Henry Cutler

6/27/22

Module 1 Challenge

* Conclusions
  + **Conclusion** **#1**: Based on the dataset given, there are more crowdfunding campaigns for theater, film, and music, with an additional emphasis on plays. Whether the campaigns are successful or not, it looks like people are still more willing to back the arts than other areas.
  + **Conclusion** **#2**: There doesn’t seem to be a strong correlation between successful campaigns and time of year. Looking at the chart of campaigns over the course of the year, there is a slight uptick of successful campaigns in early summer and more failures at the end of summer, but it does not look to be statistically significant.
  + **Conclusion** **#3**: Crowdfunding campaigns whose goals are under $50,000 are much more likely to succeed than those that are asking for $50,000 or more. Although we may not have enough data to show the middle of this range ($10,000-$49,,999), there is certainly more success in the $1,000 to $9,999 range than over $50,000.
* **Limitations** **of** **the** **dataset**: Although we are looking at 1,000 different crowdfunding campaigns, over 75% of these are originating in the United States. If we wanted to get a global view of these campaigns, we would need a much larger sample size from outside of the US. Additionally, we don’t see many projects that have fundraising goals between $10,000-$49,999. Although these campaigns seem to be largely successful, there may not be enough data to show causation.
* **Additional table/graphs that would provide more insight:** During this project, we looked at how successful (or unsuccessful) crowdfunding campaigns are based on their cultural category, sub-category, time of year, funding goal, and even number of backers. One more thing that could be important to look at is *how long* each crowdfunding campaign lasts. There may be a difference in terms of success for a campaign that lasts one week versus one that lasts six months. Within this, we can look at timeframe by funding goal to get even further analysis. We would do this by creating a scatter plot with length of campaign on the x axis and percent funded on the y axis. We can then insert a trend line to show whether or not more time to fundraise means a higher percentage of the goals hit.
* **Bonus Statistical Analysis:** In terms of number of backers, there are a few outliers on the high and low end for both the successful campaigns and the unsuccessful campaigns. Due to these outliers, especially on the high end, we should look at the median number of backers to summarize our data. The median number of backers for successful campaigns is 201 and the median number of backers for unsuccessful campaigns is 115. There is more variability in the number of successful campaigns than unsuccessful campaigns. This makes sense, as there are simply more successful campaigns and they have a larger range of backers.