**Crowdfunding Book Analysis**

Data Analytics Bootcamp

Module 1 Challenge

Khrystyne Clifton

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

1. Entertainment options (Theater, Film & Video, and Music) are more successful in crowdfunding campaigns than other categories.
2. The vast majority of campaigns that are crowdfunded come from the US.
3. June, July, and September at the best months to launch a crowdfunding campaign.

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

1. Roughly same size sample for each year excluding 2020 (84-107 projects)

* This may not be reflective of the popularity of crowdfunding vs. other funding methods year over year.

1. Successfully funded projects do not indicate success/popularity after funding

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

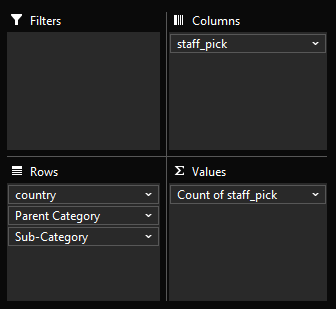
Percentage of projects that have each outcome by month. This shows that June, July and September are the best months to launch a crowdfunding campaign.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Row Labels | canceled | failed | live | successful |
| Jan | 7% | 39% | 1% | 53% |
| Feb | 9% | 35% | 0% | 56% |
| Mar | 5% | 38% | 0% | 57% |
| Apr | 1% | 38% | 1% | 59% |
| May | 3% | 41% | 2% | 53% |
| Jun | 3% | 32% | 1% | 63% |
| Jul | 4% | 33% | 1% | 62% |
| Aug | 9% | 41% | 1% | 48% |
| Sep | 7% | 32% | 0% | 62% |
| Oct | 8% | 33% | 1% | 58% |
| Nov | 4% | 35% | 4% | 58% |
| Dec | 8% | 38% | 4% | 50% |
| Grand Total | 6% | 36% | 1% | 57% |

Pivot table- Staff picks by country, Category and Sub category.

What are the trends regarding staff picks? Is there a noticeable preference for projects from one country over another?

Result: Plays and documentaries from the US are the sub-categories recommended the most by staff



A percentage table based on pivot table above shows us that though the US has the highest number of projects that are staff picks, CH is actually the clear leader by percentage of projects launched that landed a staff recommendation.

|  |  |  |
| --- | --- | --- |
| Row Labels | FALSE | TRUE |
| **AU** | 100% | 0% |
| **CA** | 93% | 7% |
| **CH** | 91% | 9% |
| **DK** | 94% | 6% |
| **GB** | 98% | 2% |
| **IT** | 96% | 4% |
| **US** | 95% | 5% |

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

Median better summarizes the data overall. When calculating the standard deviation based off of the median, the SD is smaller.

Median Calculation:

37.2=((0-186)2)/930

Mean Calculation:

600= ((0-747) 2)/930

In the examples above, 930 is the total number of records, 186 is the median, and 747 is the mean.

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

There is more variability in the numbers of backers for successful campaigns. SD is higher for successful campaigns than failed campaigns. This makes sense since some of the campaigns reached 1,000% of their goal or more. Failed campaigns only have a range of 99.684% to 0%.