BCS Homework

03-20-2021

Matthew Granson

**Question 1: Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**

1. When filtering the data to focus on category around the world, it becomes apparent that “entertainment” themed Kickstarters (film, music, and theater) are the most popular. Each of these categories contain the highest amount of successful projects compared to all other categories in the data set. Of these three, theater is the most popular project type with 38% (n=839) successful campaigns of the 2185 recorded. The data suggests that theater productions are the most common type of campaign projects on Kickstarter.

2. Looking at the history of successful campaigns stretching from 2009 to 2017, the data suggests that projects created/started in early summer (quarter 2) have the highest success rate. Regarding theater productions (as noted in question 1 above) the data shows that quarters 2 and 3 (specifically April to August) produced the highest number of successful campaigns. This would seem to suggest that future Kickstarter patrons should look into starting their campaigns during the summer months if they are planning a theater production or play.

3. On the contrary, looking at failed or canceled campaigns, technology themed projects exhibit the highest amount of cancellations. Of the 2185 successful campaigns only 10% were technology projects (n=209). In fact, every category other than film, music, and theater displayed success rates at 10% or lower. Journalism projects were shown to be the least common type of Kickstarter campaign all of which were canceled (n=24). Based on this data set, media content such as music, plays, and film was the most popular content hosted on Kickstarter from 2009 to 2017 and was most likely to succeed in being funded. This trend makes sense as those dates encompass the explosion of social media and internet hosted, (indie) content creation.

**Question 2: What are some limitations of this dataset?**

First is the source of the data; where is this data coming from? How was it collected? There is always potential for bias/error in data curation. Second is the extremely varied sample sizes. It is hard to make meaningful comparisons with categories like “journalism” which has a sample size of only 24. Data coming from the US makes up such a significant portion of the data set it will be hard to draw connections between the geographical origins of campaigns to their funding success rates or user donation behavior. This is primarily due to other crowd sourcing alternatives such as GoFundMe which may be more popular in foreign countries.

**Question 3: What are some other possible tables and/or graphs that we could create?**

While it would be informative to show scatter plots comparing the average donations of users across different countries or project successes between continents, I would personally like to see some data visualization depicting occurrences of fraud or “scam campaigns”. Comparing the frequency of these occurrences with other sites like GoFundMe or Indiegogo would allow us to see the reliability of each site’s respective method of crowd funding.