Ellen Malaki

Data Analytics

Module 1 Challenge - Crowdfunding

Given the provided data, we can conclude that theater based crowdfunding campaigns are overrepresented when compared to other categories, with plays being the largest sub-category within the parent category. Additionally, we can conclude that crowdfunding is not a commonly used source of funding for journalism. We can also conclude that campaigns beginning in late June and early July have the highest chance of being successful. Another conclusion is that the number of failed campaigns has generally decreased over the years, likely due to more people becoming aware of crowdfunding campaigns in general.

One limitation of this dataset is that some campaigns may fall under one or more parent categories and sub-categories. The data may be skewed by the necessity of choosing one category and sub-category for each campaign. Another limitation is the fact that not all campaigns are using the same currency.

Another interesting graph we could create could look at the number of outcomes (failed, successful, live, cancelled) based on the time between the start date and the end date of the campaigns. This would provide some insight into whether longer or shorter campaigns are generally more successful. It would also be interesting to compare the percent funded or the average donation of each campaign based on the parent category and sub-category. This would provide information on how much people are willing to donate to certain causes.