Questions about Crowdfunding Data

1. From the crowdfunding data tables and charts we created, we can draw at least three conclusions from this data. One such conclusions we can create is that an overwhelming amount of successful projects’ parent category was theater, and the sub-category that was most successful was plays. Another conclusion we can create is that from 2010-2020, July has been the month with the most successful campaigns with a big dip occurring in August, being the month with the least successful campaigns. A third conclusion we can draw from the data is the US was the only country to contribute to the journalism category with 4 projects, all of which were successful.
2. A limitation of this data set includes not being able to see where most of the money was coming from for the projects, as in, where the backers were located. Another limitation is that although we know if the project reached its goals or not and the deadline of each project, we don’t know how fast it took a project to reach its goal, if it did, within the time span of its crowdfunding process.
3. I think having a table showing the number of projects that have been in the spotlight vs the amount that weren’t in the spotlight and the percentage that have been successful in each category would show us if projects that are in the spotlight get successfully funded. I also think that having a table that shows which category of project was created in a certain month would also help to explain trends in outcomes during specific time periods.

Questions About Backer Analysis

1. I think that the median is better suited to summarize the data because using the mean is more prone to be manipulated by outliers, whereas the median is not affected much by outliers since it is the middle point of all the data.
2. There seems to be more variability with successful campaigns versus unsuccessful ones, which makes sense because each project has different amounts donated by a different number of backers to become successful and generally, more backers mean a higher likelihood of success, which correlates with the maximum number of backers in the successful category being significantly higher than the maximum number of backers in the unsuccessful category. With unsuccessful campaigns, it seems more difficult to gain enough backers which leads to a lower variability since there aren’t enough backers to make the project successful.