Create a report in Microsoft Word, and answer the following questions:

* + Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
    1. In parent categories theater, film & video, and music had the greatest number of successful funding.
    2. In subcategories, play had the highest number of attempts at nearly 350 while all other categories had under 100. Plays had 187 successful outcomes while all others had under 50.
    3. June and July are the best months for successful crowdfunding. The rest of the months are generally similar.
  + What are some limitations of this dataset?
    1. It is unclear what crowdfunding platform the campaign is from. Per background instructions they give examples like Kickstarter and Indiegogo. Does the data only include data for platforms like Kickstarter and Indiegogo. If so, which one is more successful?
    2. The blurbs provide information on the “reason” for crowdfunding however the information is not sortable or categorized. It is possible that certain reasons for funding are more successful than others. However we cannot extract the correlation of reason for funding and success rate.
  + What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
    1. A table and graph with a new calculation of average amount donated per day for each category. This would help estimate how long you need to crowd fund for to reach your specific goal.
    2. A table and graph summarizing outcomes per country. This provides value on determining which countries are most successful or unsuccessful in crowdfunding.
    3. A table and graph summarizing outcomes per year to show a trend over time.

STATISTICAL ANALYSIS

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

Our data is not symmetrically distributed and has several outliers in the upper quartile so our data is skewed. The mean is inflated due to the high outliers. Therefore, the median is the realistic and typical data point to summarize the successful and unsuccessful crowdfunding.

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

The successful campaigns variance (1,603,373.7) is higher than the failed campaign variance (921,574.7). The higher the variance, the greater the variability. Therefore, there is more variability with successful campaigns. As an additional checkpoint, higher standard deviation means greater variability. The successful standard deviation (1,267.4) is also higher than the failed standard deviation (961.3) This makes sense as the range for MIN and MAX is also larger for successful campaigns (MIN: 16, MAX: 7,295, Delta:7,279) in comparison to failed campaigns (MIN:0, MAX:6,080, Delta:6,080).