Name: Henry Dinh

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?

Answer: We can make the following conclusions about Kickstarter campaign based on the given data:

1. From table 1 (Excel sheet 2) we see that out of all the parent-categories, theater-based projects have the highest successful rate with the ratio 839 out of 1400. On the other hand, food and game-based projects have some of the lowest successful rate with ratio 34/200 and 80/220 respectively. Interestingly, all journalism-based projects were canceled. Based on this information from table 1, we can make the following conclusion; theater themed projects have the highest demand and success rates which lead to higher supply of new starter projects. On the other hand, the lack of demand for journalism projects leads to the cancelation of all of its projects.
2. From table 2 (Excel sheet 6) out of all the subcategories, plays has the highest successful rate as well as failed rate. This is due to the law of large numbers since the number of startup projects here out weights all other sub-categories. Based on the information given, we can conclude that plays is the most demanded sub-category of theater.
3. From Table 3 (Excel sheet 13) we see huge fluctuations in the successful rates from January to December in comparison to other ratios. Interestingly, we don’t see many fluctuations in the canceling rate and failed rate. We also observe that the number of successful projects out weights the number of projects that were canceled or failed. Based on the given information, we can conclude that fluctuations in successful rates, cancelation rates, and failed rates can be explained by the law of large numbers.
4. What are some of the limitations of this dataset?

Answer: This data set only allows us to make estimations based on successful rate and failed rate of the projects created in their respective categories. Some of the questions that we would like to answer are:

1. What factor(s) determine the success and failed rate of a project?
2. What are some aspects that determine the number of startup projects in a given category/sub-category?
3. What factors determine the market trend?
4. What factors determine the demand for startup projects in a given category or sub-category?

The limitation of the original data table is that it only shows the number of successful, live, canceled and failed rate but omit the factors that determine the above variables. Therefore, making some of the questions listed difficult to answer.

1. What are some other possible tables/graphs that we could create?

Answer: As established in 2, in order to answer some of the questions from a to d, we going to need more information such as locations of the startup projects, the market demand for a certain category in a given location, etc. Next, we will choose our controlled and dependent variables based on the correlation that we want to study. In accomplish this task; we want to construct a Regression table showcasing the correlations between the controlled and dependent variables. Next, we can construct a p-test or t-test in order to reject or confirm the null hypothesis.