Asal Valisoltani

Excel assignment- Crowdfunding Analysis

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

Based on provided data, US has the most campaigns (76%). The remaining 24% are in Australia, Canada and Europe (England, Denmark, Switzerland and Italy).

The companies operating in the theater, film and video, and music industries had the most fundraising campaigns with %34, %18 and %18 respectively. If we breakdown data for these 3 industries, we can see as below:

|  |  |  |
| --- | --- | --- |
| Theater | %54 Successful | %38 Failed |
| Film & Video | %57 Successful | %34 Failed |
| Music | %56 Successful | %38 Failed |

The rate of successful campaigns is around %50-%60 for all the industries. The lowest number of campaigns belongs to Journalism industry with only 4 campaigns in US which all of them were successful.

If we look to the sub-categories for the highest 3 ranked industries:

* Plays is the only sub-category of theater, which has the most successful campaigns (%54).
* Film and Video is the next highest industry in which Drama, Documentary, and Animation have the successful rate of %60 , %57 and %62 respectively.
* In the music industry, the most successful campaigns belong to Rock with %58 and Indie Rock with %51.

In some countries some specific industries are more successful. For example, in Australia, technology industry has more successful campaigns compare to the Music.

In one of pivot tables, we can see campaigns as a function of time. Data has been collected from 2010 to the first month of January 2020. In average for each year same number of campaigns had been run, about 100 per year.

In June and July, we see an increasing in number of successful campaigns. July has the most successful campaigns, (58). And the less numbers for failed campaigns are in September, (23). According to this graph, June and July might be a good time to run a new campaign and probability of success is optimal. Success rate hit a low in August while at the same time both failed and cancelled campaigns reach a high. Hence August might not be a good time to run a new campaign.

Surprisingly, with increasing fund, rate of successful has been increased too. From $15000 to $49999 with successful rate more than 67% to 100%.

The highest failed rate belongs to the campaigns with more than $50000 (%53).

1. What are some limitations of this dataset?

That would be beneficial if we had distribution of data for different states in US, that could help to determine which states could be a better target for each campaign industries.

In addition, if we have some data about genders and age of backers, we could have better analysis.

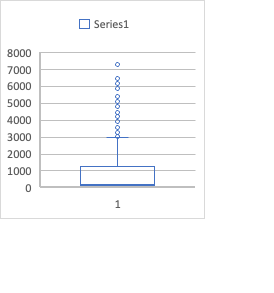
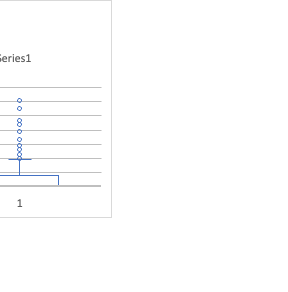
1. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

We could have considered the relation between successful rate and period for each campaign in each industry. Also, we could’ve filtered data by year and country.

We can explore the relationship between average donation and chance of successful. We could have defined a range for average donation and analyse data based on that.

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

According to box and whisker plot, the data distribution is skewed, and median give a better insight in distribution of data.



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

In successful cases we have more variability. Standard deviation is 1266.24 and it shows more variation compared to failed cases with standard deviation of 959.99. In successful cases we have more outliers and wider interquartile range.