1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
2. What are some limitations of this dataset?
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
4. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

1)The three conclusions we can draw from the Kickstarter campaigns are:

1. Conclusion:

The data shows clearly that some categories are easier to fundraise for than others. The most popular categories, for example, theaters raised the most money and had the most campaigns. This is also shown when analyzing sub-categories. When looking at sub-categories of entertainment animations and nonfiction were both greater than 62% successful.

Analysis also shows that photography, publishing and technology were more successful than others. On the other hand, although journalism had only four campaigns, they were all the successful. This is difficult to evaluate because of the small sample size.

1. Conclusion:

The larger the campaigns goal the smaller the probability of success.

1. Conclusion:

The data also shows the months of the year influences the success rate

of campaign. For example, June and July were the best months for fundraising and May and August were the worst.

1. Limitations:

We are limited by data samples size when trying to compare categories, for example, journalism has only four campaigns and had a 100% success rate, whereas theater with almost 350 campaigns had a success rate of 54%. It is possible that journalism’s success rate would be lower with a higher a number of campaigns.

1. Charts and Graphics:

One could make histograms and box plots to compare the different in distributions of successful verses unsuccessful campaigns in regard to the length of campaign times. This could discover the best length time needed to run a successful campaign.

1. Variability:

Successful campaigns had an approximately 30% higher variance than unsuccessful campaigns. This could be due to the fact that the successful campaigns variance had more backers than unsuccessful campaigns. However, the histograms were both strongly skewed to the left, so they had very similar distributions. The box and whisker plots for successful and unsuccessful campaign distributions were also very similar except for absolute value quantities.