# Crowdfunding Campaign Write-Up

#### Written Report (20 points)

* Presents a cohesive written analysis that:
  + Draws three conclusions from the data (10 points)
  + States limitations of the dataset and suggestions for additional tables of graph (10 points)
  + Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

The first pivot table labeled, ‘**Crowdfunding Pivot Table**’ highlights the three primary category campaigns that generate the most successful, failed, live, and canceled results for a sample list of countries worldwide (i.e. Australia, Canada, Denmark, Great Britain, Italy, Switzerland, and the U.S.) It is clearly evident that the theatre category generates the most results, while the music and the film & video categories are almost similar in the number of outcomes generated for the list of countries worldwide.

It is clearly evident that close to 76% of the results are generated from one country (i.e. the **U.S.**), while **Great Britain** and **Italy** are almost close to each other when generating the 2nd and 3rd most results. More importantly, 80% of the world’s successful campaigns are generated in the **U.S.**

Only in **Great Britain**, the outcomes of the campaigns are more defined and realized for the music and film & video categories; and the theatre crowdfunding campaign is the lowest from this data set. (But it is only slightly behind).

Lastly, the theatre campaign being the most successful, could be skewed by the financial profile of people living in the **U.S.** Essentially, what is the type of consumer that supports these campaigns, and the level of disposable income that is available to support these crowdfunding events is another aspect to consider as well.

From the 2nd pivot table labeled, ‘**Crowdfunding Pivot Table (Sub)**’ It is clearly evident and reinforces the facts from the first table; that the subcategory ‘**plays**’ under the primary category ‘**theatre**’ has the most successful and failed outcomes for a given campaign.

The 2nd most important campaign comes from the sub-category ‘**rock**’ from the music category. These generate the most results for every country besides Australia, which prioritizes the sub-categories of ‘**drama**’ and ‘**web**’ based campaigns to be successful. As also seen by the fact that, ‘**documentaries**’ generate the 3rd most successful campaigns for the film & video category for every country in this sample, besides Australia.

From the 3rd pivot table labeled, ‘**Crowdfunding Pivot Time Scale**’. The most successful campaigns for the theatre category are realized during the months of June and September on average for the years 2010-2020. However, the most failed campaigns happen to occur during the months of January, May, and October for the years 2010-2020 on average.

From the 3rd pivot table labeled, ‘**Crowdfunding Pivot Time Scale**’. The most successful campaigns for the film & video category are realized during the months of February and August on average for the years 2010-2020. However, the most failed campaigns happen to occur during these months of December and January for the years 2010-2020 on average.

From the 3rd pivot table labeled, ‘**Crowdfunding Pivot Time Scale**’. The most successful campaigns for the music category are realized during the months of February, July, and August on average for the years 2010-2020. However, the most failed campaigns happen to occur during these months January-April for the years 2010-2020 on average.

* + What are some limitations of this dataset?

The dataset isn’t able to identify and quantify the demographic of people who are able to influence the different outcomes for these campaigns for a sample list of countries. It also does not take into account the age of the population sample from various countries, and it cannot be used to determine what is the spending capability of these individuals depending on how much money they are able to retain and support various crowdfunding campaigns.

The limitations of this report can also be tied to the fact that there is no explanation, as to what the columns ‘**staff\_pick’** and ‘**spotlight**’ should indicate.

Essentially if these columns are appropriately defined, then it would possible to draw better conclusions as to what columns can be correlated to each other in order to generate meaningful insights in the report.

From my perspective ‘**staff\_pick’** could indicate whether the campaign was unanimously picked by members within the organization or the campaigns were consulted and agreed upon by external parties.

From my perspective **‘spotlight’** could indicate that these campaigns were marketed to reach a broader audience, by using a specific platform or service. However, in this case; it would be unknown to understand how much might have been spent to create an effective campaign that resulted in 4 possible outcomes.

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

From the dataset provided, we can investigate which campaigns would rank in the top 10 for the most successful, failed, and canceled campaign outcomes. More importantly, it would be ideal to investigate where the most popular outcomes are generated to investigate the data further.

Also, comparing the number of successful and failed campaigns from when the campaign started and when the campaign ended, as a time series graph; could be used to identify what is the most effective period for campaigns to be successful. It also gives an idea as to how long the consumer is willing to wait before proceeding with a certain campaign.

Another useful graph might be to explore the use of a Pareto Chart, along with campaign titles and the countries where most results are generated from successful, failed, and canceled campaigns. In this way, the focus can be centered around concentrating on where to focus the crowdfunding campaign search.

Also, having a cluster chart to indicate the most popular titles relative to what was actually pledged, can help to identify the word titles that are popular in some countries, and how can marketing be used effectively to promote more campaigns.

#### Statistical Analysis (20 points)

* Computed calculations of the mean, median, min, max, variance, and stdev using Excel formulas (15 points)
* A brief and compelling justification of whether the mean or median better summarizes the data (5 points)
* 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 **median** is a better measure of the data set; as there are many outliers present for both the number of successful and failed outcomes in the data set. Since there is more variability, the median is a better statistical measure, as it ignores more of the outliers that are present. This is also evident by the fact that the data is not symmetrical from the box plot analysis, since a majority of the outcomes are generated from the US, and the amount of revenue generated from the various campaigns from 1 country does not compare equally to the rest of the sample population in the U.S.

Also, the histogram clearly indicates that the data is positively skewed, where the mean is considerably higher for both the number of successful and failed outcomes in the data. Also, since some campaigns generate considerably more funding than others, the mean is not a good measure of the data, and it would be beneficial to proceed with utilizing the median as a better measuring stick to tabulate a set of analyses for the number of successful and failed outcomes.

When it comes to assessing the variability of both successful and unsuccessful outcomes, the number of **successful campaigns** has a higher variance than the number of failed campaigns. This does make some sense, since the total # of observed successful outcomes is much higher than the number of failed campaign outcomes. In other words, there were more backers supporting successful campaigns with the number being 7295, while the minimum number of backers was 16 for successful campaigns. Unsuccessful campaigns, however, had a maximum number of backers of 6080, while the minimum number of backers was virtually non-existent with 0. This also helps to provide more evidence as to why the data is more skewed for successful campaign outcomes vs. failed campaign outcomes.