**Conclusions drawn about crowdfunding campaigns:**

* Campaigns that are started in June or July may have a higher chance to be successful than campaigns started in other months.
* The “film & video”, “music”, and “theater” parent categories are by far the most popular campaign categories, with nearly 70% of all campaigns being one of the three.
* The “plays” sub-category is the most popular, with 34% of campaigns being for plays.

**Limitations of the dataset:**

* We were only able to calculate an average donation per backer, therefore it is impossible to know if one backer made a huge contribution that skewed the average donation value of a campaign.
* A disproportionate number of the campaigns were under the sub-category of “plays”. This likely makes the data fairly accurate for predicting the outcome of crowdsourcing a play, but not as useful when predicting the outcome of a campaign of other sub-categories.

**Other tables/graphs and insight they may provide:**

* A comparison of the amount of time a campaign was live vs the “outcome” would provide a sense if there was a correlation between time spent collecting donations and success or failure.
* A comparison of “staff\_pick” and “spotlight” vs the “backers\_count” and “outcome” could potentially show if the staff pick and spotlight is an effective way to increase the odds of recruiting more backers and having a successful campaign.

**Comparison of mean and median as a summary of the data:** The median is a better representation of the backer count’s effect on the outcome of a campaign. In both successful and failed campaigns, the maximum number of backers is several standard deviations higher than the mean. This indicates that the data is heavily skewed right, and the mean will be strongly influenced by outliers. Therefore, the medians which are not as easily swayed by outliers are the better summary.

**Comparison of variability in successful and failed campaigns:** The variance, and therefore variability, of successful campaigns is higher than that of unsuccessful ones. This makes sense due to the fact that campaigns can be funded in excess of 100%. Failed campaigns have to fall somewhere in the range of 0 to 99% funded, whereas successful ones will be 100% or more funded (some earning well over 10x their goal). The range of successful funding percentages being so much larger than the range of failed ones lends itself to having a more varied data set.