1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

From the data we can see from the main categories food, games, and publishing have the highest failure rates with the greatest being 70 percent (food). Furthermore out of all the categories technology has the highest cancellation rate of about 36 percent. Based on the data about project creation there does not seem to be any benefit or disadvantage to launching a project during any time of the year except for August. All the other months had similar failure rates for projects throughout the years but August stood out with the highest failure rate of 60 percent.

2. What are some limitations of this dataset?

One of the main limitations of this dataset could be that we only looked at data from 4000 projects out of the total 300000 which means our sample set could be an outlier that is not distributed like the population of projects. The instructions stated only about a third of the projects were successful but in our data set, every goal under 45000 saw a success rate of almost 40 percent or higher. Additionally I noticed the amount of data collected on projects prior to 2015 is much smaller than data collected on projects after 2015. This could have potentially skewed our graphs and analysis if we were to filter them on a yearly basis.

3. What are some other possible tables and/or graphs that we could create?

Another table we could have made that would have could’ve helped us better understand the data would be to set up a table that take the launch and deadline dates and figure out how much time was allotted to reach the funding goal. From this we would be able to see if projects that allowed more time to reach funding goals were more or less successful than projects that allowed less time.