**Data Science & Visualization Boot Camp**

Excel Homework #1

Amir Vazquez

11/13/2020

1. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**
2. The 3 most successful categories are Music, Technology and Theater
3. Music KS Campaigns are 79% most likely to succeed in funding and project completion
4. Theater KS Campaigns are 61% more likely to get funding and complete their project
5. Technology KS campaigns are 35% more likely to get finding and complete their projects.

Hence theater and technology are 36% more likely to fail in their campaigns. And Music are more likely to cancel their campaigns by 30%.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Years | (All) |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **Count of outcome** | **Column Labels** |  |  |  |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **successful** | **Grand Total** |  | **Cancelled** | **Failed** | **Successful** |
| music | 20 | 120 | 540 | 680 |  | 3% | 18% | 79% |
| technology | 178 | 213 | 209 | 600 |  | 30% | 36% | 35% |
| theater | 37 | 493 | 839 | 1369 |  | 3% | 36% | 61% |
| **Grand Total** | **235** | **826** | **1588** | **2649** |  |  |  |  |

We can note that the countries with the most KS campaigns are 1st world countries like USA and UK.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Years | (All) |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **Count of outcome** | **Column Labels** |  |  |  |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **successful** | **Grand Total** |  | **Cancelled** | **Failed** | **Successful** |
| **GB** | **19** | **115** | **305** | **439** |  | **4%** | **26%** | **69%** |
| music | 1 | 7 | 25 | 33 |  | 3% | 21% | 76% |
| technology | 12 | 19 | 22 | 53 |  | 23% | 36% | 42% |
| theater | 6 | 89 | 258 | 353 |  | 2% | 25% | 73% |
| **US** | **166** | **597** | **1173** | **1936** |  | **9%** | **31%** | **61%** |
| music | 19 | 110 | 490 | 619 |  | 3% | 18% | 79% |
| technology | 121 | 138 | 158 | 417 |  | 29% | 33% | 38% |
| theater | 26 | 349 | 525 | 900 |  | 3% | 39% | 58% |
| **Grand Total** | **185** | **712** | **1478** | **2375** |  | **8%** | **30%** | **62%** |

When focusing on UK/GB and USA as the primary countries we notice that GB has a greater success rate on the mentioned categories vs the USA.

**B)** The distribution of data for successful and failed campaigns has a right positively skewed. Meaning that the mean was lager and the mode in all of the cases

For both scenarios the standard deviation was extremely off from the mean causing the data to show a dispersed distribution of the data as seen in the probability distribution charts

The data had some outliers that created some noise. But even when removing these outliers from the data, I found that it didn’t helped to find a balance in the data.

**C)** The data showed that the largest percentage of successful projects come from campaigns that had goals less than 10,000. And that campaigns that have a large expectation goal tend to fail as their pledgers cannot meet the original goal.

Although the cancelled projects show as only 9% of the whole population of data. It is uncertain why the project was cancelled. Lack of funds or other reason could be the reason why the project could have cancelled.

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

* Need more demographics on the data. Such as Sex of the pledgers, education degree, etc.
* Is the result of this KS campaign started anything bigger? No data to back this up
* Are these KS campaigns from already made corporations or are they all form non corporations/ small business?
* There are no bonus tiers in projects when pledging bigger amounts
* Success factor of same type of projects across different countries.

1. **What are some other possible tables and/or graphs that we could create?**

* Distribution of data for subcategories by countries.
* Heat maps by:
  + Category
  + Country
    - Successful, Failed, Cancelled, Live
* See if there is any impact when the Spotlight or Staff pick are set to a project
* Comparison of Smaller campaigns vs large campaigns