Excel Challenge-Crowd Funding

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

* Even though “film and video, “music” and “theatre” have the most successful outcomes with 102, 99 and 187 respectively. I can conclude that the most successful parent category was “technology” since they had 66% of their outcomes to be successful compared to 56%, 57% and 54% for “film and video, “music” and “theatre” respectively. The Kick-starters and Indiegogo should focus more on “technology” campaigns and find out ways to improve success rate for film and video, “music” and “theatre”.
* Similarly, even though the sub-category “plays” has the most successful outcomes at 187, it also has the most failed outcomes at 132. Based on the data, I can conclude that the most successful sub-categories are “web”, “wearables”, “translation” and television with a success rate of 71%, 62%, 66% and 64% respectively.
* Based on outcomes per year pivot and chart, I can see a clear trend in data where we can see peak in successful outcomes in June/July and a peak in cancelled and failed outcomes in August/September. This could be a result of the projects aiming to be completed June before end of financial which resulted in more unsuccessful campaigns in end of July and August.

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

* No data on the reason the project is cancelled or failed. If this was possible, we can create a pivot/chart based on reason for failed/cancelled projects per category or year to see the main causes.
* Unable to confirm the actual average donation per backers/currency since currency conversion rates are not utilised.
* Timestamp for time taken to complete the project not mentioned in the dataset. This can identify how much time taken to complete each project based on category.

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

* Pivot/chart showing count number of backers per Parent category and sub-category to see which category backers are contributing the most for. This will show the popularity of different type of projects.
* Pivot/chart showing total donations/pledged amount per currency and country. This will show which countries are generating the most donations/pledges for each category and what’s missing in other countries. In addition to this, another chart can show number of categories that are completed in each country and percentage of completion projects in each country.
* We can also do a chart on pledged amount and categories to see which category requires the most and least funding. So, Kick-starters are aware what type of projects to get into.

**Bonus Statistical Analysis**

1. **Use your data to determine whether the mean or the median better summarises the data.**

* As the number backers is from a range of 0- 7000, I feel the median better summaries the data since it shows that most of the data sets are around 200 for successful outcomes and 115 for unsuccessful outcomes. The mean doesn’t present the data as well since it’s outliers will drastically impact this measure.

1. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

* Based on the data set there Is more variability in successful campaigns compared to unsuccessful campaigns.
* The variance is 1603377.73 in successful which was drastically higher than 921574.68 in unsuccessful campaigns. This shows that the values in the successful campaigns are much further from the mean compared to the values in unsuccessful data set. The standard deviation also supports this, since the standard deviation in successful campaigns (1266.24) is higher than the standard deviation in unsuccessful campaigns (959.99). This also proves that the values in the successful campaign data set are far from the mean in comparison with the unsuccessful data set.