# 01-Excel Homework: Kickstart My Chart

Sumukh Kumble

## Analytics Report



### Given the provided data, what are three conclusions that we can draw about Kickstarter Campaigns.

The following observations from the dataset have been noted:

* The industries that gained the most number of successful projects were in theatre, film and music.
* The timeframe indicates that most of the successful projects occur within a specific month of the year, mainly in May, which is the summertime in the norther hemisphere.
* The proportion of successful projects appear to occur when the target goal has been set to a range that is less than 1000. It can also be observed that as the target goals increase, the ratio of success begins to decline inversely proportional to the rate of failure.

### What are some of the limitations of this dataset

* Lack of granularity within the dataset as we cannot make what the age, gender or ethnicity had a role in the success of the projects.
* Additional datasets from earlier years in which to make a comparison and determine if preferences for difference industries follow a specific trend or vary from year to year

### What are some other possible tables and/or graphs that we could create

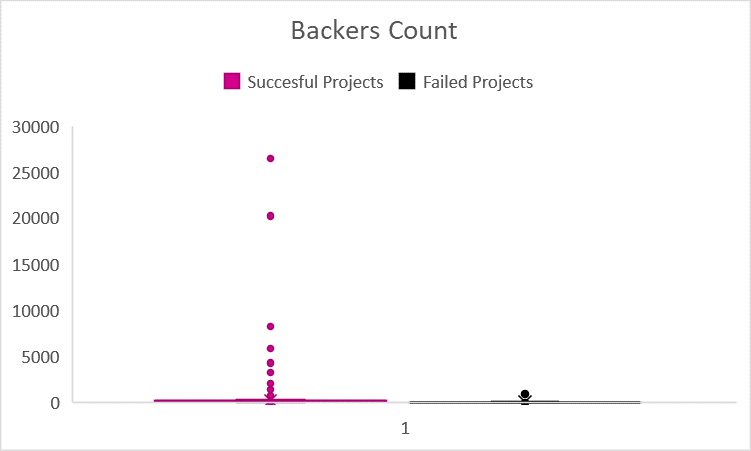
* Create a pivot graph to determine which countries have the most number of successful projects, backers, goals and pledges
* Which countries have focussed on which industries and the success rate in those categories and sub-categories
* Which countries had the most cumulative goals and successful pledges
* How many projects that had pledges that met or exceeded the goals

## Bonus Statistical Analysis

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

* Due to the greater variability that is seen within the datasets for both successful and unsuccessful projects, the median provides a more accurate measure on the general numbers of backers that were available for all projects within the Kickstarted campaign.
* For projects that were successful, the median number of backers were 62 whereas for projects that failed, the median numbers were 4.

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



* From the box and whisker plots, the amount of variability are highest in the projects that were successful. The variability in this dataset does not provide any true insight which can be based on the following reasons:
* The number of backers within a certain goal range are higher when the target is between 1000-1500
* Which industries have drawn more backers for successful campaigns
* Which countries were able to draw the most number of backers and in which months of the year?
* If the datasets were broken down and analysed in different layers, it would be possible to determine if there was a variability in different projects based on industry, country and time of the year.