Sofia Ysita

Data Science and Visualization Bootcamp

Northwestern University

**Crowdfunding Analysis**

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + The greatest number of projects were theatre with subcategory of plays with 34.4% of all the projects. Theatre had 54.3% successful projects, 38.4% failed, and 6.7% were canceled. Theatre is the category with the most successful, as well as the most failed, and canceled number of projects.
  + Journalism projects were the least projects done with only 4 projects that amounts for 0.4% of the total of projects. Although there were less projects, 100% of those projects were successful.
  + January was the month with the most failed projects, while July was the month with the most successful projects. There were no more than 8 canceled projects in a month throughout the year.
* What are some limitations of this dataset?
  + You can’t see the reason why the projects did not make it to the goal.
  + Average donation does not tell you anything about the success of the project.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + Outcome of projects by country: you can see the relation between the country and the outcome of the projects
  + Category of projects each month: You can see what category of projects are more or less successful in each month

**BONUS**

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

The Median better summarized the data since the distribution of data values is skewed.

* 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 with successful campaigns. It makes sense because we have more successful than failed campaigns.