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

1. Data shows that under parent categories, all of the parent categories have successful outcomes and **theater** is the one with the most successful outcome. This result doesn’t mean campaigns will succeed more if they use this data as reference because the failed outcome is also high across the board except for **journalism**.
2. Moving to sub-category vs. outcome, the sub-category **play** showed the most successful outcome and drastically outperform the rest of the category. The sub-category that had the lease outcome are **world music** and **audio**.
3. Date shows that the highest successful outcome of the year is in July and the lowest successful outcome is in August; the most failed outcome is in August and the least failed outcome is in September. There is no direct correlation between success and fail across the year but there are some data shows that the more successful outcome tends to correlate to higher chance of failure to occur. The canceled outcome data seemed very constant, and it showed that the canalization of the project is not correlated with successful outcomes or failed outcomes.

What are some limitations of this dataset?

* The data didn’t include how those campaigns reached the target audiences for engagement will which also affect the outcome of the project.
* The data didn’t emphasize the importance of the year which will also affect the funding.
* Need to show the amount of funding and successful and failed rates for outcomes.

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

* Create a pivot table that can be filtered by country with campaign launched year and outcomes as row and value. This will analyze the relationship between launch year and the outcome across different countries.
* Create a pivot table that can be filtered by country with campaign launched year and pledged as row and value. This will analyze the relationship between launch year and the pledged amount across different countries.
* Create a pivot table that can be filtered by year and country with pledged and outcomes as row and value. This will analyze the relationship between pledged and outcome across different years and countries.
* Create a pivot table that can be filtered by country with parent category and pledged as row and value. This will show the trend that indicate which parent categories might get more fund based on the parent categories across different countries.
* Create a new sheet and out pledged amount into categories such as “less than 1000”, “100 to 4999”, etc. to see the relationship between how pledged amount affect the outcomes.

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

* If you create a box and whisker plot using the date for both successful and failed outcome, the chart indicates both have lots of outliers which means that using mean is not a good way to summarize the data. The mean value in this case is affected by extreme. However, median is a better choice as it will not be affected by outliers.

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

* Variance in successful backer count is only a bit greater than backer count with failed outcome. This means that the variability of the successful and unsuccessful campaigns is similar, the spread of two means is very close and not significant.