Crowd Funding Answers

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

1. crowd funding is successful 56.5% of the time, while fails only 36.4% of the time, which makes it a fairly useful means of raising funds
2. plays are the most crowd funded of all the subcategories accounting for 34.4% of the entire study. The second most crowd funded sub-category is rock at only 8.5%
3. broken up by months, the most successful month to crowd fund would be June which shows a success rate of about 63.2%. While August has the lowest success rate of only about 48.2%

* What are some limitations of this dataset?

1. Just because a campaign was cancelled doesn’t mean it failed or was unsuccessful, the data is incomplete so a canceled project could mean the people/company searching for funding received funding elsewhere
2. the data set doesn’t show us who the biggest donators are. If we had that information we could use it to show what people should be targeted by which category. This could speed up the process and show clients were to advertise their crowdfunding efforts.

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

1. It would be useful to create pivot charts to show which categories had the largest “goals” and “pledges”. The value behind this could have multiple uses such as showing possible bodies like governments or big companies which categories have the most demand and require the most help.
2. Aside from creating and changing the colour of the column “percentage funded” it would be useful to create a pivot chart that showed which category and subcategory are most funded beyond their goal. This information would show what the most popular crowd funding options are and what people are most willing to contribute financial assets towards. This could persuade/discourage some future crowd funding ventures from taking place.

Statistical Analysis

* Use your data to determine whether the mean or the median better summarizes the data.

The Median is a better representation of the data. The Mean is substantially larger than the Median because there is a lot of variability in the data which makes the Mean less reliable.

* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Due to the data that has been collected, it shows me that there is more variability with the successful campaigns. The variance number and relative standard deviation are both much larger on the successful side which shows a greater distance from the Mean. This does make sense because there is a much larger amount of successful campaigns as opposed to failed campaigns so it is reasonable that the data would be more spread out amongst the successful campaigns.