**Module 1 Challenge**

**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?

Based on the given data, it can be observed that there are multiple categories listed, which primarily belong to the entertainment and technology industries. Upon analyzing the pivot tables and graphs, it becomes apparent that there are three businesses that stand out from the rest as there is a significant differentiation in performance. Those three businesses are the Theater, Music, and Film & Video industries.

Even among these three businesses, most of the people have chosen to do theater business however, looking at the data theater business has lot of failures compare to other two businesses.

* What are some limitations of this dataset?

After reviewing data, these are the following limitations of the report which could simplify data more promptly.

* All category of businesses’ percentage of failed, live, successful and canceled outcomes.

Eg:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Category | Live | Canceled | Successful | Failed | Total |
| Theater | 0.58% | 6.68% | 54.36% | 38.37 | 100% |

* Add another column to convert different currency into USD to get precise data.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
* There should be pivot table to filter not just yearly, but quarterly data. Also create column chart for better visual.

**Statistical Analysis**

* Use your data to determine whether the mean or the median better summarizes the data.
  + Upon analyzing data and examining the chart, there are few numbers of successful campaigns that has lot more support than the rest. As a result, the average number of backers appears higher. And due to that reason, the large number of backers significantly biases the mean to the one side. Therefore, the median provides a more meaningful summary of the data, not mean.
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
* The data sets exhibit higher variability. due the higher variability, the concept of mean meaningless, which doesn’t make sense for both successful and unsuccessful.