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

The total number of successful campaigns is 565 out of the 1000 entries, which accounts for just over half or 56% of the total entries. Out of all these the parent Category “theater” has the most successful entries with 187, while also having the highest number of failures at 132 out of 364 total failed campaigns. At first glance because of this “theater” appears” to be the most successful but only has 54% of its entries being successful. While the category “Technology” has 67% of its entries as successful. From this information I can determine that even though “theater” has a high amount of success due to having more entries within its category.

When comparing the months in which Crowdfunding campaigns were started, I see that the number of canceled campaigns stays consistent. The number of successful campaigns stays consistent through May and starts to increase in June and peaks in July, dropping immediately in August to its lowest point. From this information it does seem like one could decide on when the best time frame to start a campaign would be.

Looking into the data comparing subcategories only two categories have a 100% success rate, “audio” and “world music”. I do not believe that these two would in turn be the most successful due to their small number of entries, with only 7 entries combined.

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

A limitation I noticed while answering the previous question is the high number of entries in the “theater” category. I found it difficult to accurately analyze the data to find out which category would be the best to invest in. The question I wonder is what the data would look like if the totals for each parent category where all closer to the same number.

A larger dataset could help us more accurately analyze funding, 1000 entries I don’t believe would be enough to draw conclusions on which to make decisions accurately.

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

We could create a table to compare the amount of time a campaign ran to see if there is a correlation between time and whether the campaign failed, was successful, or was canceled.

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

From this data I would determine that the mean would better summarize this data for us. The mean shows the average number of backers for each campaign, while the median would just tell us what number of backers falls in the middle of all backers in the data we’re analyzing. So, if we are looking to see and compare the number backers from successful campaigns to failed campaigns the mean would be the best to use.

**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. I do believe this makes sense as the number of backers only tells us how many backers took part in the campaign. A successful campaign with a small goal would have less backers. This tells me that variability wouldn’t play a big role on what makes a campaign fail or succeed. It serves us to only compare or with each other.