Tianchi Zhang

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. See Tab “pivot1”: The majority of the projects selected in this dataset are successful at reaching or exceeding their fund goal (more than 50%)
   2. See Tab “pivot1”: Among all the categories, “music” has the highest success rate (77%). So as an organization that focuses on a music related project, the likelihood of reaching the fund goal is very high compared with all the other categories. The category that has the lowest success rate is journalism (0%). However, there are only 24 projects in this database and they were even all canceled! It’s likely that there aren’t many fundraisers focusing on journalism on Kickstarter at all.
   3. See Tab “pivot3”: Projects created in May has the highest success rate (61%), April ranks as the second (60%).
2. What are some limitations of this dataset?
   1. It doesn’t represent the entire dataset of all the projects because it’s only a sample of 4000+ projects.
   2. Canceled projects could skew down the success rate and not as meaningful since we do not know why they were canceled. The could have a chance to making it to success but for some reason they were canceled in the process. But for the sake of data analysis we could potentially exclude them from our analysis.
   3. The information that the “Average Donation” can reflect is limited. If we have the dataset of how much each backer donated, the dataset could potentially be skewed instead of being a typical normal distribution. So if the average donation of a project is $50 and there are 10000 backers, there could very well be someone who donated $90,000 but the rest 9999 backers only donated from $41 each. Therefore, it should be carefully evaluated before we use that information to come up with any conclusion.
3. What are some other possible tables and/or graphs that we could create?
   1. We can calculate the difference between the end date and the created date to get the duration of the projects, and then create a scatter plot between that and the amount pledged (funded) to see if there’s any correlation. This could serve as insights to decide how long the crowdfunding should last since longer duration doesn’t necessarily mean more funds.
   2. We can make a pivot table with “Country” as the row input, “Category” as the columns, “Average Donation” as the values, and “State” as the filter for comparison. With this table we can see which category has the most per capita donation across countries. Then, doing a 100% stacked column chart could give us visualizations of each country’s primary interest of funding…