**Rosanna Reza-Giles**

**Module 1 Challenge**

**Written Assignment:**

**1.Given the provided data, what are three conclusions that we can draw about crowdfunding Campaigns?**

Crowdfunding campaigns are an ideal way for individual and companies to deliver timely funding to conventional and sometimes unconventional initiatives, judging by the 9 parent categories in the “Parent Category Outcomes” table and chart. The span of categories also totaled 22 in the data and each and every category had success to some degree, the highest being plays and the lowest success being world music. The successes were no bound by country of, as the results from includes United States, Canada, Denmark, Sweden, United Kingdom, and countries in Europe.

Crowdfunding plays is a great way to get your project off the ground. When compared to all the other 22 subcategories it was head and shoulders above the rest of the 348 attempts at plays more than half of the attempts landed in the successful category. Every subcategory had a level of success but plays our performed them by double digits.

Crowdfunding initiatives are short lived, the data does not lend insight if this is by design or a symptom of the process but of all the 1000 campaigns 566 got a “successful” result but only 14 showed the “live” condition as per the data in the chart. I believe more insight needs to be concentrated on this population of campaigns to answer some important questions around sustainability and longevity for the campaigns.

**2.What are some limitations of this datasheet?**

Limitation:

Description of campaign need better representation

Evidence:

Most of the entries are non sensical blurbs this information was not useful in any analytical sense, perhaps creating a description template that lends more insight into the purpose or values of the campaign. Some descriptions like “Pre-emptive tertiary standardization”() a blurb from one of the least successful campaigns or “Monitored incremental info-mediaries” () description form one of the most successful campaigns leaves the donor with questions that may hinder donation.

Limitation:

There is too much data spanning multiple countries and over many years, resulting in a lot of data cleaning and possible irrelevance. We also don’t know how the data was gathered and how it relates to the change in the way crowdfunding has evolved over time.

**What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

1.Do a chart comparison by currency and country which may be able to tell us if one countries currency or policy is more beneficial to crowdfunding success than another’s.

2.Tending graphs that compare outcome by date: I think dates are really important because they capture temperament, financial viability and I believe certain campaigns would benefit from soliciting funds during an economic downturn (charities for example)

3.Bar charts to compare by method of backer investment: Sources of funding can be further investigated to see if there is a correlation between source of funding and categories of funding.

4.Tables that show how funds were solicited: Electronic funds are definitely going to get a certain level of response but are there populations who are being left out.

5.Gleen data that extracts categories success to initial goal to find if goals can be optimized, and whether there is a science to how much an initial goal should be set to.

6. Tables and graph to chart categories that are more advantageous than others for backer (i.e., tax breaks, symbolic, cultural, prestigious, good for laundering money)

**Statistical Analysis Questions:**

**1.Use your data to determine whether the mean or the median better summarizes the data.**

When you compare the mean value of successful campaigns to the mean value of failed campaigns you notice a stark difference in the ratio of the 2 means: 851 (successful mean) is 1.45 times larger than 586 (failed mean) while 7295 (successful maximum) is only 1.2 times larger than 6080 (failed maximum).

When the medians (average of backers) are compared to the 201 (successful median) to 115 (failed median), successful backer count is 1.75 times larger than the failed backer count. I believe the numbers are not dramatic enough to draw a conclusion based on the mean or the median. Neither the mean or the median are indicators by them selves can be used with confidence to draw anything but conjecture on the topic.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **successful** | **differences** | failed |
| **Mean** | 851 | 1.45% | 586 |
| **Median** | 201 | 1.75% | 115 |
| **Minimum** | 16 |  | 0 |
| **Maximum** | 7295 | 1.2% | 6080 |
| **Variance** | 12165647 |  | 16427630 |
| **Standard deviation** | 439 |  | 3040 |

**2. Use your data to determine if there is more reliability with successful or unsuccessful campaigns. Does this make sense why/why not?**

Reliability is statistics is defined by Wikipedia (<https://en.m.wikipedia/wiki>) as the “overall consistency of a measure… A measure is said to have a high reliability if it produces similar results under consistent conditions.”

Successful campaigns had a total of 7295 out of backers a total of 200,442 total backers which approximately 36% of the total. Judging by the huge variances there and the small variances between the means and median seems to indicate a certain level of *unreliability*: the two figures (a difference of 193,147) which is inconsistent with the less dramatic differences between the mean and the medians (1.45% and 1.75% respectively). Based on the jump in the differences and considering the definition of reliability, I conclude that there is *not* much reliability in either figure. It does however make sense that there be a consistence, or insignificant deviation in the statistics across data groups for to merit them reliable.