**Assignment 1 - KickStart My Chart**

Due: February 16, 2019

Submitted by: **Polina Zasheva**

TU-TH-SA Session

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

There are many key insights we can derive about the Kickstarter campaigns from the data provided. **First,** the majority of campaigns were successful with the largest success attributed to ‘Plays’ within the ‘Theater” category. **Second,** the month of May seems to generate the greatest number of successful campaigns, followed by June, or in other words, late spring through summer is when most campaigns see success. These are also months with the largest number of campaigns overall. The months that generate the least successful campaigns are December and September and those two months have the lowest number of campaigns overall. The months with most failed campaigns are October and January. **Third,** the category with the lowest engagement with KickStart campaigns in Journalism, followed by Food. Within that, Journalism actually had no successful campaigns between 2009 and 2015.

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

In general, the answer would depend on what an analyst is trying to understand from the data. In the specific case, we are asked to understand past trends of successful projects in order to determine trends that may drive that outcome in the present. Based on that, here are some potential limitations:

* The data cuts of in 2015 so we can’t use this information to understand more recent performance of Kickstart campaigns. With rapid advancement in technology that lead to ever evolving changes in market trends and consumer preferences, four-year-old data may not be as meaningful as data from prior or 2 prior years.
* The goal and pledge amounts were in different currencies, depending upon the ‘country’ within which the campaigns were held. Though we can compare how monetarily large a campaign is within a county, it may not be a good ‘apples to apples’ comparison across the entire data-set because of the currency variances. To understand trends of success, it may be more useful to review the data one country at a time.
* In calculating the average donation per backer, we don’t know if the backer is an individual or an entity. It may be good to have that delineation in order to understand if individuals within a group are more likely to donate more on average compared to individual donors.
* We are missing some key attributes that tell us what drives success, like number of participants running the initiative, the outreach method for generating donations, the demographic of those participating, was there a geographical part of a country that saw more success (i.e. are urban Kickstarted campaigns potentially more successful than those that we run in rural areas)

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

* Successful campaigns by country

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Row Labels | successful | live | failed | canceled | Grand Total |
| US | 1651 | 33 | 1097 | 257 | 3038 |
| GB | 366 | 8 | 205 | 25 | 604 |
| CA | 64 | 1 | 64 | 17 | 146 |
| DE | 23 |  | 27 | 3 | 53 |
| AU | 19 |  | 41 | 14 | 74 |
| ES | 11 |  | 9 | 1 | 21 |
| FR | 10 | 2 | 10 | 5 | 27 |
| IE | 8 | 1 | 4 | 2 | 15 |
| IT | 7 |  | 19 | 3 | 29 |
| SE | 7 |  | 9 | 5 | 21 |
| DK | 4 |  | 6 | 4 | 14 |
| NZ | 3 |  | 5 | 4 | 12 |
| AT | 3 | 1 | 2 |  | 6 |
| NL | 2 | 1 | 14 | 4 | 21 |
| LU | 2 |  |  |  | 2 |
| NO | 2 |  | 5 |  | 7 |
| BE | 1 |  | 1 |  | 2 |
| SG | 1 |  |  |  | 1 |
| CH | 1 |  | 2 | 3 | 6 |
| MX |  | 3 | 8 | 1 | 12 |
| HK |  |  | 2 | 1 | 3 |
| **Grand Total** | **2185** | **50** | **1530** | **349** | **4114** |

* Successful campaigns by year
* Overall breakdown of campaigns by state:
* Top 10 most successful sub-categories by count of successful campaigns

|  |  |
| --- | --- |
| state | successful |
|  |  |
| Row Labels | Count of name |
| plays | 694 |
| rock | 260 |
| documentary | 180 |
| hardware | 140 |
| indie rock | 140 |
| photobooks | 103 |
| spaces | 85 |
| tabletop games | 80 |
| television | 60 |
| musical | 60 |
| nonfiction | 60 |
| shorts | 60 |
| **Grand Total** | **1922** |