Ashley Drayton

Q1. Given the provided data, what are three conclusions we can draw about Kickstater campaigns?

Based off the provided data I believe the following conclusions can be made in regards to, (a) the success of kick-starters in based off the category they have applied for, (b) the success of a kick-starter in relation to which country they live and finally (c), the date created impacting the level of success.

1. The given data has allowed us to see that the success and failure of any kick-starter is swayed by which category it belongs to. This statement is evident through the vast difference in successful vs failed kick-starters when divided into categories.

Concluded kick-starters in the category of Music, with a total of 680 (with 20 live), have been able to achieve a succession rate of 79.41%, with 540 being successful, whilst only having 17.64% (120) failing. The same could be said for Theatre with 61.28% (839) being successful whilst have a failure rate of 36.01% (493).

Alternatively, we can see the opposing effect that the category has when looking at both Publishing and Food. The data shows that the percentage of failing kick-starters is rather high, with only 53.58% (127) in Journalism and 72.16% (140) in Food, with only 33.76% (80) and 17.5% (34) respectively, being successful.

Although this is not the absolute determining factor to the success and failure of these particular categories, we can still assume that for any currently live or future kick starters, that Music and Theatre will more than likely have a successful outcome, with the opposite for Journalism and Food, in relation to their goal for total amount funded.

1. Sifting through the data has also made clear of a distinctively high succession average being connected to the origin of the kick-starters (shown in “Conclusion b” sheet). This high succession rate can perhaps be attributed to the total amount of kick-starters being applied. This is shown by the US and Great Britain having the first and second total kick-starters with 3038 and 604 as well as a 54.94% (1651) and 61.41% (366) success average for completed cases.

I believe what the data shows has the potential to accurate amongst all countries given a similarly high amount of total cases. The high amount of cases shows an awareness amongst the population of the service, which could equally be said about the countries with a low amount of cases as this may be because of a lack of awareness for the service. Where there are an increasing number of projects, this could potentially improve word of mouth as well as other marketing types to promote the service to then lead to a wider audience being made aware and thus more donations.

1. From the data given, it can be found that 76.59% of all projects were created between 2014 and 2016, which is can be responsible for 69.38% of all successful projects.

A possible reason for this extreme number of cases during only 3 years could be due to an improved understanding of how applicants promote/ devise an informative and more well-written description that will improve their chance of people donating.

Q2 What are some of the limitations of this dataset?

As the data set provides an over whelming amount of projects it is unrealistic to take time to sort through all of the proposed projects to looks for blank spots of data as well as verifying if all of the projects are legitimate and are not inherent outliers that will sway the data. This is illustrated with cell “F124”, where a project has set a goal of $100,000,000, stating “My ambition knows no bounds”.

Another limitation for this data set is that although only a small amount of kick-starters were canceled (349), having the data available for how they were tracking in terms of reaching their goal could have made some differences to the overall success and fail rates of particular categories. This is evident when looking at the Journalism and Food categories, with 100% and 10% respectively being canceled.

Q3 What are some other possible tables and/or graphs that we could create?

I believe that a scatterplot could help improve the scope of the data due to its ability to provide a clearer linear relationship between certain factors. This alternative view could then bring light to more accurate or unique similarities that are not shown in the current graphs and tables.