# Conclusions regarding crowdfunding campaigns

Below you will find three primary conclusions we can draw from the data available to us within the Crowdfunding Campaign data set provided.

* The most straight forward conclusion drawn from the crowdfunding campaign data when illustrated in a bar graph format is that crowdfunding campaigns are more likely to be successful than not.
* There is an obvious trend in the success of crowdfunding campaigns across the globe that revolves around the specific time of year in which the campaign is launched. Campaigns launched in August have historically performed poorly. This is illustrated by the peak in both the failed and canceled campaigns launched during August. This conclusion is further supported by the dramatic dip in the number of successful campaigns launched in the August timeframe. The definitive conclusion is that ensure the best results possible, crowdfunding campaigns should not be launched in August.
* Another conclusion that can easily be drawn is that specific categories tend to lend themselves to the crowdfunding arena, specifically theater. It can be concluded from the data provided and graphs constructed that the theater industry is the dominating player within the crowdfunding industry, accounting for an extreme make-up of the industry.

# limitations of the crowdfunding campaign dataset

* The data provided illustrates a common problem that occurs in working with real world data, in that there tend to be extremes. As noted in the data and graphs the theater industry is by far the category that dominates the crowdfunding landscape. The overwhelmingly large number of campaigns initiated in this category appear to create an outlier in the data and thus the results from the theater industry should not be considered when conducting analysis on the other categories included in the study as it inaccurately skews the data. Outliers can impact many statistical evaluations such as the mean resulting in an inaccurate depiction of the data exclusive of the outliers.

# POTENTIAL OTHER TABLES AND GRAPHS AND ADDITIONAL VALUE ADDED

* A scatter plot would have been a beneficial graph to use considering the large amount of data present. A scatter plot is beneficial in that it easily draws attention to potential outliers. Scatter plots are also beneficial in that they can illustrate a correlation between two variables, either positively or inversely correlated. Scatter plots are also valuable in showcasing trends in the data. This is most illustrated through the addition of a trend line within the scatter plot.

# statistical analysis: the mean versus median and variability

* The data provided is best summarized by the median as opposed to the mean. The reason for this is that the mean is subject to impact from outliers, which this dataset has, which ultimately skews the data and is reflected in the mean. The median on the other hand, which still considers all data sources, is not as heavily impacted by the extreme nature of the outliers. Their impact is somewhat minimized with the use of the median as the central tendency indicator.
* The box plot gives a clear and quick illustration of the difference in variance between the successful and unsuccessful campaign data. A key feature of a box plot is the size of the actual “box.” The larger the “box” then the more variability exists within the data. This makes for a very quick analysis of the difference in variance amongst data types via a quick visual scan of the box size. The data and graphs indicate that there is a much larger variance within the successful campaigns. I feel this data is indicative of the real work and makes sense. I would assume that many campaign failures have multiple commonalities, specifically focusing on a lack of organization leading to an early failure rate. The success of the other campaigns, simply based on the time it takes to get to market would allow for a higher level of variance amongst themselves.