. Generally plays stay in a specific location, whether that be one venue, city, or state. Generally, they aren’t available across the nation like film and music would be and especially not with the same cast and crew. This would likely show who is showing support for these projects. With that in mind the very high average donation for these theater projects could have some light shown on it. With the average donation being so high a box and whisker plot or a histogram could show possible outliers in the data set. I’m inclined to believe that a select few backers are carrying the average donation very high and these two visuals would go a long way to showing whether that belief is true or not.

Mean vs. Median

When looking at the mean and median of successful backers for both successful and unsuccessful campaigns I believe that the median is a better representation of the data. If you were to look at the two datasets, both successful and failed, on a histogram or a box and whisker plot it would show a peak low in the data set below one thousand and would trail far off to the right going over six-thousand backers. There is a massive amount of variance is this data set meaning the mean is being skewed higher than it should be. Whereas the median is the exact middle point of all points in the data with little to no skew because its not based on input data but the number of data inputs.

Variance in the Data

There is much more variance in the successful campaigns than compared to the unsuccessful campaigns. To me this does make sense, where some campaigns may be overambitious and not reach there goals the opposite is true of the successful ones where low requirements are met with relative ease and low backer counts. It also makes sense that if a campaign is unsuccessful then there is more than likely less backers compacting the data towards the lower end. The opposite would also seem true for successful campaigns where if a campaign is receiving lots of backers then the likelihood of it being a success rises inflating the maximum number of backers in the data set. These two factors are where its higher variance comes from in my opinion. The combination of low requirement projects becoming a success easily and high amounts of backers increasing success chances stretches the data. In comparison there is no minimum amount of backers to a failed campaign, but there is a maximum before it turns into a success.