



Georgia Tech Data Science Bootcamp
800 Spring St NW, Atlanta, GA 30308

Internal Memo

To: Georgia Tech Data Science Bootcamp Teachers
Georgia Tech Data Science Bootcamp Teaching Assistants

From: Ivan Escalona-Faria

Subject: Homework 1 (Report and Excel Workbook)

Date: March 2, 2019

Distinguished teachers and teaching assistants:

As requested, I am attaching the worksheet with the processed kickstart data. I am analyzing the data and providing you with the top 3 things to be on the look for regarding this information. I am also including recommendations for further analysis.

Year analysis:

There seems a clear trend for most of the years the number of successful projects are always greater than the failed ones, this is especially true when looking at more than a single year data. There are some years which don't follow this trend; years 2009, and 2017 don't have a significant amount of data points, and therefore not a big enough sample to extract behavior that is meaningful; years 2014, 2015, and 2016 are years in which the successful and failed campaigns are drawing closer to each other.

Recommendation: It would be important to look whether it is a decrease of successful projects, or an increase of failed projects.

Category and Sub-category breakdown analysis:

Regarding the category breakdown it is seen that the majority of projects are: film and video, music, technology, and theater is the largest one. That holds true for most of the countries with larger number of projects. It is true for those as well that the general trend is that in these larger categories, the number of successful projects is larger than the rest of the projects, except for technology.

In the Sub-category breakdown, it's noteworthy that almost every sub-category falls below 200 projects in the combined data, except for "rock" and largely so "theater". The rate of successful vs failed is very much diverse and doesn't seem to follow a uniform pattern. It seems that the majority of largely populated sub-categories have more successful projects than any other status.

Recommendation: Special attention should be drawn toward the smaller countries regarding the successful/failed/cancelled ratio to have a better idea of how representative are the failed/cancelled projects for those with high project count, as well as those in which the failed amount surpasses the amount of successful projects.

Breakdown of outcome vs goal:

It is significant that those campaigns in which goal is lower seem to be more successful than campaigns in which goal is higher. There seems to be both a negative trend for successful campaigns and a positive trend for both failed and cancelled campaigns as the goal increases. If such a trend is proven to be statistically significant, then a prediction could be made as to whether a project will be successful or not based on the set goal for said project.

Please let me know of any additional requirements or follow up questions you may have.

Yours truly,

Ivan Escalona-Faria
Data Scientist