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

1. The three major campaign categories appear to be ‘Film & Video’, ‘Music’ and ‘Theater’.
2. When digging a little deeper, the subcategory ‘Plays’ is significantly more prominent than any other subcategory. When looking closer at the parent categories, the ‘Film & Video’ the subcategory ‘Documentary’ is most prominent, and in ‘Music’, the subcategory ‘Rock’ is most prominent
3. When looking at the LaunchDateOutcomes, when all years and categories are combined June has a spike in successful outcomes. When filtering further, it appears this spike comes from the ‘Theater’ category. Music, has a spike of successful campaigns in July.
4. There are more successful campaigns than there are failed, even when combining cancelled with failed and live with successful.

**\* What are some limitations of this dataset?**

1. This is only a sample of data of campaigns. There could be an error on how this data was gathered or a bias in the sample based on what kind of projects the collector is interested in.
2. Although there are two dates, ‘Date Created’ and ‘Date Ended’, it would be interesting and valuable to see how the backer are accumulated throughout their campaigns.
3. There are a limited amount of data from different countries. Our data only accounts for 7 countries. Furthermore, it would be interesting to see if there are difference’s in states if we want to drill further down.
4. It would be helpful to know how the different campaigns got their backers. Example, networking or marketing? If marketing then how did they market their campaigns and how much did they invest into marketing?
5. How much does each

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

1. A pivot table which shows the relationship between the average donation and the success rate of the project. Make filters with Categories and Subcategories to see which categories tend to have higher donations. It can reveal a trend that success of the campaign is not dependent on the number of backers but the amount of funds it receives.
2. Two scatter plots for each successful and failed displaying how many days a campaign was in session. (From date created to date ended) This can reveal a trend if some campaigns had more time to get backers, there could be more successful outcomes.
3. Two box and whisker plots for each successful and failed campaigns which includes backer count. This can be a great visualization to see if the data is skewed and if there are many outliers.

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

Median is more accurate and summarizes this data set better because the data is skewed and contains outliers. I came to that conclusion due to the significant difference in mean and median. I also created a box and whisker plot in the other excel file but had to delete it due to slowing down excel too much.

**\* 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 more variability with successful campaigns. Some successful campaigns have a larger number of backers but don’t necessarily need to exceed their goal in order to be successful whereas unsuccessful campaigns seem to fail because they did not get enough backers and did not achieve their goals. When looking at the box and whisker plots there appears to be a larger number of outliers in successful campaigns compared to unsuccessful which explains the variability which also gives more visibility.