**Written Report – excel-challenge**

**Conclusion one:** *(based on Outcome\_per\_category tab in workbook)*

**Between 2010 to 2020, crowdfunding campaigns have had the most popularity in theatre. More specifically, this category has had the most project proposals within the US.**

Given the provided data set, theatre has the highest number of crowdfunding campaigns across all countries. This category has a total of 344 campaigns associated with it out of the 1000 in total, meaning it accounts for 34.4% of the total campaigns. When drilling down into the US alone, the theatre campaigns for this country accounts for 79% of the total campaigns in this category.

Rationally, as theatre has the highest popularity, this may also explain why it had the highest count of: ‘cancelled’ (17), ‘failed’ (106), and ‘successful’ (149) outcomes for the US.

If you were to set up a crowdfunding campaign in theatre, you would be more likely to outcome in a successful campaign than have a failed, or cancelled outcome. However, it should also be considered why this category has the highest number of campaigns? It may suggest that independent content creators struggle to find funding themselves through other mediums such as individual savings and there are higher prospects for this industry if within the US. Alternatively, it could suggest that it is so popular due to the number of famous celebrities linked to the industry within this country, who have generated its popularity to try out projects; this is backed up by the second and third highest categories being linked to film & video, and music but further correlating analysis would be required to verify this trend.

More interestingly, the publishing category only had a total of 67 campaigns across all countries. Out of these 67 campaigns, 49 of these are accounted for within the US region whereby 28 were successful, 18 had failed, and the remaining 3 were either still live or were cancelled. When acknowledging the successful and failed outcomes alone, this means 64% of outcomes were successful in this category and region, warranting the highest success rate in proportion to failed when comparing each category.

**Limitations:**

The data set observed is biased towards the US and it’s cultural influences, the majority of the population sample resides in this region. It should be considered that the other countries patterns and trends would be better analysed by obtaining a data set with more information from these countries before reaching conclusions about them.

With that being said, even the sample size of the 273 theatre related crowdfunding campaigns analysed above, is also not a sufficient representation to reach conclusions with confidence. The US is an extremely large country, and it is very likely that there are many more crowdfunding campaigns, a bigger data set would help to better represent an entire category population.

The stacked bar graph may mislead the end user slightly; just because the highest number of successful campaigns were linked to the theatre industry, does not mean that you should necessarily begin a campaign in this category as shown when looking at the publishing category. Some additional analysis would be required firstly to try and find out more information about the other mediums creators raise funds for their projects on, and you would also need to check if there were any significant events such as celebrity announcements during the time period observed which may have impacted the results.

Using the current data set, I would make the future recommendation of creating a scatter plot. The chart would be filtered for successful outcomes of theatre campaigns in the US and would compare goal amount’s to the percentage funded. Conducting various regressions on a graph like this would help visualise the relationship between these two variables and perhaps help to determine what goal amounts may result in a higher percentage to be funded and therefore, more likely to be successful if someone were to start a campaign in this category.

**Conclusion two:** *(based on Outcome\_per\_sub\_category tab in workbook)*

**The plays sub-category outcomes are directly proportional to its parent category: theatre; there is more diversification of sub-categories in film & video.**

Plays has the most voluminous count of outcomes in all countries. This is expected as this sub-category falls within the parent category of theatre which has already been established as having the highest count of outcomes, in fact, plays is the only subcategory within theatre and so they are directly proportional.

The 273 play related crowdfunding campaigns in the United States alone accounts for 27.3% of all campaigns across all countries and 36% of all campaigns in the US. Out of these 273 projects: 149 were successful, 106 failed, and 17 were cancelled. If we omit the 17 cancelled and focus on the ratio of successful to failed, this amounts to 61:39 respectively. It can be inferred that if an individual sets up a crowdfunding campaign for plays in the US, they will be weighted towards being successful.

When observing the second most popular crowdfunding campaign category from our first analysis: film & video, in the US it was found to consist of 6 sub-categories: animation, documentary, drama, science fiction, shorts, and television. When comparing success: failure rates, the sub-categories: television and animation have a 75% and 71% chance of succeeding respectively, whereas the subcategory science fiction had a 63% chance of failing. Therefore, raising a crowdfunding science fiction film/video campaign would be weighted more towards failing.

**Limitations:**

There is further room to drill down ‘plays’, this subcategory consists of many genres; one of which may be more likely to gain backers in comparison to another. Therefore, the data analysed in this section has the potential to provide false security for an individual looking to set up a campaign in this area. Although time consuming, sentiment analysis on the ‘blurb’ column in the initial data set could help find trends. For example, it could help group commonly appearing phrases associated with a particular genre, which can then help further categorise this sub-category and then enable us to analyse different play genres against outcome counts for further interrogation to help users better understand what drives it’s success.

As a running theme, the more data the better, it would also be beneficial to gain more data for each of the sub-categories in film & video to help provide more assurance to the conclusions raised, it would also help to find other sub-categories within theatre to see if they also are as successful as plays e.g., opera’s.

**Conclusion three:** *(based on Outcome\_per\_creation\_date tab in workbook)*

**Proceeding to create a crowdfunding campaign in the theatre category within the month: May, should be done so with caution.**

When analysing the outcomes of crowdfunding theatre campaigns with their creation date, June was the month which had the highest number of successful campaigns across all years and accounts for 11% of the total successful campaigns in theatre, it had the most success in 2010, 2013, and 2014.

It can also be deduced that within the theatre category, the lowest number of successful campaigns occurred in May and August across all years. May also happened to be the month that had the most failed campaigns occur for this category. Therefore, whilst it may not be secure to say that a campaign should be created in June as its percentage of the total doesn’t account for so much, it is probably better to suggest to avoid creating a campaign for theatre in May due to it consistently having at least one failed campaign per year (with the exception of 2015).

**Limitations:**

This data dates up to a creation date of January 2020 which is now over three years old. A lot of circumstances may have changed since; developments may have been made in the industry which could have influenced our findings. The exclusion of more recent data also ignores the impact that external influences such as the pandemic’s impacts on people’s views of raising crowdfunding campaigns may have had. This data also has some lines which do not join when comparing outcome by creation dates, this would not likely happen in a bigger sample size.

I would also conduct further analysis for the month of May in future to compare the duration that a campaign had between its creation and deadline (in days) along with the count of outcome status on a column chart to check if there is a relationship between duration and success, it could help understand if duration is an influential factor which should be considered.