**Report**

**Conclusions:**

The category ‘theatre’ and sub-category ‘plays’, had the highest number of successful projects, suggesting that projects may be more successful if they are within these groups. Following theatre, ‘music’ had the next highest number of successful projects with the sub-category ‘rock’ being the most successful. All projects for jazz music had failed. Out of the 4,115 projects, more projects succeeded than failed, this is shown in the graph (Outcomes based on launch date spreadsheet). Projects which were launched in the second quarter had the highest number of successful cases, this gradually decreased in quarter 3 and 4.

**Limitations:**

The dataset only contains 4,115 projects out of the supposed 300,000+. It is difficult to draw a conclusion as there are a large number of projects from the years 2014-2016 with fewer projects for all of the other years. It has not been made clear whether there were fewer projects in the other years or if the selection of projects was biased. As a result, any trends identified may not be accurate. Also, collecting data from ‘live’ projects

**Other possible tables and/or graphs:**

A scatter plot which compares the amount of money raised vs the number of backers. This will tell us whether certain categories/sub-categories receive larger amounts of money from each backer as opposed to numerous backers that each donate a smaller sum of money. From this we can learn which projects appear to be more profitable, thus encouraging larger donations. In addition, the time between project launch date and deadline can be calculated in a table. The table could show the amount of days each project had to reach their goal, and then an average can be calculated for each state to see whether there is a relationship between the number of days and the success rate of a project.

**BONUS:**

**Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

In this case the median may be more representative of the dataset. There is a large amount of variability in both the successful and unsuccessful campaigns. Large or small value would skew the mean. The median would be more accurate in representing the data as it will not be skewed.

**Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

There is more variability in the successful campaigns. The goal differs largely for each project with some projects requiring a small amount which can be raised by one backer, while others may require a very large amount. Due to this, the successful campaigns would have greater variability, the unsuccessful campaigns did not meet their goals which may have been very high, as a result they would have fewer backers on average compared to the successful projects.