Module 1 Challenge: Data Report

First set of questions:

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + Out of the 986 campaigns 565 were successful. Therefore, we can conclude crowdfunding campaigns are successful approximately 57% of the time, based on this data set.
  + Theatre had the most demand for crowdfunding over the time in which this data was collected.
  + Based on the data set provided, the failure rate for crowdfunding campaigns was approximately 37%. The failure rate was closely distributed across all categories, with the highest failure rate being 48% and the lowest being 26%.
* What are some limitations to this data set?
  + The data set is limited because it does not include any data on how the campaigns were marketed. Information such as, dollars spent on marketing, platform in which the campaign was marketed on, and other marketing information would provide additional insight into why these campaigns succeeded, failed, or were cancelled.
  + It is unclear what the data in the blurb represents, which limits the person analyzing the data set.
  + The data is limited because it does not offer any perspective into the people making the donations. Additional data such as, census data on the people making donations, or data on whether the backers knew those who created the campaigns, would allow for more accurate conclusions.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide.
  + We could have created a bar graph showing the percentage failed and succeeded for each crowdfunding campaign category. This would have provided a closer look into how each category performed.
  + We could have created a chart showing the relationship between the amount raised and the total length of time a campaign ran. This would allow us to analyze the relationship between time and dollars raised and determine if this is a factor in the amount raised.
  + We could have created a chart or graph that better defines the relationship between the size of the average donation and the duration of the campaign and/or the goal of the campaign. This relationship would allow us to highlight another potential key factor into why campaigns fail or succeed.

Second set of questions:

* Use your data to determine whether the mean or the median better summarizes the data.
  + For this data set the median is a better summary of the data as there are significant outliers past the second standard deviation.
* 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 the successful campaigns than there is with the failed campaigns. This is determined by the higher standard deviation and variance in the successful campaigns. This does makes because when a campaign is successful because there is no cap to how much successful campaign earns. Whereas a failed campaign can only be 0 to the goal, decreasing the potential for total variability from one campaign to another.