Meng Zhao

9/11/18

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

The dataset was composed of 4114 projects listed on Kickstarter created between 5/17/09 and 3/15/17 from 21 countries, where projects from US made up 74% of the data. Between 5/17/09 and 3/15/18, and across the 21 countries, the most popular projects created on Kickstarter were theater projects (n=1393), which was almost twice the number of the second most popular projects, music projects (n=700). However, music projects had the highest rate of being successfully funded (77%), whereas theater projects had the second highest successful rate (60%). All journalism projects were canceled which resulted 0% successful rate. Food projects also had a low successful rate that only 17% projects were successfully funded in the past. Within theater projects, the most popular subcategory was play that 1066 plays were listed on Kickstarter, which was 77% of the total number of theater projects. Within the second most listed category, music, the most popular subcategory was rock (n=260). In terms of rate of being successfully funded, 11 subcategories that had a 100% successful rate. Rock and documentary were the two subcategories that had the greatest number of successful cases (Rock n=260; Documentary n=180). In terms of time span, 2015 was the booming year of Kickstarter that 1225 projects were created. July was when projects created the most on Kickstarter, that 387 projects were created in history. However, May was the month having the greatest number of projects being successfully funded. 234 projects created in May were successfully funded. Projects created in December were least likely being successfully funded that only 111 projects were successfully funded in the past. Looking at how funding goal affected the success of funding, the rate of being successfully funded decreased as the goal of funding increased. However, the pattern changed when it came to the goal of 35000-44999 that success rate bumped back to 47%-49% but dropped significantly after 45000.

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

One significant limitation of the dataset is that the data is mostly composed of US data. I’m not sure how the data was sampled that whether having US as the major component in the dataset reflected the real condition of Kickstarter. Also, 2015 was the year where a lot of data coming from. I’m not sure whether it was due to sampling error or reflected the real condition of Kickstarter. Another limitation of the data is that currency from different countries was not converted to the same currencies that drawing any money related conclusion would be inaccurate. The amount of data may also be one of the limitations.

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

There’s a variable in the dataset called “staff\_pick”. How I understand this variable is that some projects on Kickstarter are listed in a special section called “staff pick” on the website, and may receive more attention. The question I have for this variable is if staff pick may affect the possibility of being successfully funded. To answer this question, we can do a pivot table like below, doing column %, to get the successful rate within being staff picked versus not staff picked:

|  |  |  |  |
| --- | --- | --- | --- |
| **Count of state** | **Column Labels** |  |  |
| **Row Labels** | **TRUE** | **FALSE** | **Grand Total** |
| failed | 11.96% | 46.29% | 41.18% |
| successful | 88.04% | 53.71% | 58.82% |
| **Grand Total** | **100.00%** | **100.00%** | **100.00%** |

From the table we can see that the successful rate of being staff picked (88%) is much higher than not being staff picked (53%). We can follow up with T-test to see if this difference is statistically significant.