**Excel Homework**

Tim St. Onge

11/28/2018

1. **What are three conclusions we can make about Kickstarter campaigns given the provided data?**

One conclusion that can be made about Kickstarter campaigns is that “Theater” is the largest category of campaigns and “Plays” (a subset of “Theater”) is the largest sub-category of campaigns. There were 1,393 total Theater campaigns, of which a staggering 1,066 were Plays. This type of sub-category campaign was far and away the most common, with “Rock” music campaigns being the second largest campaign sub-type at 260 campaigns. This potentially points to a prominent use of Kickstarter for fundraising in the theatrical arts more broadly.

A second conclusion is that late spring may be the best time to start a Kickstarter campaign, with December being the worst time. This conclusion is informed by the “Kickstarter Campaign Outcomes by Starting Month” graph, which shows that the number of successful campaigns by starting month peak in May and are still relatively high in April and June. The number of successful campaigns largely decrease as the year goes on, culminating in December, in which more campaigns that started in this month failed than succeeded. It remains difficult to pin down the seasonality of this data or what specific annual trends might be in play, but this information may still be valuable to a campaigner looking for the best time to open funding.

A third conclusion is that while Theater and Plays make up a very large number of campaigns, a far as success rate goes, campaigners involved in “Music” themed campaigns may be more likely to find success on Kickstarter. As a percentage, about %77 of Music campaigns were successful, compared to about 60% of Theater campaigns. These percentage calculations are from the Pivot Table (and hinted at in the graph) in the “Kickstarter Campaign Outcomes by Category” section. Each category had no more than 61 Live and Cancelled campaigns. Within the “Music” category, “Rock” music may be the best bet, with the greatest number of campaigns within the Music category as well as a 100% success rate. Outside of these categories, “Technology” was the worst performing category, with only about 35% of its 600 campaigns being successful (35% failed, 30% cancelled). This type of information could suggest to a campaigner their chances of fundraising success in Kickstarter, based on the nature of their campaign. Theater campaigners may be encouraged by the large number of similarly-minded campaigners, although competition for funding may be tighter. Music campaigns, and especially Rock campaigns, may be encouraged by the high rate of success. Campaigners dealing with Technology campaigns may need to temper their expectations about fundraising success on the Kickstarter platform.

1. **What are some of the limitations of this dataset?**

There are at least several distinct limitations of this dataset which can inhibit our full understanding of Kickstarter campaign trends and insights. One potentially valuable piece of information that is missing is any ideas of initial project funding. While this data may not naturally be in Kickstarter’s database, we don’t know the impact of a campaigner’s starting funding or general wealth. Are well-financed projects (ie. Financed by other sources) asking for less money on Kickstarter, leading to more success? Are projects that start from scratch more likely to fail, and are these campaigners setting too high a funding goal?

We are also limited by only being able to calculate the average donation of a campaign, which still does not give us the best understanding of how each campaign is funded. Are there Kickstarter “superdonors” supporting many projects that we don’t know about, and that are hidden by the simple, skewed average? If we had the specific donor information (e.g. Donor X gave $5, Donor Y gave $50, etc.), we could calculate the standard deviation of this data and see with greater granularity how the campaigns are funded.

1. **What are some other possible tables/graphs that we could create?**

There are many other types of graphs and tables that could be created with this dataset, in addition to the ones created in this homework, that can provide insights into Kickstarter campaigns. Updated data tables could be created with a potentially useful, new calculated field: “Time Duration of Campaign”. The analysis has thus far not taken into account how long campaigns were open, which might be an interesting variable to consider. Are long-open campaigns more likely to be successful, or is there a more complicated relationship between campaign duration and success? Taking the difference between the “Date Created Conversion” and the “Date Ended Conversion” date fields would yield this calculated field. A scatterplot could be introduced to the analysis to visualize the “Time Duration of Campaign” variable against other variables such as goal or average donation. Points in the scatterplot could be colored by outcome or category to potentially yield more insights.

There are many other types of graph configurations that could be created. A bar graph showing average donation by category and sub-category may be illuminating, perhaps pointing to the existence of “high spending” or “low spending” donors taking interest in certain types of campaigns. While the homework included a line graph of Kickstarter Campaign Outcomes by Starter Month, it may be interesting to see a similar line graph, but charting the full history of Kickstarter campaigns (by creation date), going back to 2009. Such a graph might be able to reveal historical trends and tell us more about the Kickstarter platform overall: Are Kickstarter campaigns more or less lucrative (or successful) now, as opposed to in its early days? Lastly, a basic TreeMap chart may be a helpful way of visualizing the number of campaigns in the dataset organized by category and sub-category.