Crowdfunding Analysis

Crowdfunding Data Conclusions:

With the given data for the 1,000 crowdfunding campaigns one of the first conclusions that can be derived is that based upon the data that most of the successful campaigns are in the film & video, music, and theater categories. These 3 categories have a significantly larger success rate for crowdfunding funds then any of the other categories out of the 9 total presented in the data.

The second data driven conclusion is that theater plays are the most successful at crowdfunding which can be shown in the number of successful plays crowdfunding. It would appear that more people are interested in supporting things that they enjoy doing which is clearly plays, listening to music, or watching movies/films. These categories also try to crowdfund significantly more than others which helps to show the correlation between crowdfunding more often and as a result of doing so getting more people engaged to want to help fund it.

The third data driven conclusion is that crowdfunding is done more in the US and more successful in the US. This will help to show us that if we need to crowdfund, we should be doing so in the US as that will give us the most opportunity at successfully funding our ideas/venture. Crowdfunding campaigns in the US gets your ideas/ventures in front of more people which in turn proves to be more successful at bringing on more people who in the future could help you fund other ideas within the company as you move forward.

Limitations:

One of the limitations of the dataset I see off hand is the limited number of campaigns provided. At having only 1,000 crowdfunding projects, while giving us an idea of how the campaigns success/failure rate is, if we had 10,000-50,000 campaigns, we could better provide the client(s) with more of an accurate depiction of the data and trends.

We also are provided a lot of the data in the form of foreign currency which doesn’t allow for us to compare them in the same currency would could keep us from seeing the correlation between the various countries as we are only provided the amount they wanted to raise in their local currency. Given that some currency is valued lower than the US dollar is would be imperative that we have some way to have all foreign countries currency converted to the US dollar format if we are presenting this to a US based company, or vice versa, convert the US currency to their local currency so that they can have a true idea of how much money was actually raise via the crowdfunding and can compare the categories based on the same currency.

The last limitation of the dataset is that a lot of the campaigns time conversion start and finish dates are very small among the data give. We need to have data provided that shows campaigns of longer than 1 month from start to finish so that we can take those into account as well.

Other Tables/Graphs:

Some of the other tables and graphs I might have used for a presentation of this data would be line and pie charts. I would use the pie chart to compare the different countries to the entire data set so that you can show how each country performs when crowdfunding. The line chart can also show this for as well. I think the addition of these would allow us to showcase a lot of the data better. We can use the pie chart to show countries to other countries, we can show categories vs other categories, we can show the different average donation sizes vs the whole.

As far as tables go, I would do tables for each respective country so that in case I was asked to show how the different countries perform I could do so at in moments notice. I would also provide tables for all the categories and sub-categories.

Bonus:

In my opinion due to the small amount of data provided and after looking at the stand mean and median for both successful and failed, I would most likely use the mean for both. After putting the data into graphs, I can see that both successful and failed have a lot of outliers which shows me that I should be using the median which would give me a truer depiction of the data provided.

Looking at the data provided I can see that with successful the mean(851.15) and median(201) are very far away from one another while the failed mean(585.62) and failed median(114.5) are a lot closer. So based on that I feel the successful would have more variability. Whenever there are large gaps between mean and median you will start to see the data is more varied with more outliers. Since you want to provide the client with the most accurate depiction of what the data is doing you would want to have the smallest amount of difference between these two numbers so that you can predict your what you are providing the client with the highest and most accurate information you can.