Juhi Chinthapatla

Module 1 Challenge

1. **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?** Firstly, theater has the greatest number of campaigns, while journalism has the least number of campaigns. All the journalism campaigns are successful. On the contrary, theater has almost similar number of successful (187) and failed campaigns (132). Secondly, plays, animation, and video games are all tied in sub-category for having the highest number of live campaigns. Thirdly, July is the month with the highest number of successful campaigns, and August is the month with the highest number of canceled campaigns.
2. **What are some limitations of this dataset?** A limitation of this dataset is that it is only a small sample of campaign information out of all the campaign information gathered from crowdfunding platforms. Additionally, since this is fake data generated by edX Boot Camps, it is hard to extrapolate analysis from this assignment to how campaigns perform using Kickstarter or Indiegogo.
3. **What are some other possible tables and/or graphs that we could create, and what additional value would they provide?** A possible table to create is of the names of the backers and the outcome of the campaigns they supported. The table will show which backers have either good or bad judgement in regards to the campaign they support and what the campaign’s outcome is.
4. **Use your data to determine whether the mean or the median better summarizes the data.** The data set is skewed right with many outliers, which makes median the better summary statistics over mean to summarize the data.
5. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense?** **Why or why not?** There is more variability with successful campaigns compared to unsuccessful campaigns. This does not make sense because high variability implies there is an inconsistent number of backers pledging money for a campaign. For a campaign to be successful, there needs to be consistent number of backers. So, a successful campaign in theory should have low variability, while a failed campaign should have high variability. Therefore, the summary statistics based on the given data set does not make sense.