Crowdfunding Questions and Conclusions

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   1. Crowdfunding campaigns by category generally have a 50/50 success rate, with the exception of the journalism category. Each category has successful campaigns.
   2. Theater/Plays are a high percentage of the data set, ~30% of the total campaigns.
   3. Generally, as the goal amount rises so does the chance for failure. However, due to the variation in number of campaigns for the different goal levels it’s difficult to determine if with a larger data set, the failure percentage would rise over the successful percentage earlier than $50,000 range.
2. What are some limitations of this dataset?
   1. The unbalanced data set is skewing the data. Some examples are the small number of journalism campaigns leading to a 100% success rate, and the significant number of campaigns for plays creating outliers and false campaign spikes (June).
   2. Due to the low number of campaigns in the different goal levels, it’s difficult to determine truth around success/failure for the campaigns in the $10,000-$49,999 range.
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   1. A few bar graphs charting percent funded, backer count, and categories to help determine skewness of the data. Due to the mean/median analysis this data doesn’t have a normal distribution.
   2. It would be beneficial to create box plots using the percent funded category to determine any outliers.
   3. In addition to the date created line chart it would be interesting to see another line chart comparing the date ended to see if the end time of a campaign has any effect on funding.
4. Does mean or median better summarize the data?
   1. The median better summarizes the data since it is different from the mean. The higher backed campaigns are pulling the mean up and the median demonstrates that the mid-point of the data is much lower than the mean suggests. Like the mean, the median also represents that overall, more backers are funding successful campaigns than failed ones (median of 201 backers for successful campaigns versus median of 114.5 for failed campaigns).
5. Do successful or failed campaigns have more variability?
   1. Successful campaigns have more variability in backer count.
      1. Does this make sense? Why or why not?
         1. It makes sense that backer count would be variable since each goal for any given campaign is different and backers can pledge different amounts each time. Since successful campaigns are funded with more backers on average, more variation is expected. Also, failed campaigns tend to remain at the lower end of backers (and funding) giving a sad sort of stability.