Northwestern University Data Science Boot Camp Homework 1

1. The music category has the highest success rate of all the categories. Within the music category, rock and indie rock have the highest number of successes. The amount of success for the music category starts off the highest in January and progressively decreases over the year.
2. The sub-categories are still vague. It would be helpful to have more sub-categories within the sub-categories, such as the kinds of foods and plays. It would be interesting to see which type of plays have higher success rates. The difference between the number of plays in the dataset compared to other sub-categories is extremely large. This makes me think that the backers may have some preference and bias toward plays.
3. Another graph we could create includes a line plot with “years” in the x-axis and “state” counts in the y-axis. We could also create a line plot with the “date ended conversion” in the x-axis and “state” counts in the y-axis. We could make a table showing the average donation for each category, or a table showing average donation for each sub-category with a category filter.