Mike Suomi

Homework #1

May 8, 2018

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

1. Plays dominate the total usage for projects on Kickstarter – accounting for almost a quarter of all campaigns. This would indicate one of two possible next steps to explore: 1) increase marketing towards those with an interest in theater/plays since they are likely to use kickstarter or 2) the theater/plays market is saturated (they all know about the use of kickstarter) so diverge that money to other sector campaigns to improve those categories.
2. Kickstarter has very low international volume outside of the US (74% of total campaigns) and Great Britain (15% of total campaigns), so there is a lot of room for growth in international markets.
3. In the music sector, Rock music was clearly most successful (and popular) with 100% of the 260 campaigns getting successfully funded. I would not recommend Jazz and Faith music to try startup campaigns on Kickstarter as their campaigns have uniformly failed to date.

What are some of the limitations of this dataset?

I find these open-ended datasets to be very challenging to interpret without a lens of who is asking the question or having more specifically defined questions. Should I care about success rate and what makes a successful Kickstarter to help give advice/customer support? Or should I care more about the amount of money that is passing through our site and trying to maximize that? These are just two small possible questions of many that could be posed, and each question requires very different data and focuses that would benefit from more outside data. Some additional data that I think could add more stories would include: advertising dollars/clicks that are working for people that are hosting a Kickstarter, advertising/clicks that lead to donors funding a project, quality rating of a product’s summary/visuals (that are more subjective), etc. Overall, I didn’t find my generic conclusions very novel or useful without further context to drill down or questions to be answered.

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

1. I created a 100% stacked column based on the state of a project (success, failure, canceled) to interpret what the trend over time was of the amount of Kickstarters and whether they are more likely to get fulfilled or less likely. The results indicate that as time has gone on the failure rate of Kickstarter campaigns has increased (there is large volatility of this based on category as some categories are getting better, like games and food, and some are getting worse, like publishing and film & videos). This indicates that perhaps Kickstarter should explore these causes further – do we need to give better guidelines for successful campaigns or vet certain ones? Or does mere quantity and our algorithms of which get pushed to the top as recommended lead to success/failures?
2. I also attempted a Scatterplot of % Funded vs Goal Amount (USD), but I did not get any good correlation plots as there seems to be a lot of interesting outliers. I imagine that much of this is due to the fact that donors are often guaranteed something in return (like a product, swag, etc.), which often leads to huge over-funded projects based on interest. This makes the correlation seem a lot more random in the data and hard to pin down if there is really a goal amount that is too large (the bonus graph we created did show a pretty sharp drop-off of success over $45,000, but didn’t indicate correlations either).