Module 1 Challenge

# Crowdfunding Data Summary

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + Theater, specifically plays, makeup the largest proportion of crowdfunding projects in this dataset (34%). Additionally, they make-up the largest proportion of projects within each country.
  + 57% of projects were successfully funded, and were more likely to be funded in the months of Jun or July.
  + There were two goal ranges that had higher fail rates than success rates, these were the $10,000-$14,999 and $50,000+. It would be advisable to avoid these ranges when setting goals, even though the former range has a small n size (n=9). A goal range of $1,000-$4,999 had the highest rate of success when accounting for n size (n<20).
* What are some limitations of this dataset?
  + There are only 1000 cases
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + Additional tables and graphs that would be useful are how long the project was active, as in, what was the time between the launch date and deadline, the average donation by category and subcategory, and showing the staff pick and spotlight categories by outcome to see what kind of impact those things have on the success of a project.

# Statistical Analysis

* Use your data to determine whether the mean or the median better summarizes the data.
  + I would say the mean better summarizes the data. It is more representative of the middle of the curve because there is a large variance and the dataset is skewed to the left (with several outliers).
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
  + They were pretty close, but there is more variance with successful campaigns. There is more potential for variance among campaigns that were successful because there were more successful campaigns.