**Project #1: Excel Challenge**

**Written Report**

**1.Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**

-For the parent category, Journalism; All the crowdfundings campaigns so far have been successful

-The subcategory, plays radio & podcasts has held the most amount of crowdfunding campaigns compared to other sub categories.

-USA has held significantly more crowdfunding campaigns compared to Canada.

**2. What are some limitations of this dataset?**

- One of the limitations of this dataset is that there is no data available to find out why some categories had more successful crowdfunding compared to other categories. If data regarding how these crowdfunding were advertised within people was available, then it would have given us more insight.

- Another limitation is not knowing what time of the year were these crowdfundings hosted, the location of these crowdfundings i.e. Remote, in person, hybrid. This data would have further helped determine if outside factors had a role in the success or failure of these crowdfundings.

**3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

- To elaborate on my previous answer, I would say values regarding what kind of advertisement (by memos, email alerts, social media, etc.) was done would be very useful We could compare and see if one type of advertisements was more effective compared to another.

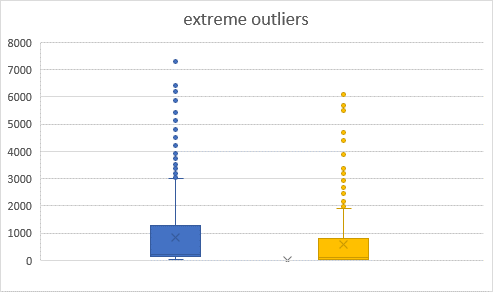
- Another table would be of the location of these crowdfundings as mentioned. Season of these crowdfundings; i.e. Fall, Spring, Summer, Winter.

**Statistical Analysis**

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

- The mean value is greater for both Successful and failed outcomes which suggests that the data is positively skewed by extreme outliers. Hence, median is less affected by these extreme values and would be a better option to summarize the data set.

- Below is the Box and Whisker chart for the Statistical Analysis data set which shows the extreme outliers present in the dataset:



**5. 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 higher variability with successful campaigns as the value of calculated variance and standard deviation is higher for successful campaigns compared to failed campaigns.