After reviewing the Crowdfunding Data Set in Excel, I was able to modify and analyze this sample using conditional formatting, formulas, and data visualization. Three conclusions that we can draw from the data on crowdfunding campaigns include:

- Plays have the highest success rate for all countries, however other types of successful campaigns vary by country. I determined this by reviewing the data and creating a Stacked Column chart. This type of chart uses aggregate (groups) data and lets you visualize each aggregate and its raw counts.
- Successful campaigns started increasing in May, June, and July, but then started falling again in August. I was able to determine this by using the Line chart. Line charts are typically used in time series to show change over time.
- Journalism appears to have a 100% success rate. Although this is true, the data only includes 4 samples and so you could not conclude that Journalism is always successful.

All data sets are going to have their own limitations. These limitations can include sample size, missing data, outdated data, and biased data. A couple of limitations that I observed from Crowdfunding data set are outdated data and sample size. This data was collected between 2010 and 2020, so the conclusions that we draw may not be accurate 4-14 years later. Also, the sample size is 1000, which is not a large data set considering that the data consists of 7 countries. The larger the data sample, the more confident we can be in our conclusions.

In this assignment, I used the Pivot Chart Stacked Column chart and Pivot Chart Line Chart. These two types are ideal for the type of data that I was working with. You could also use these charts to look at other variables within the data set by breaking outcomes down by country or determining if the time elapse between the launch and deadline had any influence on outcome.

Mean/Median/Statistics

The final part of Challenge one was comparing the number of backers for both successful and failed campaigns. I figured the Mean, Median, Minimum, Maximum, Variance, and Standard Variation for each group. The mean (or average) is very sensitive to outliers. The Mean for Successful and Failed groups is much greater than the median. When the Mean is greater than the Median, you can conclude that the high value outliers are pulling the Mean up. The Median best summarizes the successful and failed backers count because the distribution of the data is skewed right.

By looking at the calculated Variance and Standard Deviation of both successful and failed campaigns, I can determine which group has more variability. The Standard Deviation for successful campaigns is 1266 and failed is 960. Since the Standard Deviation is higher for successful campaigns, I can confidently state that the successful campaigns have more variability in the data set.

References – The information and resources that I used for this assignment include:

- Zoom classes on Excel from 4/16/24 and 4/18/24
- Google Search
- Xpert Learning Assistant
- Crowdfunding Data Set
- Excel