Excel-Challenge: Module 1

1. **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**
2. The Theatre Parent Category had the most amount of campaigns. Parent Categories such as technology and film & video have a higher success to failure ratio. Therefore the amount of campaigns does not always reflect a higher success rate.
3. The Plays Sub Category makes up 34% of the total campaigns, which is the most amount out of all the sub categories. Even when filtering per country, this is true. This could indicate that producing plays often incurs higher expenses that theatre companies might not be able to fund, which is why they require crowdfunding.
4. It appears that there is no correlation between crowdfunding at different times of the year, however the data displays a potential spike in successful campaigns around the mid-year July mark. This remains true across all years of crowdfunding displayed.
5. **What are some limitations of this dataset?**
6. One of the limitations of this dataset is that the amount pledged per backer is not recorded but is instead the sum of the total amount of backers. There is also no data to display who these backers are, and who the campaigners are reaching out to for funding. This means that there is no information to help make informed decisions around marketing to potential backers.
7. Another limitation could be that despite a few genres of music and television being present, the data is not specific enough to differentiate between successful and failed campaigns. If the goal is to analyse the data for trends in what may result in a successful campaign, there is not enough granularity for people launching media-based campaigns.
8. Another limitation includes the small amount of countries in which these campaigns were launched. The data provided only covers a total of 7 countries and will not be able to inform potential successful campaigns in other countries.
9. **What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**
10. Using the data available in the Date Created Conversion and Date Ended Conversion columns, a conclusion could be reached based on the duration of a campaign and the ratio of success to failure. It would inform the length of the campaign.
11. Using a pivot table to find the average of the pledged amount per backer for each Parent Category, we can then create a bar graph with the provided data to help predict future pledges based on the campaign’s Parent Category.
12. A Pie Chart can be used to determine the amount of campaigns held in each country.

A pie chart with different colored circles

Description automatically generated

Bonus Statistical Analysis

**Use your data to determine whether the mean or the median better summarises the data.**  
Both values are required in order to create an informed summary of the dataset. If only the mean is considered, the most common amount of backers may be interpreted as being higher than what it actually is. This can be seen in the difference between the mean and the median, especially considering the minimum and maximum amount of backers for each outcome.

**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 as shown by both the standard deviation and variance in comparison to the failed campaigns. This makes sense when comparing the minimum to the maximum number of backers which shows that both campaigns with a smaller amount of backers have been successful, as well as those with a larger amount. Additionally it’s easier for smaller campaigns to succeed due to smaller goals which would influence the variance, and there is also a likelihood that there are outliers within the successful outcomes dataset.