Samantha Hall

Module 1 Challenge – Crowdfunding Analysis

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + Performing arts, particularly theater plays are substantially more popular crowdfunding campaigns
    - 54% of theater crowdfunding campaigns are successful
    - 38% of theater crowdfunding campaigns fail
  + Crowdfunding campaigns with a goal of $50,000 or greater have the highest failure percentage at 49% of campaigns being unsuccessful
    - Crowdfunding campaigns with a goal between $15,000 to $35,000 have the highest success rate at an average of 95%
  + Time of the year does not appear to have a large effect on the outcome of crowdfunding campaigns
    - The data is mostly evenly distributed throughout the year,
      * However, in August we do a small spike in failed campaigns (31-35), followed buy a large dip from 35-23 in September for example.
* What are some limitations of this dataset?
  + We do not have indicators of why campaigns have been canceled, perhaps a campaigner had their project fall through and no longer needed funding or experienced some sort of delay in their timeline for campaign and decided to start a campaign closer to their new deadline date
  + The sub-categories are semi-descriptive
    - For example documentary is vague, was a failed documentary campaign due to nature of the documentary? What genre of theatrical play is looking for crowdfunding? What
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + A pivot table showing the number of successful or failed campaigns based on how long the campaign was active to further show possible reasons for outcome
  + A pivot table with a corresponding line or bar graph comparing the length of campaign with the outcome
* Use your data to determine whether the mean or the median better summarizes the data.
  + I believe the mean summarizes this data set better because it shows the average amount of backers and accommodates for outliers.
* 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 a larger variance with successful campaign.
    - It makes sense because there is a higher chance of variability with a larger data set, and the data set for successful campaigns is larger than the data set of failed campaigns.