Written analysis

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

Through analyzing the data one can conclude that the parent category with the highest number of participants was Theater, but journalism had the highest success rate of 100%. In the subcategory, plays had the highest number of successful campaigns but also has the greatest number of failed campaigns. Campaigns created in June and July were more likely to be successful than campaigns crated in august. Goals between 15,000 to 24,999 and 30,000 to 34,999 were 100 percent saucerful. The goal that had the lowest percent of success was 10,000 to 14,000.

What are some limitations of this dataset?

Without additional data or personal knowledge of these campaigns it is difficult to infer why some category’s performed better than others.

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

An additional graph one could create is duration of campaign vs success rate of the campaign. One could determine how the duration of a campaign effects its success rate.

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

I believe that the median would be the best method to represent this data as the variance is too high for mean to be a reliable method. Due to the high variance, outliers would skew the data.

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 than unsuccessful campaigns. This would make sense because successful campaigns bring in a wider range of backers which would lead to the higher variability.