**Kickstarter Report**

UO Data Analytics Bootcamp

Cindy Jones

12/19/20

Refer to file file CSJ\_Kickstarter Analysis Final.xlsl

**Conclusions**

1. **Full Table**
2. **Category Outcome**

The chart reveals that three project categories contain by and large the largest number of successful projects, those being film & video, music, and theater. And while theater had more projects than any other category, it is music that has the greater probability of success. This can be seen on the chart in the large proportional representation of yellow in the bar.

1. **Sub-Category Outcome**

The Sub-Category chart may give us some insight into the particular type of project leading the way. With more than 1000 projects (within the scope of this data set), plays dwarf all other project sub-categories. And the majority of them are successful.

1. **Monthly Outcome**

Throughout the course of a year (Jan-Dec) the data suggest that while the numbers of failed and canceled projects remain steady, there is a noticeable decline in the number of successful projects. And while the linear trajectory is downward, it appears that this decline is occurring during the second half of the year. (If I could, it would be nice to present a trendline for the first six months of the year, and a second trendline for the last.)

1. **Project Size**

The chart shows that for most project goal sizes, successful projects outweigh unsuccessful projects. The dramatic difference is in projects with goals between 25,000 and 45,000, where the inverse is true.

1. **Backer Outcomes**

While a significant amount of variance may be clouding and distorting the precision of the data, it is clear that the successful projects had more backing. (I think we knew that before the analysis.) Due to high variance and the presence of outliers (as can be seen on the chart), it may be better to use the median as a focus for summarizing the data.

**Limitations**

1. **Full Table**

It should be noted that the currencies referred to in this data set include a mix of USD and GBP, and if any analysis is to be done with regards to project values, the currencies should be aligned. Column U includes the dollar equivalents of average donations when 1GBP is equal to 1.35USD.

1. **Category Outcome**

The proportional balance of outcomes within categories is consistent for all but technology, which has an exceptional number of canceled projects. This could be a true characterization, or it could be the result of mis-categorization or scientific error. A closer look at that subset of data is needed.

1. **Sub-Category Outcome**

In using a category filter to show just the technology sub-group, it can be seen that most failed and canceled projects come from wearable and web projects. What we don’t have is clear and delineated definitions of any of these categories. For example, if one is funding a robot, is that a gadget or hardware? If wearables are web-linked, do they fit in to the wearable or web category?

1. **Monthly Outcome**

The data occurs over a series of years, and the “seasonal” effect should be confirmed by looking at the number of projects over a longer period of time. It’s possible that the trend is consistent in different years, but it’s also possible that the downward trend is related to a more specific time.

1. **Project Size**

The total number of projects within the 25,000 to 45,000 goal range is far too small to be usefully analyzed. The small sample size may be leading to misrepresentation.

1. **Backer Outcomes**

The large variance shows that this data is not ready to be useful. It’s possible that cleaning the data could improve it’s resolution, but that may not be enough to make it reliable.

**Ideas for Further Analysis**

1. **Full Table**

See Column U: Dollar equivalent added

1. **Category Outcome**

Further study on why there are so many canceled technology projects. What do canceled projects have in common and how are the successful projects unique?

1. **Sub-Category Outcome**

Design a multidimensional graphic of the relative sizes of categories and the relationships between categories. Like categories would be grouped closer together. Should closely related categories (like types of music) be combined? Should “dramas” and “plays” be combined? I mean, who defined these as equitable categories? And “faith?” What is that?

1. **Monthly Outcome**

I’d like to see a graph that shows the trend over continuous time. Then, if I see seasonal trends I’ll analyze them as cycles, and display them in a way that shows they are cycles.

1. **Project Size**

It would make more sense if there weren’t a relative spike in 40,000 to 45,000 range. If the project number went down consistently, I could explain that projects decrease as they get more expensive. But the way it is, that’s an anomaly that needs to be explained.

1. **Backer Outcomes**

More backers support successful projects, and successful projects have more backers. I don’t think we learned anything from this sheet, and I’d just edit it out if it were a real report.