Module 1

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

We can conclude that the number of successful campaigns were much higher than those who failed and or cancelled. The peaks of each dataset were at their highest from May to August. The number of failed and cancelled campaigns increased towards the end of the year.

This dataset is too broad. It should be broken down into smaller categories to have a more finite answer towards the different campaigns being started. Instead of country wide it could be done by province/state so we could determine which area of each country is more inclined to fail vs succeed.

The dataset could be represented in a bar/column graph. The information provided in the dataset is easier to compare and measure the distinct categories of data. The dataset would be more visually pleasing to view as well as offer a more definitive way to show the direct comparisons.

Statistical Analysis

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

The dataset provided would be better represented by using the median. The median is more suited for skewed outcomes to determine the pattern of the information provided. It is a more ideal way of viewing the data as it is a more realistic view to gauge the information due to the variances in the dataset.

There is a slight edge in variability towards successful. The reason being that there are much more successful campaigns to compare versus the failed campaigns. This data makes sense since both receive different levels of backing. Some in single digit and some in the high multi digit. Therefore, there is no definitive differences in the variances due to the mismatched quantity in the dataset. Being that the dataset is too broad of a spectrum it should be categorized into more finite categories before an actual comparison can be performed.