1). Given the provided data what are three conclusions that we can draw about crowdfunding campaigns

- First conclusion I made related to the months that the campaigns were launched the first and second most successful months to launch a crowdfunding campaign were June and July. Therefore, we can conclude that if you want a better chance of having a successful campaign you should launch your crowdfunding in the summer.

- Second conclusion when looking at crowdfunding by parent category theater had the most successful launches at 187. This is a deceptive stat as theater also had the most entries. Therefore, theater had a total of 344 entries and 187 of those were successful. This might lead people to think that theater was the most successful category of crowdfunding. This is wrong as film and video had a total of 178 entries and 102 successful. Therefore, film and video had a higher successful percentage at 57% rather than theaters 54%. In conclusion to have a more successful crowdfunding campaign you will want to enter under the category of film and video.

-Third conclusion is based off sub categories in these sub categories the highest percent success rate was photography books but , documentaries were a close second. This leads me to conclude that if you want to have a more successful crowdfunding campaign you will need to include pictures or film into your idea.

2) what are some limitations of this dataset.

- the first limitation I saw is that out of the 1001 entries we have not all categories are represented equally which leads to discrepancies in the data.

- second another limitation is its limited sample size a data set with only 1000 entries may not be enough to understand the full scope of a crowdfunding enterprise.

- third bias a smaller dataset may tend to skew data towards or away from certain areas of the crowdfunding scope without truly representing the entirety of the population.

3) what are some other possible tables and or graphs that we could create and what additional value would they provide.

- First, I would make a pie chart that shows the demographics of the backers. This would help people pick out backers that relate the most to their crowdfunding. This data would be very valuable to crowdfunding with a small number of backers you want to make sure they are the right ones.

-second I would want to add a time series analysis which would be a line graph that shows success and failure rates over time. This would be helpful for finding trends as well as looking for rises and dips in rates depending on season, category, number of backers etc.

4) use data to determine whether the mean or median better summarizes the data.

- I believe for this data set that the mean better summarizes the data because of the many outliers in the data set. Just finding the middle number from a set of backers that goes from 1 to the thousands doesn’t really tell you anything about the data itself. Rather the mean can tell you the average number of backers a successful crowdfunding startup had.

5) use your data to determine if there is more variability with successful or unsuccessful campaigns

Based of the standard deviation of the two groups successful having 1266 and unsuccessful having 961 we can conclude that the successful crowdfunding campaigns have more variability. I think this makes since one because there were more successful campaigns than unsuccessful. Therefore, the more data points the more variability a certain variable will have.