Crystal Burkhardt

Challenge 1: Crowd-Funding Analysis

April 20, 2023

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

We can conclude that when speaking in terms of general category, Theatre receives the most donations, followed by Film & Video, and then Music closely behind in third. We can also conclude that the most popular crowd-funding campaigns are to support plays, but many of the campaigns started for plays end up being canceled or failing (45% of all campaigns are unsuccessful). A third conclusion that can be drawn is that the highest number of successful campaigns are held during the summer months of June and July.

1. What are some limitations of this dataset?

Some limitations of this dataset are that we don’t have any demographical, geographical, or source information on the backers. In other words, age, sex, location, or the channel(s) that were used to secure their donation pledges (social media, in person, email, networking events, etc). This type of information could help campaign marketers better reach their most successful audiences.

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

Another table/graph that we could create would be to analyze the percentage of total campaigns that ended up failed, canceled, or both values combined to generate a more realistic insight into the chances that a campaign within a center category is going to fail, be canceled, or either of the two which both essentially equate to non-successful campaigns. This could spark an investigation by campaign managers into why particular campaigns end up failing or being canceled, and help the managers/marketers generate ideas on potential market research to be conducted, and also solutions to increasing success rates. This potential insight seems particularly valuable in particular when looking at the sub-categories because if you only consider the grand total of campaigns, you’re only really considering the popularity of certain sub-category campaigns and ignoring the fact that a more accurate assessment of the overall success/effectiveness of the campaigns by sub-category would be to calculate a percentage of successful campaigns out of total started. Another table/graph that could be ran would be to determine if particular companies ran multiple campaigns in a year (consolidate column 1 by name), and also to calculate the percentages of the total funds raised by each category or sub-category to see which categories are the most successful from a monetary standpoint versus numerical count of campaigns.

Statistical Analysis: I believe that it would be wise to use median values in this case as opposed to the means. With the standard deviations both being higher than the means, this tells me that the data is assymetrical and right-skewed by more positive outliers. As a result, if data does not have a normal distribution, the mean cannot provide a good measure of central tendency. There is a lot of variability in both the successful and unsuccessful campaigns. However, there is more variability with the successful campaigns. I think this makes sense because the data set is larger and the variance of the backer counts should vary more when these campaigns actually raised enough funds to met their goal amounts, which also vary, making the number of backer counts needed to meet varying goals amounts more diverse.