CROWDFUNDING BOOK

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

* Theater/plays have the greatest number of crowdfunding project applications – this category will be highly competitive but at the same time has higher chances of having a successful funding.
* Journalism seems to be the most successful, with 100% approved projects and games as the least successful among all the categories with 47.92% failure rate, and film has the most projects that are live (5 out of the 14 live projects).
* Most of the projects were created in January, probably because it is a project that they want to advertise and launch within the year, and in July, probably because this is the start of a new financial year, which might mean that more fundings or budget are available.

# What are some limitations of this dataset?

* Percent Funded is not actually a good metric to measure the success of a project.
* A project being live does not actually correlate to it being successfully funded by the crowdfunding platforms – they might have other agenda as to why they applied to these avenues, such as to gain exposure (or to have “free advertisements”) or to check initial feedbacks of their projects.
* Cancelled projects are also misleading in providing proper metrics on analysing if a project was a failure or a success. Providing the reason why it failed would be helpful in analysing which type of projects are more likely to fail.
* The number of days a project needs to be funded is not the same for all. Some projects have a few days to run their funding in the platform, while others have at least a month to do it. This must be considered as well when investigating as to why a project failed or succeeded, or as to why they did not achieve their goal.
* The crowdfunding platform used for each project should be included as well, as some might be more popular than the other. This might have affected the success of the project, or in achieving the fund goal, if they used a less or more popular platform.

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

* A table or (pie) chart of the fundings made by each country to know which ones have backers that are more likely to contribute.
* A graph of the average days a project gets funded and the percentage of the goal (per category) that got funded should be a good tool to see the typical length a project should be in a crowdfunding platform.

BONUS

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

For this particular data, median better summarises the backers count to set the expectations of the project creators looking for funding. It is closer to the actual most number of backers of both the failed and success projects. Using the mean would mislead the creator as there are only a handful of projects that are funded by a high number of backers, which skews the data as the difference is too great compared to the usual number of backers per project, which are 0-470 for failed projects and 0-540 for successful projects.

# 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 data variability with the successful projects than the failed projects. This makes sense as there are more successful projects than failed ones. Taking the mix of the category types in those two data sets into consideration would provide more clarity as how it affects their variability. There are project types that receive more attention and funding than the other and since there are more successful projects, then this means that there is more variation in the category types included in those projects.