**Unit 1 | Assignment - KickStart My Chart**

**Dan Boulden**

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

2.What are some of the limitations of this dataset? Pages 1-2

3.What are some other possible tables/graphs that we could create? Pages 3-4

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

1.1We can conclude that, from a percent successful perspective, music is the most successful category and food is the least successful category, and Journalism seems to have an issue as it is all canceled.

1.2There are (according to our data) far more campaigns about plays than any other sub-category, but the success rate for plays is only 66% compared to most music sub category’s, as most music sub categories have a success rate of 100%

1.3 The best months to start a campaign and the success percentage are May (61%), February (60%), and April (60%). The worst months to start a campaign are December (44%), September (49%), and Jan, Jul and Aug all at 50%. The most cancellations happen in July and November both at 11% and the least cancellations happen in October with only 6%.

1.4 The bonus question shows us that the best success to failed ratios are at $4,999 and under campaigns (<1,000 = 71%, and 1k to 4,999 = 66%) it then becomes roughly 50/50 until it drops off sharply at 45,000 (45k to 5k = 29%, and 5k < = 19% success)

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

With the current data, we are unable to do the following

2.1 Figure out what percentage of people looked at a campaign and ether decided to fund or not to fund. We have the “backers\_count” field, but if we had some additional data, like number of unique users who viewed the campaign page (and perhaps how long they viewed the page on average) we could look into stats like how many people do you need to get to view a page to get X$’s pledged. In other words, what is the average number of dollars a visitor spends, not just visitors who pledge.

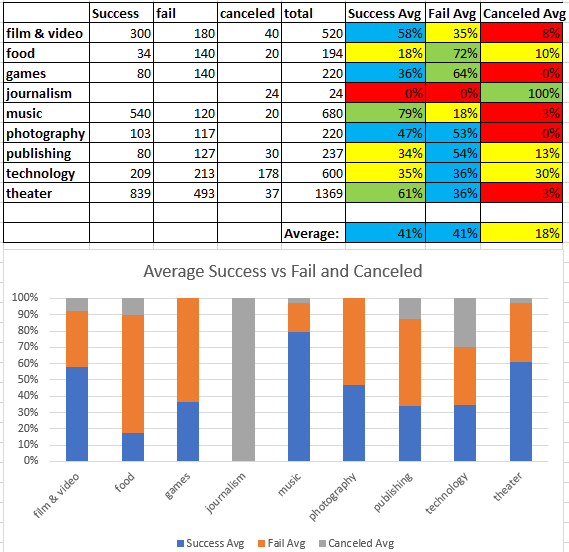
2.2 If we had more data behind each campaign, we could get a better idea of when to advertise a campaign on something like Facebook or Instagram. So if we had data that showed the date, time, amount per backer pledge, and where the backer came from (IE Facebook, etc.) we could figure out for each type of campaign what time of day and day of the week and month has the highest percent of things like pledge to no pledge, or if they pledge, what is the best day/time to get the largest pledge broken out by social media.

2.3 We could take 2.2 a step further and look at (if we had additional data) what is the best amount of social media advertising to do for a demographic. In other words, we could look at if bombarding someone with adds and information about the campaign works or turn them off, what is the right amount of advertising based on their demographic.

NOTE: for 2.2 and 2.3 we will define advertising as both paid and unpaid ads as well as posting information in written, photo and or videos. In other words, any direct from the campaign owners’ postings about the Kickstarter campaign.

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

3.1 For some reason, within our sample, all Kickstarter campaigns for journalism get canceled. No photography or games campaigns got canceled in the data. So one table, we could make, would be the averages for all the categories so that we could more clearly see the average vs the total for that category. We can also build a bar chart that visually shows us the percent of success vs the others. Note that Live was removed as it is inconclusive data for this table/chart. The value of this table and chart over the one from the project is that it shows the data for each category in relation to all categories. It is quick to compare all categories together.



3.2 The same thing could also be done for the Sub-Category giving us more granular information about overall averages.

3.3 What is the percent under goal that a campaign is at a certain number of days after the start date that the project is canceled. In other words, if a project is canceled, what is the average number of days out from the start date was it canceled and what was the average amount of $’s under the goal that was pledged at the time the project was canceled? This could be very useful in helping to pick an achievable budget.

3.4 For successful complains we could compare the amount of days on average it took to reach the goal in relation to the category and sub category and the total raised. In other words, for each category and also for each sub category how long did it take to reach the goal on average. Knowing this could help campaigns make better stop/go decisions, as well as help Kickstarter guide users on how to setup and manage their campaigns.

3.5 We could look more into what types of campaigns succeeded in the 45,000 and up categories and try to determine why they succeeded over the others. Basically doing the same thing as the category and sub-category pivot tables, but adding the $’s element to the pivot tables in a way that would allow for grouping breakdowns based on the dollar groupings from the bonus question.