Report Module 1 Challenge

Questions:

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
2. What are some limitations of this dataset?
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Answers:

1. We can see that the three categories with the most projects are: film & video with 178 projects, music with 175 projects and theater with 344 projects. This last one tends to be the category with the most projects, assuming that users tend to be more inclined to create projects related to plays.

Just as we found the most popular projects, we can also find the category with the least impact, which would be journalism, which of the 1000 projects only the 0.4% are related to this category, but in the same way with 100% of success.

Regarding the monthly graph, we can observe a maximum number of successes in the month of July, followed by the minimum in the month of August. This does not happen with failed projects, but it may be an indicator that these months may depend on the success of the project.

1. It may be that one of the limitations of this dataset is the country of origin of the backers of the projects, it may be that the tendency is that many are from the United States but if you are from another country you may get more sponsorships from your country because you base your project on the needs that you see in that place and you may not get a lot of financing from other countries because they do not know the story you are trying to tell, or your project may be is directed at a foreign audience and you want to reach more attention with strategies from other projects already carried out.
2. We could create a graph per year and see what the trend is in each one of them, see if there has been any increase in a specific category, which has gained strength in recent years, for example, the technological advances of the digital age opened the door to more people who want to develop projects like film & video that with today's tools is a very popular project, in the same way, if any category is no longer taken many attention and find the reason by which people opt for these new trends and ignore others.

* Determine whether the mean or the median better summarizes the data.

The obtained data was the next:

|  |  |
| --- | --- |
| successful | **565** |
| **Mean** | 851 |
| **Median** | 201 |

|  |  |
| --- | --- |
| failed | **364** |
| **Mean** | 586 |
| **Median** | 115 |

It seems to me that the mean gives us more useful data, since it shows us that half of the projects, in the case of successful ones, are financed by <=201 backers, while the other half, the maximum number of sponsors would be 7,295, a huge gap compared to the first half, the same occurs with failed projects, we have an average of 115 and reach a maximum of 6080. This data can help us when preparing the budget and the goal we want to reach.