Analyzing the Data

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
2. The most engaged category with the highest numbers of failed and successful campaigns is the theater with a total number of 319 and they were all part of one sub-category which is “plays” , while the least engaged category with no fails and low number of successes is journalism with a total number of 4 campaigns and they were all part of one sub-category which is “audio”.
3. The most successful campaigns were conducted in July with a. number of 58 successful campaigns, while the lowest number of successful campaigns were conducted in August with a number of 41 successful campaigns.
4. The most failed campaigns were conducted in January with a. number of 36 failed campaigns, while the lowest number of failed campaigns were conducted in October with a number of 26 failed campaigns.
5. The outcome of the campaign is directly related to the percent funded that we calculated, when we look at the failed outcomes, we see that the percent funded is bellow or equal to 100 ( there are only 3 campaigns that equals to 100), on the other hand the successful campaign’s percent funded is always bigger or equal to 100 (there are only 4 campaigns that equals to 100).
6. What are some limitations of this dataset?

Some limitation we can consider is the country the data was collected from ( each country has different currency so when we look at the money collected or donated, we can’t really compare because they’re not uniformed) and the date the event was launched at ( we have data collected from 10 years ago with different circumstances effecting it, so when we compare the data from different decades we should consider the environment that effected it).

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

We can look at each campaign’s online presence and check how much time they spend on marketing on social media. We can analyze if they’re presence gave them a successful outcome or failed outcome. We can add the age of the target audience and look at what age group is more engaged in these campaigns and how the age outcome changed from one age group to another.

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

The mean of the successful campaign is 265 backers-count higher than the ones of the failed campaigns. The standard deviation of the number of backers in the successful campaigns are larger than the failed ones, so the data is more spread out from the mean. The median of the successful campaign is 86 backers-count higher than the ones of the failed campaigns. This data helps you visualize the overall summery of the data you collected, and it helps you understand it better.

* 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 because there is a larger difference among the scores of the backers-count and this makes sense since the number of successful campaigns is higher than the failed ones, and we concluded that by looking at the mean, median, variance, and standard deviation of both campaigns.