Module 1 Challenge

# Crowdfunding

1. **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**

Conclusion 1: The Count\_per\_Category and Sub-Category show that many campaigns were for theater projects, especially plays. Perhaps because there were so many play campaigns (344), while many failed (132), there were many successes as well (187). Film/video and music categories were a close second and third.

Conclusion 2: The pivot table and line graph of Outcome\_by\_Year, does not highlight any major correlation between campaign creation and rates of success, failure or cancelation. However, there is an interesting spike of successful campaigns that were started the summer months of June and July.

Conclusion 3: Journalism campaigns represented a very small fraction of campaigns, but they were all successful.

1. **What are some limitations of this dataset?**

It’s difficult to find a relationship that shows what made the campaigns successful. More data about the method of crowdsourcing might be helpful. The difference between theater campaigns and other campaigns is so great that some of the nuance is lost for those other campaign types.

1. **What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

A table that shows how long a campaign lasted might show a relationship between duration and the chances of reaching the goal. Likewise, a chart showing the method of crowdsourcing might highlight why a campaign was successful or not.

Statistical Analysis

**Mean vs Median**

For both the successful and failed crowdfunding campaigns, the median is going to be the best value for summarizing the data. This is because medians are useful when there are extreme scores present as it is not influenced by outliers or extreme data. There were *significant* differences in the number of backers for each campaign. For example, for successful campaigns, the summary table shows us that the minimum number of backers was 16 while the maximum was 7,295. Similarly, the minimum provided on failed campaigns was 0 while the maximum was 6,080. Therefore, the medians of 201 for successful campaigns and 114.50 for failed campaigns are the best values to use in summarizing the data.

**Determining Variability**

There is greater variability with successful campaigns. This makes sense because there were more backers overall for successful campaigns (shown by both the max values and the median number of backers). Compared to the lower number of backers of failed campaigns, there will be more distance between each number of backers from the mean.