Leon Tan August 18th, 2018

**Kickstarter Program Report**

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

The three most successful crowdfunding based on parent category are fil and videos, music and theatre.

Audience would prefer to attend to a performance with a plot. Plays such as in theaters have a much successful effect in crowdfunding process. Other sub-categories in theatres are known to have a high failed rate (i.e. musical, spaces) since the plot are harder to comprehend. That is why in film and video category, the success rate is high as well and not to mentioned, rock music that usually have meaningful lyrics to understand that can be related to the audience.

Audience prefer to attend performance that are relatable in real life. It seems that in music, the rate of successful crowdfunding is high. Indie rock and rock music took up the highest success rate as the audience would feel connected to the lyrics in the songs. Same thing goes for plays in theatre. It is a live play that is telling a story to the audience and that is what makes it a high successful rate. In film and videos, documentary has a high success rate because it is documented with proven data to tell the audience a story.

All categories have a low success rate at year end, December. This same condition applies to three parent categories mentioned above. It seems like the possible explanation is that everyone is busy with the end year holidays and new year’s is just around the corner, not allowing the potential donors to prioritize the campaigns.

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

Some limitations would be not having enough variables to compare and tell a story. Without a doubt, there are timelines and parent categories to filter through the datasets, but the correlation between the states of outcomes are might be affected from other variables. For example, did Kickstarter hired popular celebrities to play in the theatres that is why their success rate is high? Or why do the graphs have a peak during May? Is it because of summer or other reasons?

There are other limitations such as not knowing have a precise goal percentage. The percentage that we are asked to set is 200% but there are goals that have been achieved in more than 20000%, which really confuses the user that analyzed this data. Also, the currency is different and the only thing that have in common are their percentage funded. But that is still not enough to pull out any relatable variables.

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

We could create a scatter plot as well, to see what type of relationships for the variables have with each other. Or we could also create a graph bar of parent category based on the years. Because there are different generations throughout the years and some people in this generation might not have interest in political or fundraising campaigns.

Chart 1: Filtered by Country

Chart 2: Filtered by Country and Parent Category

Chart 3: Filtered by Parent Category and Years