Kickstarter Campaign Data Visualization

Thomas J. Schantz[A]

[A] *Data Science Graduate Student, Lewis University, Romeoville, IL*

Abstract

The use of sites like Kickstarter for entrepreneurs and inventors has changed the way new products come to market. Taking advantage of the crowdsourcing phenomenon, Kickstarter has become a significant platform for anyone who's looking to share his or her ideas with the world but may lack the initial start-up money to get their business model off the ground. This study explores some Kickstarter data at a high-level in order to make some inferences on the scale of which their services are being leveraged across all industries, big and small.

Keywords: *Kickstarter, campaign, crowd, pledge, launch, backer*

1 INTRODUCTION

1.1 Data Set Description

The data contained within this study was provided by Kaggle as a free, downloadable source of information [1]. The source offered data within a timeseries range starting with its inception in April of 2009 and ending January 2018 [2]. The dataset contains 321,616 data points along with 13 fields consisting of categorical, locational, time-series, and continuous data types. A full list of each field along with their descriptions can be seen in Table 1.

Although the source does not provide much description of what the data has been or will be used for in the future, one would conclude that this set of attributes would be very useful for classification algorithms that seek to identify trends and marketing needs within a particular pledge category or region of the world.

1.2 Scope and Objective

The scope of this study is purely to visualize the dataset at hand so that insights might be gleaned by the reader as to the results of any given Kickstarter campaign. The primary objective would be to provide readers with an introductory understanding to the recent, historical success and failures by category, and to better understand the risk he or she might take in using the platform for any future projects within a given category (e.g. film, art, etc.).

2 ANALYSIS

2.1 Data Visualizations

Given the data spanned a number of years related to dollars pledged and dollars targeted, the first visualization thought to render was a time series plot by year which overlays these figures in order to see how the business has performed year over year (Figure 1). This identified some missing data within the *Launched* field as there were a few points with a value of 1/1/1970. These were removed from view, along with 2018 given there was only one month's worth of data to display.

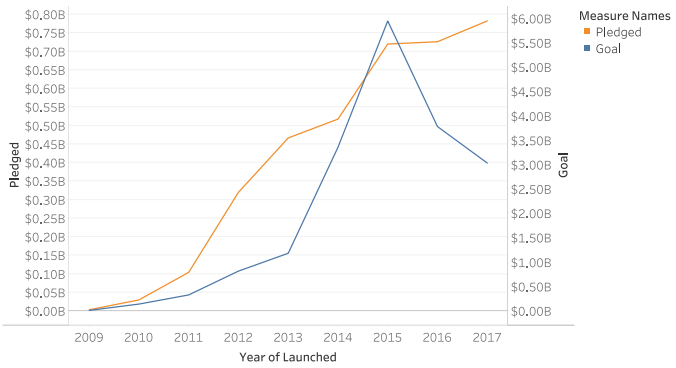

Figure 1. Time series plot of pledged $'s with backer volume over goal $'s by year

From this, one can see how the rise of pledged dollars is on a continual climb, with a slight dip in slope starting in 2015. The goal line, on the other hand, shows a steady rise up until 2013. At which point, a steep shift in slope takes hold for 2 years and turns inverse. It would seem there was something that may have caused a drop in goal dollars between 2015 and 2017.

Table 1. List of dataset field names

#	Field	Description	#	Field	Description
1	**ID**	Internal kickstart id	8	**launched**	Date launched
2	**name**	Name of campaign (i.e. a finite work with a clear goal)	9	**pledged**	Amount pledged by "crowd"
3	**category**	Sub-categor of campaign	10	**state**	Current condition the project is in
4	**main_category**	Category of campaign	11	**backers**	Number of backers
5	**currency**	Currency used to support	12	**country**	Country pledged from
6	**deadline**	Deadline for crowdfunding	13	**usd pledged**	Amount of money pledged
7	**goal**	Amount of money a creator needs to complete the campaign			

Such a drop may have been caused by a decrease in campaign starts. Plotting the volume of campaigns and backers shows a slight negative slope starting around the same time; however, it doesn't appear to be as significant (Figure 2). Never the less, it stands to reason that Kickstarter continues to grow by shear dollars pledged.

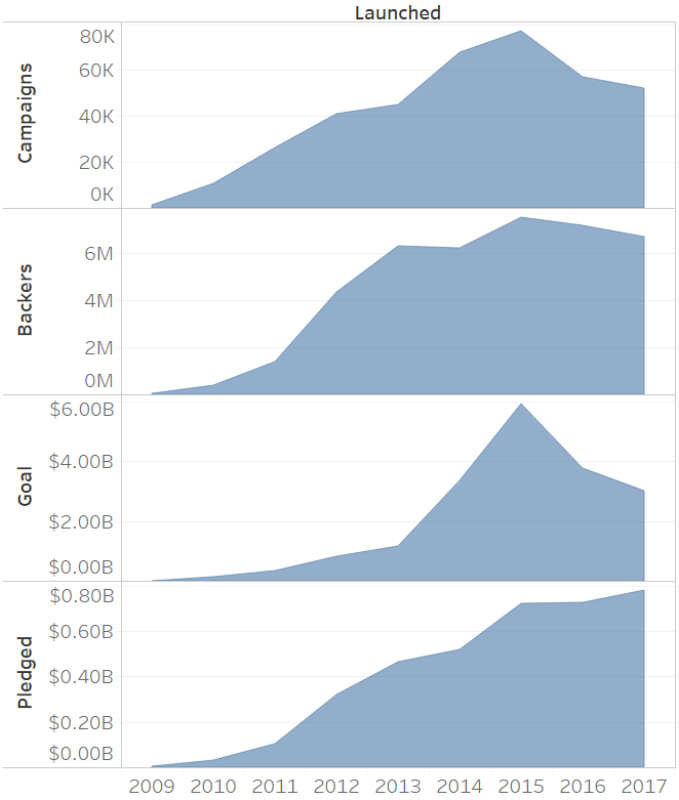

Figure 2. Time series, shaded plots of campaign vol., backer vol. goal $'s, and pledged $'s

Another way in which to measure the success of the company, as well as determine what might drive the success of a campaign, is to look at the different campaign states within the data (e.g. successful, failed, etc.) along with their associated average number of backers. It would stand to reason that the more people who believe in an idea, the more successful the campaign will be in achieving its targeted financial goal. From Figure 3, one can see that the vast majority of campaigns areas successful, with the small portion of them having failed to meet the goal. There are many reasons for why a goal might never have been met, one of which being too lofty and unrealistic. From this same figure, one can also see the large difference in the average number of backers between a successful (i.e. 264) and an unsuccessful (i.e. 16) campaign.

After validating a few business performance metrics for the company as whole, the next thought was to look at which parts of the world contribute most to this success. Given Kaggle's download showed country by their ISO code, a crossmapping of country-to-ISO value was required to display the full country name [3]. Table 2 shows the results of the link created between these two data sets. From this view, it is clear that the US is the cause of the majority of this success with the 77.4% contribution of total pledged dollars. The next highest contributor is the UK at 4.4%. A clear indication that Kickstarter is most prominently used in its country of origin.

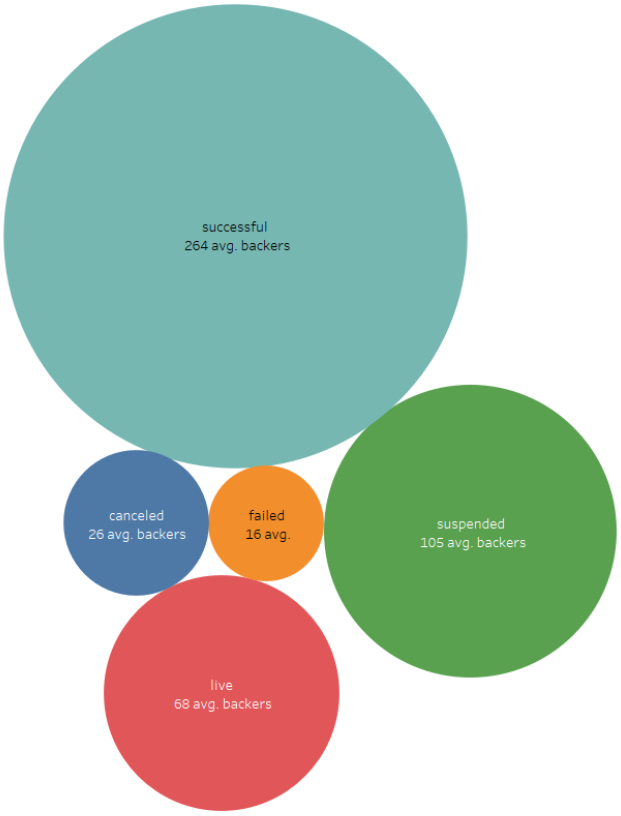

Figure 3. Bubble Plot of average number of backers by campaign state

Table 2. Pledge totals with percent of total pledges by country

Country Name	
Australia	$59.35M (1.6%)
Austria	$8.50M (0.2%)
Belgium	$3.75M (0.1%)
Canada	$113.35M (3.1%)
Denmark	$50.31M (1.4%)
France	$30.78M (0.8%)
Germany	$35.58M (1.0%)
Hong Kong	$78.19M (2.1%)
Ireland	$3.73M (0.1%)
Italy	$14.72M (0.4%)
Japan	$13.08M (0.4%)
Luxembourg	$0.41M (0.0%)
Mexico	$46.59M (1.3%)
Netherlands	$26.06M (0.7%)
New Zealand	$10.79M (0.3%)
Norway	$22.14M (0.6%)
Singapore	$8.10M (0.2%)
Spain	$11.56M (0.3%)
Sweden	$114.01M (3.1%)
Switzerland	$12.86M (0.4%)
United Kingdom	$161.48M (4.4%)
United States	$2,829.76M (77.4%)

In addition, the interesting thing about this view by country is that the average number of backers is much more evenly spread (Figure 4). This might mean that Kickstarter has opportunity to somehow market to more high-dollar contributors in these areas, given their services are well known and utilized by more frugle spenders.

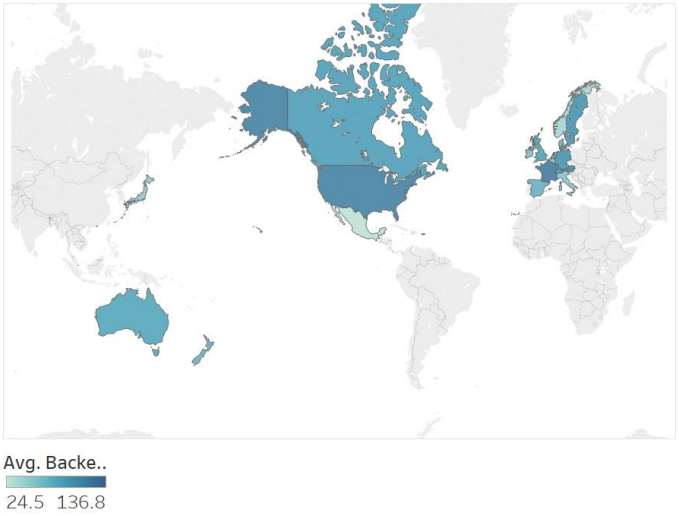

Figure 4. Heat map of average number of backers by country

The last feature which seemed interesting to visualize was the main category of campaign, and which ones contributed the most to Kickstarter's success. From Figure 5, one can see that four categories make of just over 75% of all pledged dollars: *Design*, *Games*, *Technology*, and *Film & Video*. Interestingly, the average pledged amount for the 2nd highest contributor (i.e. games) is ranked 12th with an average pledged amount of $61; however, it holds its standing in overall pledge contributions through sheer volume of backers. It appears to be a very popular category amongst low-dollar pledgers.

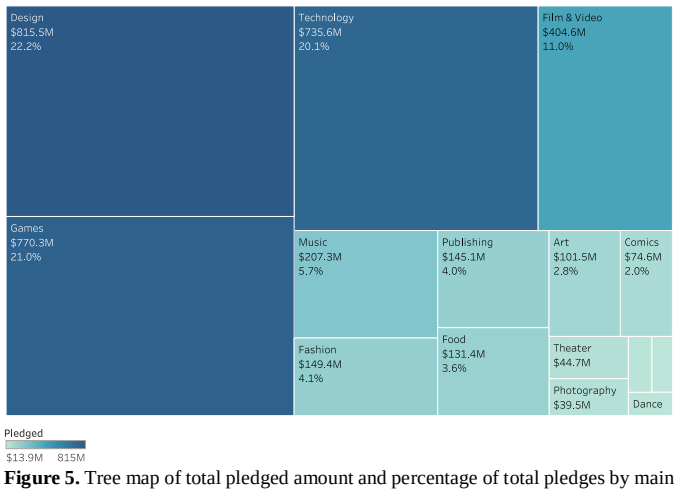

Figure 5. Tree map of total pledged amount and percentage of total pledges by main category

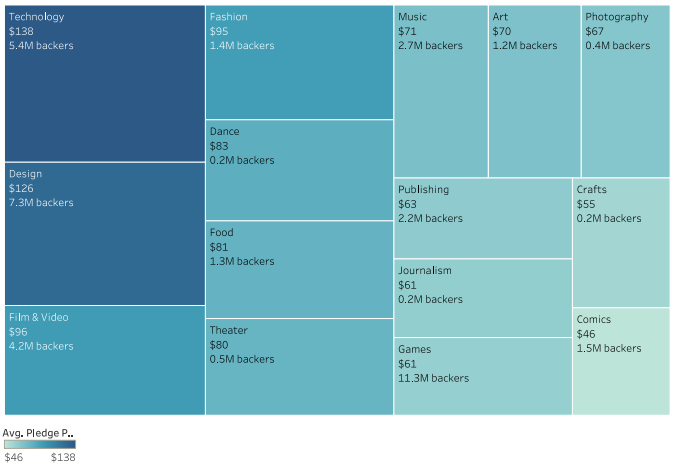

Figure 6. Tree map of average pledged amount and number of backers by main category

Looking at just these four, highest contributors to pledge dollars, the next thought was to look at how each performs in terms of pledged dollars to goal. From Figure 7, it would appear that the *Film & Video* category sets some lofty goals that are difficult to attain, even though it it's one of the top four contributors in terms of pledge dollars. This is intuitive since film productions are known to be very costly undertakings. It also indicates that the *Design* category has the shortest gap to close in terms of pledge dollars meeting the goal, overall. However, Figure 8 and Figure 9 tell a slightly different story.

Figure 7. Bar graph of total goal vs. pledged amount by main category with n indicating the number of campaigns included

The median distance to goal (i.e. how close the pledge amount for a campaign was to meeting the goal amount) in Figure 8 shows to be the lowest for the *Film & Video* category. This would indicate that there are likely some high value outliers that are skewing the gap between pledge amount and distance to goal, overall. Figure 9 confirms this with an outlier plotted above the $4 billion line. With an outlier that massive, it's no wonder the category has such a high variance in this area.

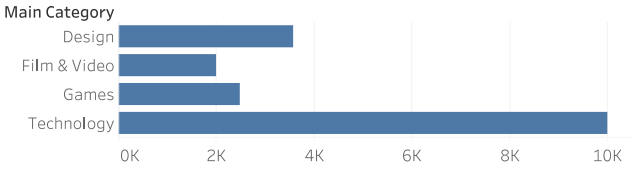

Figure 8. Bar graph of median distance to goal by Main Category

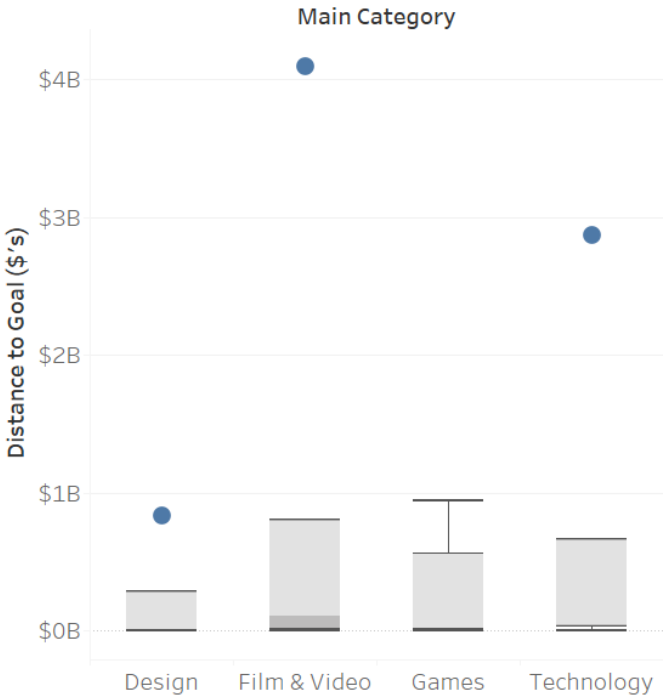

Figure 9. Boxplot of distance to goal by Main Category

2.2 Application of Course Material

The main application to the material provided within week one of the Data Visualization course within Lewis University's MS in Data Science program is that representations can be misleading. As was seen with the comparison of pledged to goal values, one graphic can tell a much different story than another.

As far as the approach the designer took within this study, it followed an exploratory type of visualization more than it did an explanatory one. This is because by definition of the objective, there is no intended action that's being required other than for the design to interpret and understand the data. In a real-life scenario where an analyst or data scientist is using Kickstarter's data to help draw meaningful business conclusions that might drive responsive action, there would almost certainly have been more signs of explanatory visualization involved.

3 Conslusion

Based on the available data, the Kickstarter company appears to be a steady growth trend since inception over 9 years ago. Given there is no shortage of new ideas out in the world, it's likely this trend will continue unless a competitor finds a way to help up-and-coming entrepreneurs bring their product to market at a fast, easier, and less costly way to them and their investors. Given it lends itself nicely to some fairly straight-forward exploration through visualization, the data that the company has shared provides a great point of view on how one might gauge success of his or her campaign before starting, as well as determine some realistic expectations while in pursuit of the entrepreneurial dream.

References

[1] "Kickstarter Projects", Kaggle.com, 2018. [Online]. Available: https://www.kaggle.com/kemical/kickstarter-projects. [Accessed: 27- Oct- 2018].

[2] R. Wauters, "Kickstarter Launches Another Social Fundraising Platform", Techcrunch.com, 2009. [Online]. Available: https://www.techcrunch.com/2009/04/29/kickstarter-launches-another-social-fundraising-platform/. [Accessed: 27-Oct-2018].

[3] "Country Codes, Phone Codes, Dialing Codes, Telephone Codes, ISO Country Codes", Countrycode.org, 2018. [Online]. Available: https://countrycode.org/. [Accessed: 27- Oct- 2018].