Homework 1 Report

**What are three conclusions we can make about Kickstarter campaigns given the provided data?**

Based on the data provided, there are few trends that are apparent from the analysis done. Out of all the category invested in, Theater category is most invested in by users. While theater is the most invested category, it comes in second in success rate with a value of 60%. The most successful invested category was music with a value of 77%, which is 17% different from second position by success rate. Music was also the second most invested in by users after theater.

Another interesting trend that is apparent by sub-category is that the plays sub-category is the most invested in by a significant margin from position two. Plays had 1066 of the count out of total 4114 data points, this puts plays at 25% of all the investments made my users. Second position for most invested after the plays is rock which is only 6% of the total investments.

By viewing the data by date created for the projects, it seems that projects started in December have a significant disadvantage for success rate compared to rest of the year. The average for rest of the year for success rate is 53% while the success rate in December is around 43%, which is 10% lower than rest of the year’s average.

**What are some of the limitations of this dataset?**

One major Limitation for the analysis and dataset is that all the projects are grouped by the categories and sub-categories; this over simplifies the projects details which might mask important details about the specifics of the project that might have contributed to its state and pledge amount.

Limited information given on what the staff\_pick and spotlight columns indicated for the data set, without context of how that was relevant to the project it is hard to correlate what kind of information we can extrapolate by interrogating the pledged or the state data against staff\_pick and spotlight groupings.

**What are some other possible tables/graphs that we could create?**

Look at average pledge = pledged / backers\_count, this will give insight into what class is investing in what category and sub-category of projects. There might be a correlation of types of projects that are invested in by social economic class. This can give insight into types of project to focus on to gear towards a particular economic class, it also might reveal some information regarding the topics that are important to that particular group.

Calculate the days project was open for by subtracting deadline from launched\_at. Graph that information with the state of the project. This graph will give important information for how each category/ sub-category might have days relating to its final state. If there is a distinct difference in days and state, this information could be used to help more projects be successful in each group.