Unit 1 Homework: Kickstart My Chart

# Intro

Over $2 billion has been raised using the massively successful crowdfunding service, Kickstarter, but not every project has found success. Of the more than 300,000 projects launched on Kickstarter, only a third have made it through the funding process with a positive outcome.

Getting funded on Kickstarter requires meeting or exceeding the project’s initial goal, so many organizations spend months looking through past projects in an attempt to discover some trick for finding success. For this week’s homework, you will organize and analyze a database of 4,000 past projects in order to uncover any hidden trends.

# Conclusions

1. The most successful categories are Film & Video, Music, and Theater, where nearly 2x as many campaigns have been successful as opposed to failed. Within those categories, Plays were clearly the most successful, with 694 successful campaigns and 353 failed. Next most successful would be rock music and then documentaries, where all 260 and 180 campaigns, respectively, were successful.

Figure Successful subcategories of Film&Video, Music, and Theater and their outcomes

1. Successful campaigns most often had goals between $5000 and $9999, or under $1000. When goals for the campaign are set at $20,000, they become more likely to fail than succeed, with a significant increase in failures for goals $45,000 or more. The one exception for this seems to be the $35,000 to $45,000 range.
2. Time of year may play a role in success. When we look at a line graph showing trends of canceled, failed and successful campaigns over time, there are three points where there are almost twice as many successful campaigns as failed ones – February, May, and to a lesser extent November. It might be worth doing hypothesis testing (chi squared) at a future date to validate this observation and see if it is a statistically significant difference.

Figure Number of campaigns by outcome over time 2009 - 2017

One factor to consider, however, when we look at the relationship of time of year and success vs. failure for campaigns is that the ratio of successful to failed campaigns has changed over Kickstarter’s life span to date. The data shows that in the beginning (2009 – 2013) successful campaigns outnumbered failed and cancelled substantially. However, starting in 2014, there started to be a 1:1 trend of successes for failures. It is hard to know if this relates to changes in the platform and its popularity or ease of use? Or to Kickstarter expanding into countries around the world? Or is it simply regressing to the mean? Before drawing too many conclusions from the data, it is worth investigating more current data to see if the trend holds, or if it changes.

Figure Outcomes of campaigns across life of Kickstarter

Another key piece to consider with this data is that cancelations start occurring in 2014, where concerns around Fraud become a higher issue. Kickstarter campaigns are cancelled when they are considered fraudulent in some way. Overall, the technology category has had the highest number of cancellations, in a 1: 1.2 : 1.2 ratio of cancellations : failed : successful campaigns.

*In summary, if you want to have a successful Kickstarter campaign, choose a play to fund, with a goal of between $5000 and $9000, and go live with the campaign in February.*

# Limitations

* Our data is old by today’s standards. The most recent data is from Q1 of 2017, and so is now 2.5 years old. As we saw previously, the trends have changed year over year for Kickstarter, so the lack of more current data limits our ability to generalize to current day. This in turn impacts our ability to draw conclusions and make predictions.
* Our data is far from complete. Only 4,000 data points were examined from over 300,000. To complicate matters, we do not know how these data points were chosen in the first place (for example, if they were truly random, etc). Additionally, Categories and subcategories do not align with the current roster on the website. For example, in the data there are only 3 subcategories for games – mobile, tabletop and video. On the current site, there appear to be more and different categories – RPGs, Console games, PC Games, Strategy games, and Visual Novels. Similar changes exist for other categories as well. This means that the data and the project classification (as well as the conclusions drawn from it) may not be generalizable to Kickstarter’s current system. And lastly, we are missing data from our data points such as country of origin for the campaigns, which might be a necessary component of gaining further understanding of trends.
* As mentioned before, we don’t know what platform changes were made over time to Kickstarter. Any policy changes or user experience changes to increase ease of use or creating campaigns could be a confounding variable for the data examined. Platform changes are common for technology companies, and could be controlled for if we knew the changes and time points. However, without that data, we may be missing factors in our analysis and conclusions.

# Other Possible Tables/Graphs

* **The impact of spotlight and staff pick features**. A bar graph showing the number of campaigns that succeeded or failed categorized by whether or not they were given the spotlight would allow us to see if that had a positive impact on generating support. Likewise, I’d also like to see one of campaigns that succeeded or failed categorized by whether or not they were a staff pick. My hypothesis is that those campaigns who were given these opportunities were more successful than those not given the spotlight or designated a staff pick.
* **How number of backers impacts success.** A line graph of average number of backers by category, with different lines based on success, failure, canceled would help us to dig further and understand the impact of number of backers on the campaign. Looking at a cross section of this data with who was given the spotlight vs not, and who was a staff pick vs not, would also give us an idea of whether additional support was actually generated by these features. We could even go so far as to do hypothesis testing with mean number of backers for spotlight vs. none, and mean number of backers for staff pick vs. not in order to see if it approaches statistical significance. My hypothesis is that those given the spotlight or a staff pick had higher numbers of backers and more success than those without.
* **How length of campaign impacts success or failure.** Using a stacked bar chart, with length of campaign divided into time bands (similar to what we did with goals), I’d want to see the proportion of successful campaigns to failed campaigns. My hypothesis is that shorter campaigns were more successful due to the pressure and sense of urgency, but only so long as they had enough time and resources to advertise and market the campaign.