Alec Mintz

Module 1 Assignment

1. Given the data, three conclusions that can be drawn are:
   1. The largest number of crowdfunding campaigns fall into the film/video, music and theater categories. Conversely, the least number of crowdfunding campaigns fall into the journalism, photography, and food categories.
   2. Crowdfunding campaigns with goals between 15,000 and 35,000 are successful at a rate of 87%, while campaigns with the goal of 50,000+ fail or are cancelled at a rate of 63%. This indicates campaigns with goals between 15,000 and 35,000 have a 50% greater chance of achieving success.
   3. In the United States, the most popular campaigns are plays (sub-category), which fall into the category of theater. However, even though they are the most popular, they are only successful 54% of the time.
2. Some limitations on the dataset include:
   1. Selection bias – Since crowdfunding platforms might only attract certain categories of business/products, it is not a full representation of the broader community since other categories of business/products tend to shy away from using crowdfunding campaigns.
   2. Since the dataset only contains data for campaigns between 2010 – 2020, it is not a full representation of the crowdfunding community as there is no data from 2021 – 2023, or from before 2010.
   3. We do not know any details about the companies behind the crowdfunding campaigns, therefore that is an unknown variable that could certainly affect the level of support a given campaign receives.
3. Additional tables/graphs that can be created for this dataset are:
   1. A histogram that traces the success rate of each different category over the datasets 10-year span. This would show how each category has either become more successful or less as time has progressed.
   2. A box and whisker plot that shows the distribution and outliers in the statistical analysis. Having specific outliers in the number of backers for successful vs failed campaigns would give a better representation of whether what unit of measurement best summarizes the data.