Module 1 Challenge

Doug Short

Question 1 – Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

By looking at the sub-categories entrepreneurs and investors can understand the risk associated with various types of projects. For example, projects like Food Truck and Plays are either failed or canceled 52% of the time, which means these projects have a greater probability of failing than succeeding – which shows that people are not as interested in funding these projects (a Good piece of information for people looking to have their project funded). On the other hand, Nonfiction is failed or canceled only 33% of the time, making it a product with a higher probability of success. Similarly, Web is cancelled or failed only 27% of the time.

The data shows us what sub-categories are most likely to receive support from the crowdfunding community.

The data shows us that there is NO optimal time for creating a crowdfunding campaign. Some people would inherently look at the data to see if there was a best time to attempt crowdfunding for a project (i.e. at the beginning or end of the calendar year, or during summer months.)

The data shows us that failure rates on a monthly basis range between 31% (the low) and 40% (the high). Failure rates are somewhat consistent month over month.

We can also see that projects with an initial funding goal of greater than $50,000 are the least likely to succeed. This is very important data for any entrepreneur or investor to understand. People that are looking to launch o project or invest in crowdfunding projects need to use this data to assess their risk.

Question 2 – What are some limitations of the dataset?

The dataset is limited by the information available, as more data would be optimal. For example, there is nothing in the dataset that tells us about the experience of the leadership/founder of the company launching the projects for crowdfunding. Success and failure in business is certainly tied to leadership, and I think a column that identified leadership experience OR a “True” “False” column to tell us if the person launching each project had ever launched a project before.

I would also like to see a 1 person MAX pledge as a column. The average donation is too much of a generalization. I think it would be better to see the 1 person MAX pledge…that way an entrepreneur can see if certain categories and/or sub-categories have a few big “investors” that are really keen on that category or sub-category. Eliminating the 1 person Max would also clean up the average donation column, giving what I think is a more accurate viewpoint.

Question 3 – what are some other tables and/or graphs that we could create, and what additional value would they provide?

I think a Summary Statistics Table per Sub-Category would be very beneficial. It would show the entrepreneur and the crowdfunding investors a more realistic view of their expected outcome.

Another significant chart would be Percent Funded by Parent Category and/or Sub-Category. It would be great for entrepreneurs and crowdfunding investors to understand if there are specific areas where funding is far exceeding the original goals. This would allow investors to feel more comfortable about their personal crowdfunding investments if they see that they are investing in a sub-category that is going over its goal by 200% (or more).

Use your data to determine whether the mean or the median better summarizes the data.

I think the Median is the better summary of the data. Given that the data does not have a normal distribution, the Mean is not a good summary of the data. We have a high variance, which impacts how far away our data points are from the Mean, therefore we do not have a tight data set.

Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is less variability with the unsuccessful campaigns. Its variance and standard deviation were lower. I think this makes sense because unsuccessful campaigns have not reached their Target Goal, and therefore are going to have underperformed which means consistently lower numbers of backers. By contrast, successful campaigns have succeeded in raising their funding and may have raised much more than their target goal – therefore having many backers, making the numbers of backers in successful campaigns more varied. There is no CAP or maximum on the number of backers a successful campaign can have.