It can be observed that U.S based projects are the most frequent. There are 763 projects that are based in the U.S while the other countries were at or below 48 projects. The make up about 76% which is more than half of the total projects within this population. The data indicates that crowdfunding projects are most frequently held within the US. Another conclusion is that most of the projects in this data set come from the theater category at 344 projects falling within theater. Theater/play projects make up about 34% of the total population of projects while the other categories made up smaller amounts of the total. It can also be observed that projects happened the most frequently during the month of March. 108 projects occurred in March which was about 11% of the total population while the other months saw 9% of the total projects or lower. Other conclusions that can be drawn from the data is that audio and world music sub-categories had projects with 100% successful projects relative to all other sub-categories.

For this data set that has to do with crowdfunding someone’s project, a limitation could be outliers. For example, sub-categories world music and audio have 100% successful projects. These outliers may not be helpful in determining whether a project is going to be successful or not. Just because the project that happens to be under the sub-category world music or audio is going to become successful. What these outliers can tell us is that there is a niche in the market for opportunity in these two sub-categories. Some other general limitations of data include data quality, missing data, bias, data migration, measurement errors, etc.

Other possible visual forms of data could be comparing whether projects were successful or failed based on the amount funded and using a bar graph as well as a box graph to portray that. It would help the reviewer get a better grasp of which how much a project needs to have a chance at being successful. When comparing whether projects were successful or failed based on country, it would be easier for the reviewer to utilize a pie chart. The reviewer could easily see that U.S projects are more common, and it may be competitive to enter the U.S market.

Statistical Analysis

Based on the box graph, both successful and failed projects have skewed data towards having around or below 1,000 backers. The outliers appear to be at or above 3,000 backers for successful projects and at or above 2,000 backers for failed projects. Outliers are the rare occurrences of data unlike the data that is considered normal (most frequent). Outliers pull the average towards it. The median best summarizes the data because the impact of outliers pulls the mean away from the actual middle of the data. There is more variation for successful projects. This maybe due to the volatility of successful projects. Projects can become successful for reasons other than the amount of backers. The amount of backers doesn’t always determine the outcome of the project.