1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   - There are a high number of campaigns in the parent category of film & video, music, and theater. This shows that many projects are being created in these categories and could be due to demand or popularity which might need more data to support.
   - There is only one sub-category named play under category of theater, and this category has the highest number of campaigns with successful rate of 54%. Under film & video, documentary has the highest total number of campaigns with a successful rate of 57% which is above average successful rate within all the sub-categories under film & video. Under music, the rock music sub-category has the highest number of campaigns with a successful rate of 58% which is below the average of the successful rate within all the sub-categories under music. These shows which sub-category under the parent categories have more projects and this might be a sign that those sub-categories are getting more attention and have more demand than the others.
   - The number of successful campaigns in each month is approximately in the range of 40-50, however, there seems to be a slight peak in successful outcomes during the months of May to July, then the number of outcomes returns to the range of 40-50 in August. This could mean there are more campaigns/ projects being created/initiated during May to July or people are seeking new work in each field during those few months.

2. What are come limitations of this dataset?

   The possible limitation for this dataset would be that there are not enough data since the range of the data covers around nine years. I would suggest a bigger number of data set that covers a wider range of years, also the years which the data is collected could also affect the outcome due to the different age group.

3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   - Could create a pivot table and graphs regarding the percent funded for each parent category. By creating this table/graph could help us see which category/sub-category is getting more recognition by other backers as they are putting in more money to invest compared to the initial funding the category is looking for.
   - Could create table/graph of a percentage of successful, failed, canceled, and live for each parent/sub-category to compare data. This is because just knowing the number of successful campaigns does not really tell much since the chances of it being

successful depends on the total number of campaigns that’s being counted. Knowing the percentage of successful and failed could help determine which category/sub-category has a higher chance of being invested or not being invested. This data could help during analysis as it shows the likeliness of the category/sub-category succeeding.

Statistical Analysis

[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?]

- Based on the variance and standard deviation that are calculated from the successful and failed data, it’s shown that it is better to summarize the data with median instead of mean. This is because for both successful and failed data, the variance and standard deviation are both very large and this indicates that the data set is very spread out from the mean, which will not be very accurate to use for representation of the data set.
- Based on my data, there is more variability with successful campaigns, and this make sense. This is because with failed campaigns, there is a way lower average of backers count in most category since the campaign itself is less supported by backers. Whereas for successful campaigns there are a bigger difference in each categories, such as, there could be a huge number of backers count, 181835, for successful theater campaign, and there could also be a relatively lower backers count, 12761, for food campaigns.