**Kickstarter Report**

1. Dataset Description

The sample data is pass 4114 projects from Kickstarter which are created from May 2009 to March 2017 and ended from May 2009 to Aug 2017. These sample projects are including in 7 parent categories - film & video, food, games, journalism, music, photography, publishing, technology and theater within 20 countries. The minimum goal is 1, maximum goal is 100000000 and medium is 5000. There are 2185 succeed projects, 1530 failed projects, 349 canceled projects and 50 live projects.

1. Data Analysis
   1. Countries and Categories

US is country which has most projects from the sample dataset. There 3038 Kickstarter projects from US, it is 73.85% among all the countries. The second most-projects country is GB. It has 604 projects within the time frame, 14.68% in total.

Theater is the parent category which has most projects created within the time frame among all 7 parent categories – 1393 of 4114 projects. 60% of them are successfully reached goal, 35% failed, 3% canceled and 2% in live. The music category has the most success rate which is 77%. The journalism has least projects created and least success rate. Only 24 projects created and 24 out of 24 are failed.

Figure Countries and categories

* 1. Categories and Subcategories

There are 2 subcategories have more than 200 projects – plays and rock. Most of the projects are in subcategories -plays. Plays is the only subcategory which has more than 1000 projects created with 65% success funding rate. Classical music, documentary, shorts, rock, electronic music, hardware, metal, nonfiction, pop, radio & podcasts, tabletop games and television have 100% success rate in pass 9 years.

Figure Categories and Subcategories

* 1. Months and status

Based in figure 3 below, we can see that the total counts of canceled projects didn’t change too much across months. Counts of succeed projects increased in March and reached the highest in May then decreased after that. December, the total succeed counts is less than total failed counts.

Figure Months and Status

1. Limitations of dataset
   1. Dataset Size

This sample dataset contains 4114 row in the time frame of 9 years. Variability is determined by the standard deviation of the population. Reducing the sample size reduces the confidence level of the study. A smaller dataset also can lead to cases of bias which is uncovered in the sample. Getting larger dataset could help to insure accuracy by avoiding sampling errors.

* 1. Countries

Among 20 countries, more than half of sample, which is 73.85%, from this dataset are from US. It will increase the bias of the data if we would like to use this dataset for countries which is not US. Furthermore, there are 7 countries has less than 10 projects in 9 years. It will also decrease the accuracy of our analysis for those countries.

* 1. Categories and Audience

Kickstarter is a global crowdfunding platform focused on creativity and merchandising. It means the type of projects which funded in Kickstarter is influenced by its audience. In the sample dataset, all the projects are belonging to 7 parent categories. Most of the projects created in Kickstarter are in Theater “”, Music “” and film & video which are more trend to be in entertainment type. This may impact our analysis with other categories due to small sample data of other categories.

1. Other tables and/or graphs

I would add the counts of projects created in each year line chart. It is interesting that 2013 to 2016 has much projects created than 2009 to 2013. After reached to the highest count in 2017, the total projects decreased rapidly. It could be related to trend, expectation of economic or the popularity of Kickstarter if we do more research on it.