Homework question 1:

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

Based on the chart of outcome by category, the count of successful outcome of “theater” category is significantly higher than the other categories. Campaigns in the “Music” category have the highest successful rate. Meanwhile, there is no successful precedent in “journalism” category.

Based on the outcome per sub category chart, it can be concluded that the amount of successful campaigns in “plays” sub-category is higher than other sub-categories. Meanwhile, the successful rate of “documentary” and “rock” and “hardware” sub category is 100% with sufficient counts of successes (counts>140)

Generally speaking, there is no significant difference in terms of the starting months of the campaigns towards success. The accumulating counts of successful campaigns from 2009 to 2017 basically evenly distributed through the year.

Homework question 2:

What are some of the limitations of this dataset?

First, the sample size of the dataset is ~4000 cases compared with over 300,000 projects launched on kickstarter. Clearly the sample size is relatively small, which might not be able to well present the true outcomes of all the projects.

Second, most of the cases are from 2011 to 2016, only 59 of them are newer projects launched in 2017. With the fast-changing economy from 2017 to 2018, these datasets could be outdated and cannot be used to properly present the trends of successful projects.

Third, according to the statistics provided by kickstarter(<https://www.kickstarter.com/help/stats>), the categories we examined at here didn’t include some of the categories raised significant amount of money (etc. Design) or had highest successful rate (dance), therefore the analysis based on the dataset given could be inaccurate.

Homework question 3:

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

We could check backer counts per category to see if any of the categories (sub category) being especially attractive to backers. Also, we could do average donation per category (sub category) to see if there is any category attracts big money.

We could make graph comparing the duration of the campaign per category, to see if any of the categories raises money faster than the others.

We could also check the region of the successful campaigns to see if there is any relationship between the success and the country in which the project is launched.