**Conclusions:**

There are multiple different metrics used in this dataset and we can draw a few different conclusions:

1. **Years/Months**: As we have the data divided into different years and months looking at the number of successful campaigns by different months, we can see that the highest number of successful campaigns happen in the summer months, especially in June and July. While the lowest number of failed campaigns happen in September.
2. **Categories**: When created a pivot table for categories, we can conclude that the most successful campaigns are from Theater followed closely by Film & Video and Music. If we look a little deeper into subcategories, Plays seem to be the sub-category with the most campaigns and highest number of successful campaigns. However, these categories also have the highest number of campaigns. While the Technology and Photography category has a high number of successful campaigns by ratio as well as the smallest number of failed campaigns. Journalism also deserves special mention as it is the only category with a 100% successful campaign rate. But also has the lowest number of total campaigns.
3. **Goal $ Value**: When looking at the dataset based on the goal amount, we can conclude firstly that the highest number of campaigns are between 1000 to 9999 or above 50000, and between these 3 the 1000-4999 range is the one with the highest success percentage. Even though the range of 15000 to 24999 has the lowest number of total campaigns they both have a 100 % success rate along with range 30000-34999. The entire range between 15000 to 34999 has a very high cumulative success rate along with having low failure and cancellation percentages. The only ranges for which the success percentage rate is less than failure rate is 10000-14999 and for the range above 50000.

**Limitations**:

The main limitation of this dataset is based solely on quantitative data and only tells us that certain campaign failed or succeeded based on different dates, goal amount and categories, however it does not tell us the details of why it has failed or succeed and the qualitative side of the data.

Another limitation of this dataset is according to the goal amount; we have different values and data for campaigns with funding for less than 50000. But if we look at campaigns above this goal amount, we only have one category. This should be broken down further to get more insight into campaigns above this amount.

**Additional Tables/Graphs:**

We could look at this data based on the different countries and geographical locations to determine how that affects the success and failure of a particular campaign. Also, by making a graph based on that table we could get visual representation based on geography.

We have a column on the main table for average donations, adding this information to all our tables/graphs of the goal amounts and combining it with the backers count column we can draw more detailed conclusions.

Another graph that would add more value to the dataset is to investigate the variance. We could use a box and whisker chart to get a visualization on the variance of the data.

*The median would better summarize the data as there are a few campaigns with a high number of backers that would affect the mean/average. The median, however, shows us that most of the campaigns have a backer count in the low hundreds. The value of the mean is 3-4 times higher than that of the median. If these 2 values were closer, it would suggest an even distribution.*

*The variance is higher in that of successful campaigns. This suggests that the number of backers are further away from the mean. In some successful campaigns, the amounts collected are much higher than the goal amount, back a factor of 2 or even 3. This suggests a greater number of backers. Therefore, having a higher variance than that of failed campaigns, as they would have a lower number of backers with most of them being much closer to the mean or 0.*