

## Declaration on Plagiarism

Name:	Akhileshkumar Mutuguppe Subbarao
Student Number:	19210140
Programme:	MSc. In Computing (Data Analytics)
Module Code:	CA682
Assignment Title:	Data Visualisation
Submission Date:	16 Dec 2019
Module Coordinator:	Dr Suzanne Little

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. I have read and understood the Assignment Regulations. I have identified and included the source of all facts, ideas, opinions, and viewpoints of others in the assignment references. Direct quotations from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged and the source cited are identified in the assignment references. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

I have read and understood the referencing guidelines found at <http://www.dcu.ie/info/regulations/plagiarism.shtml>, <https://www4.dcu.ie/students/az/plagiarism> and/or recommended in the assignment guidelines

Name: Akhileshkumar

Date: 16 Dec 2019

# Kickstarter Project Outcome Analysis (2009-2018)

## Abstract

Kickstarter is a big platform which built with the aim of helping creative projects. Kickstarter help gamers, artists, filmmakers, designers and many creators to find the required resources and supports to make their ideas into reality. Thousands of projects have built with the support of Kickstarter community. The story to be depicted will be based on Kickstarter projects analysis and what was the outcome of the projects over the years (from 2009-2018).

## Dataset

Dataset is taken from Kaggle and name of the file is “Kickstarter Projects”. It is a big data and has huge amount of information on the details of Kickstarter projects from 2009 to 2018. Dataset is a 46 MB csv file and has 378661 rows and 15 columns. All the projects sub category, main category, launched date, deadline, outcome, backers, country, amount pledged (in US \$) are considered. Dataset is a combination of different datatypes like int64, float, object. Since there is a large amount of data present in the Volume of the data are taken into consideration.

## Data Exploration, Processing, Cleaning

Dataset is a CSV file. Python code was written to generate the required data in accordance with my visualization aims. All cleaning and Processing are done on jupyter notebook. First data read from the csv file and all the column details and datatypes are checked. Numerical values are converted to float to maintain a uniform format. Null values in the “name” columns were replaced by “not available” string. Data set had some junk null and “launched” date year with 1970, which are removed in excel itself. Columns like pledged, goal, currency, US\$ pledged were dropped as they are repetitive columns and are not used in the visualization. “Country” column has only two letter abbreviation for all countries (called ISO 3166-1 alpha-2 format) which were replaced with the help of pycountry package. The cleaned dataset is then visualized in Tableau desktop. All the columns are used to depicts the complete analysis and to get the outcome of the Kickstarter Projects.

## Visualisation

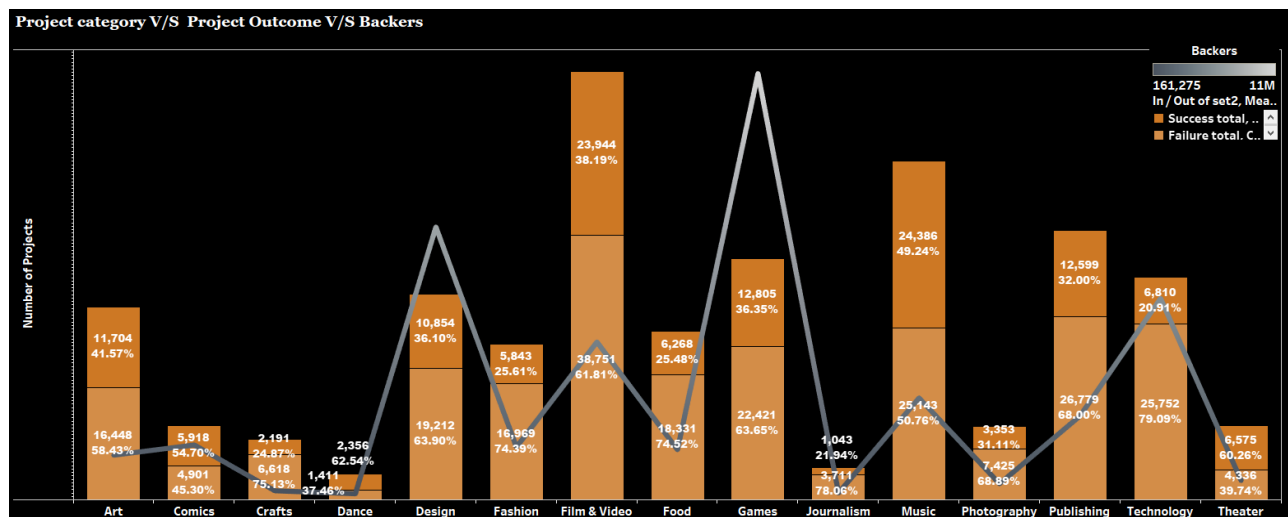
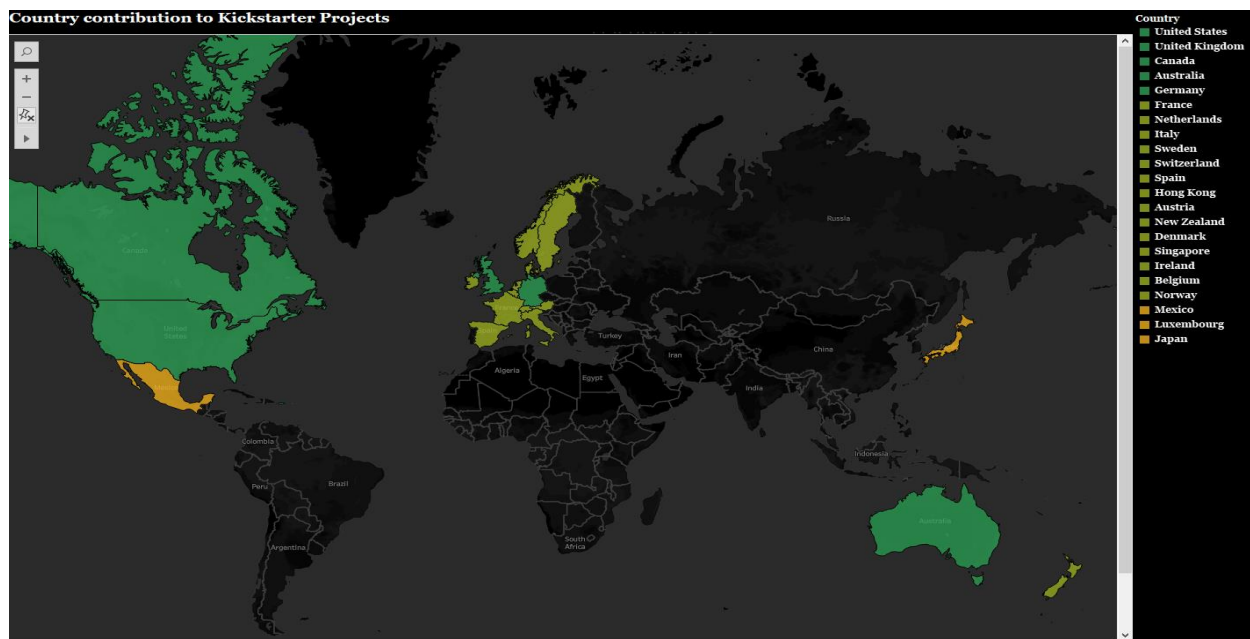


Figure 1

**Figure 1** is a complex graph which describe the multiple layers of the Kickstarter Project. Bar chart in the graph describes success and failure (either cancelled, unknown, failed, suspended) counts and percentage in different project categories. On the other hand, line graph illustrates the total number of people who backed these projects. From the graph it is very clear that highest number of Kickstarter projects were based on “Film and Video” category, but we can see the backers to this category was less as well as the success rate was less(38.19%). This implies people have carefully screened the projects before pledging the amount to the projects in this genre. This was not the case in all categories, a smaller number of people supported “Dance” and “Theater” category, but the success rate in this category was more than 60%. Similarly, Games is the most popular category with more than 11million backers, but this category does not have high success rate. Graph clearly indicates that projects with less support from people were more failed, cancelled or suspended except few genres. This shows popular categories like games were backed irrespective of the project potential. gold shaded colours have been used to represents the success and failure counts as it is a lively vibrant colour and blue is more subtle colour used to differentiate the line graph from the bar chart. A dark background is chosen to make the graph more standout and to provide positive colour insights. White font has been used to make the numbers, title and axis more enhanced on dark background.



**Figure 2**

**Figure 2** illustrate contribution of each country in the Kickstarter Projects. Since Countries involved, map graph is chosen. Country’s contributions are sorted and after sorting United States being the highest contributor to the Kickstarter community with more than \$2.8 Billion dollars and lowest being Japan with pledge amount \$1,15,888. Top 5 contributors are illustrated in green to represent highest pledges and shades of green representing the decrease in contribution. Bottom three countries Mexico, Luxembourg, Japan represented by Yellow colour as these colour compliments the green colour very well and make graphs more appealing. Dark background to map is chosen with an aim to intensify the contributed countries.



**Figure 3**

**Figure 3** is a combination of multiple aspects of Kickstarter Project. A bar graph representing Success and Failure Percentages of Kickstarter Projects over the years. It helps us to identify what were the outcome of the projects. It has two components merged together, one being the success and failure rate of all the pledged amount and other is total success and failure rate of all the backers of the Projects. By looking at the Graph it is clear that projects which are pledged and backed by the people had high percentage of success rate in all the years. This shows projects which had good potential were supported and funded by people. On the other hand, funds were used efficiently by Kickstarter project creators to make these projects successful. Green colour is used to represent the success rate and since red is a complimentary colour of green, it is used to represent the failure of the project, dark background is used to create positive intent and to make graph more imploring.

## Conclusion

Data analysis with visualization provides the best way to see and understand the different trends and aspects of the data. As volume of the data increases, data visualization tools and techniques are much needed to analyse the massive amount of information. A good visualization describes a storey by removing all the noise in the data and highlighting the required information [2]. From the above graphs, for Projects supported less by people were (except few genres as explained above in Fig 1) more likely to fail. Kickstarter projects were supported all over the world. United States, United Kingdom, Canada, Australia, Germany were the highest contributors to the Kickstarter community. These countries were responsible for 94.89% of the total pledged amount. Most of the contributions were done by these 5 countries. Even though many Kickstarter projects were failed over the years, but the projects which are funded and supported have been successful. From the analysis we can conclude that Kickstarter is a great platform for the creators to make their ideas to reality. People might have backed the topic of their interests but the project with great potential were always encouraged by the people. Projects with good potential will be supported by people all over the world. I could improve in building a more interactive graphs and features. I should work on adding the animation to the graphs in future so that the graphs will be more emphatic. I am glad that I could visualize Kickstarter project analysis to the users.

## Reference

1. <https://www.kaggle.com/kemical/kickstarter-projects>
2. <https://www.tableau.com/learn/articles/data-visualization>
3. <https://www.tutorialspoint.com/tableau/index.htm>