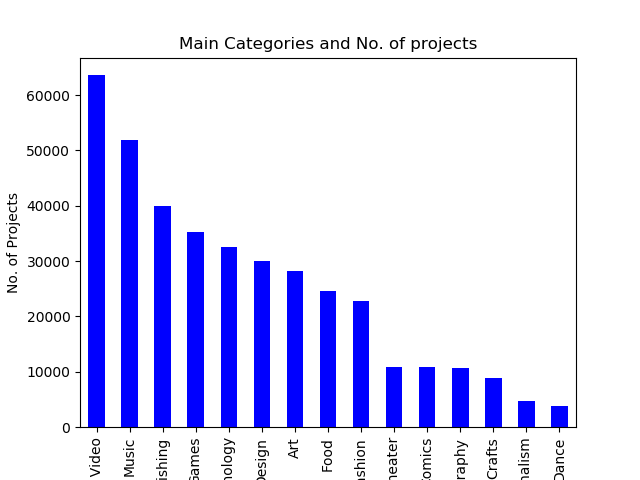
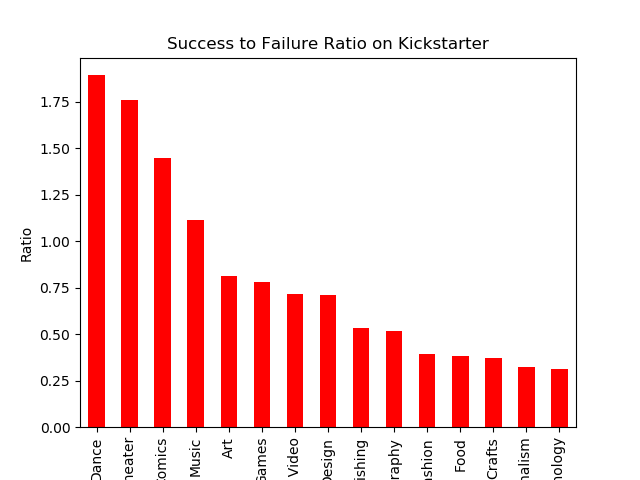
**Prepared by: Aashish Panta**

I started analyzing the dataset of 378,661 projects, each divided into either of 15 main categories, and state of either successful, failed, canceled, live or suspended. The given data set was analyzed and based on the analysis using python, I came to know about some pretty interesting facts and trends of the data. Out of these, 77% of the total projects were from the single country USA. After analyzing, I came to know that most of the projects were related to the main category of video and music, where dance and journalism were the least common category. The bar graph of main categories and the number of the projects are shown below:



Out of all these projects, about 52.2% of the projects were failed whereas almost 35% projects were successful. 10% of the total projects were canceled. Overall, these are the most significant state of the projects. ‘Dance’ category had success to failure ratio of 1.89, which is the most among these projects. It was followed by theater with ratio of almost 1.76. Technology, journalism and crafts performed the worst with the success ratio of 0.31, 0.32 and 0.37. Here is attached the bar graph of success to failure ratio:



From the graphs above, it can be concluded that even though ‘Dance’ category has the least number of projects, it has high chance of success. So, any new coming Dance projects is likely to be more successful than other categories.

Digging more, I came to know that 77.28% of the total projects was from Us, followed by Great Britain with 8.89%.

To get this information, about 300k datasets were trained and rest were tested using both naïve Bayes, decision tree and logistic algorithm. Using Naïve Bayes, the overall accuracy was found to around 0.67, whereas using LR it was around 90%. This verified the accurateness of the analysis done in the summary.