**Crowdfunding Written Analysis**

1. Based on the data, the majority of the crowdfunding campaigns are done in the US. Overall, theater was the most popular category, and thus plays are the most popular subcategory. However, removing theater as a parent category from the analysis does not seem to significantly impact the rate of success and failure based on date created. Based on the sample, the campaigns were more likely to be successful (56.5%) than failed (36.4%) and canceled (5.7%). When removing theater as a category, the proportion of outcomes remain relatively the same: successful (58.7%), failed (36%), and canceled (5.2%).
2. The sample size is relatively small and may not give a comprehensive idea of campaigns that work and don't work. For some categories there is a very small sample size, and it is difficult to determine whether this is representative of the population or due to an inadequate sampling method, and thus draw conclusions on the data. There is no information about how the sample was drawn (randomized, systematic, biased, etc.). Additionally, there is not a large representation of countries and the US constitutes a lot of the sample, which could not be representative of the population.
3. One graph could be displaying the correlation between the length of time or a campaign and the outcome of the campaign. This would give us an idea of if there is any correlation between how long campaign is open and if that campaign is successful. Additionally, a graph could look at the number of backers compared to the outcome of the campaign. This could determine a correlation between the two and determine future strategies for campaigns. For example, if a campaign is more likely to be successful when there are more backers, then the campaigners can focus on gaining investors and promoting funding.

**Statistical Analysis**

1. The median better summarizes the data. This is because there is a large skew in the data towards bigger values, which is demonstrated statistically by the large variance and standard deviation, as well as visually by the box and whisker plots. This is the same for both successful and failed campaigns.
2. There was more variability for successful campaigns, as both the variance and standard deviation are greater for successful campaigns than failed ones. This makes sense based on the data itself, as the max is higher and minimum lower for the successful data, indicating more range. Additionally, this can intuitively make sense because the failed campaign has a lower median, mean, max, and minimum, indicating that overall, the failed campaigns have consistently less backers, perhaps providing some correlation with why they failed, and thus creating less variance in the data for the failed campaigns.