Sarah Ryan’s Challenge 1 Graph Analysis:

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

From analyzing the data, I have drawn three main conclusions:

1. **Theater, Film & Video, and Music Dominate**   
   The top three most popular campaign parent categories are theater (344 campaigns), film & video (178 campaigns), and music (175 campaigns). Together, these categories represent approximately 70% of all campaigns.
2. **Success Rates Vary Across Categories**  
   The success rates for these top categories fall between 54% and 57%. Theater has a 54% success rate (187 out of 344 campaigns), while both film & video and music show slightly higher success rates at 57% (102 out of 178 and 99 out of 175, respectively). It is important to note that theater has roughly double the number of campaigns compared to the other two categories. Meanwhile, the journalism category, though small with only four campaigns, achieved a 100% success rate.
3. **Timing Influences Campaign Success**  
   Campaigns launched during the summer months, particularly in June and July, have the highest success rates, as shown in the line graph. Timing appears to play a significant role in maximizing success.

Additionally, looking at the subcategories shows that plays dominate with 344 campaigns and a 54% success rate. In music, rock stands out with 85 campaigns and a success rate of 58% (49 out of 85).

**What are some limitations of this data set?**

From my analysis, I identified three main limitations in the dataset:

1. **Uneven Distribution Across Categories**  
   The dataset is heavily skewed toward a few categories in the parent category—theater, film & video, and music—which account for 697 out of 1,000 campaigns. Despite their popularity, these categories have moderate success rates, ranging from 50% to 60%. In contrast, smaller categories tend to have higher success rates. For example, journalism achieved a 100% success rate but only included four campaigns, while publishing had 67 campaigns with a 60% success rate (40 out of 67). Technology stands out with a 67% success rate, despite being less represented. This uneven distribution limits our ability to generalize findings across all categories.
2. **Outdated Data**  
   The most recent campaigns in the dataset are from early 2020, with the last five campaigns occurring in January and February. This timing is significant because March 2020 marked the beginning of the global COVID-19 lockdowns, which likely influenced crowdfunding trends. It raises the question of how the pandemic and subsequent years have reshaped success rates and popular campaign categories, which this dataset cannot capture.
3. **Geographic Bias Toward the United States**The dataset is predominantly based on campaigns from the United States, accounting for 76% of the total. This geographic focus may skew the results toward U.S.-centric trends and preferences, limiting the applicability of the findings to other regions or global crowdfunding behavior.

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

I’m interested in examining the success rates of categories with fewer campaigns, such as journalism, which currently has a 100% success rate. This makes me wonder whether smaller categories tend to achieve higher success rates overall. To explore this, I would create a scatter plot with the number of campaigns on the x-axis and the success rate on the y-axis. Adding a trend line to the plot would help identify if there is an inverse relationship between the number of campaigns and their success rates. This analysis could guide our campaign strategy by helping us decide whether to focus on smaller, high-impact categories or larger categories with moderate success rates.