Module 1 Challenge

3 Conclusions on Crowdfunding Campaigns:

* Theater campaigns are the most common parent category of crowdfunding campaigns overall and in most countries (except in China and Great Britain). This is closely followed by film & video and music campaigns, which occur at a relatively similar rate. However, the success rates for all of these campaigns are relatively similar.
* Plays are the most common (over 1/3) sub-category of crowdfunding campaigns.
* Overall, crowdfunding campaigns have a 56.5% success rate and a 36.4 failure rate – similar success/failure rates can be observed in a majority of the category of campaigns. Crowdfunding campaigns have a higher success rate during the summer months (June – July) but face a dip in success immediately after in August.

Limitations of the Dataset:

* A majority of the crowdfunding campaigns are located in the United States so the dataset is likely not truly representative of all the included countries. We also have data from only a select number of countries (might be where the platform is used most but no information to support this).
* A majority of the crowdfunding campaigns fall into 3 major categories – theater, film & video, and music – meaning that all other campaign types may not be as well-represented in the dataset (unless those are the most common types of campaigns that occur).
* We do not have much information about fundraisers or contributors or the prevalence of the crowdfunding platform itself in all the different countries represented.

Additional Tables or Graphs:

* We can look at campaign lengths and success/fail rates to determine optimal time for a campaign to stay open.
* We can look at staff pick and/or spotlight and success/fail rates to determine if either factor is correlated to the outcome of a campaign.

Statistical Analysis

The mean does not summarize the data very well because we have very large numbers for variance and standard deviation. This means that the values in our data set are very spread out over a wider range and are not close to the mean. The median likely summarizes the data better than the mean.

There is more variability with successful campaigns compared to unsuccessful campaigns as those have higher variance and standard deviations. This makes sense because the range of values of successful campaigns are much larger than the range of values of unsuccessful campaigns, as evident by the difference in their minimum and maximum values.