Jonathan (Riley) Sayre

Project 1

Data and Analytics Bootcamp CWRU

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + We can conclude that crowdfunding campaigns in the “Theater” Category are the most popular category of crowdfunding campaigns.
  + We can conclude that “Plays” are the most popular crowdfunding sub-category as they have nearly 250 more campaigns than they next highest sub-category, “Rock”.
  + We can also conclude that July is the best month to start a crowdfunding campaign as July has the highest amount of successful campaigns.
* What are some limitations of this dataset?
  + Heavily skewed towards “Theater” parent category and “Plays” subcategory.
  + Multiple currencies, not a single standard currency to truly evaluate the value of each kickstarter
  + Some projects had much higher goals than others which would skew the data, there should be a minimum goal to be involved as a campaign asking for $1,000 compared to one asking for $100,000 are vastly different.
  + Dates are vastly different. Some projects started in 2010 while others began in 2020. This is a big difference in the value of the currencies being used due to inflation. There should be a conversion of all currencies to USD (or another standard) and then converted to value of USD in today’s USD value.
  + Some projects are still live so that data needs to be thrown out as it cannot be used in a conclusive manner.
* What are some other possible tables and/or graphs that we could *create, and what additional value would they provide?*
  + *It would be helpful to know the goal of the kickstarter compared to the outcome. Do kickstarters with lower goals have a higher chance of being successful or do more expensive ones?*
  + *The country of origin to outcome would be good to know, perhaps certain countries have a higher, or lower, success percentage.*
  + *The number of backers to category would be good to know as well. Do certain categories pull more backers, even if those still fail due to cost?*