****

Data Analytics Boot Camp Program

**Module 1: Excel**

Submitted By: Jorge Maspons

Submitted Date: Saturday, January 26, 2019

**1) What are three conclusions we can make about Kickstarter campaigns given the provided data?**

Based on the provided data of this study exercise, the three main conclusions that can be drawn from analyzing the sample data set of 4,114 Kickstarter campaigns are:

1. The number of campaigns related to Theater (33.9% of total count), Music (17.0% of total count), Technology (14.6% of total count), and Film & video (12.6% of total count) represented about 78 percent of the total sample data among all nine categories.
   * + Within the most popular category (Theater), the sub-category of “Plays” was the clear leader with 1,066 campaigns, or about 26 percent of the grand total sample data.
     + In addition to having the 2nd. highest number of campaigns (700 in total), the Music category had the highest percentage of successful campaigns (77.1% successful within category) among all nine categories, leading the runner-up —Theater campaigns with a successful rate of 60.2 percent— by a difference of about 17 percentage points.
2. The number of launched campaigns were fairly equally distributed among the twelve months of the year with a high season during the summer months and a low season in the final month of the year (December).
   * + Although the data does not provide clear evidence to suggest that the month in which campaigns were launched had an impact on the final outcome, it is relevant to mention that campaigns submitted on May experienced the highest rate of successful outcomes (about 61%) while December was the month with the lowest success rate (about 44%).
3. The percentage of successful campaign outcomes presented an inverse relationship —with a strong negative correlation coefficient equal to negative 0.68— with respect to the dollar goal amount that the campaign needed to secure from investor pledges.
   * Campaigns that required lower dollar amounts (less than $10,000) had a success rate above 50 percent, whereas projects with higher goal amounts (equal to or more than $50,000) experienced a success rate below 20 percent.

**2) What are some of the limitations of this dataset?**

First, it is important to highlight that there is an estimated twenty percentage points difference between the reported success rate of the population —noted to be about one-third from a total of 300,000+ campaigns— compared to the sample success rate —just above 53 percent from 4,100+ campaigns. This significant difference in these two success rates leads to the possibility of the presence of improper sampling; some of the most common pitfalls in sample generation are related to sampling bias, sample representativeness, and sampling methods. If the dataset were to suffer from any of these types of sampling errors, conclusions drawn from the sample can be subject to additional scrutiny and, ultimately, labeled as producing “statistically invalid” results.

In addition, there is no type of ranking information regarding the quality of the submissions besides the “true/false picked by staff” category element that is listed for all campaigns on the sample dataset. Given that all projects/campaigns are reported alike from a quality perspective, it is not possible to analyze if there is relevant connection between the final outcome of the project/campaign and the quality of the campaigns submitted, which may be very well be a critical element for the decision making process of a potential investor.

Finally, a formal definition or description related to each major category element of the dataset is not provided; this limits the understanding of certain dataset categories, such as “spotlight,” and can also lead to improper interpretation of data associated with each category.

**3) What are some other possible tables/graphs that we could create?**

Some examples of additional tables relevant to this analysis are:

Table 1: Currency denomination table by total count

* + - The US dollar (USD) and the British Pound (GBP) accounted for about 89 percent of the total number of campaigns with the USD predominantly leading the way (73.8% of total count).
    - Among the top four categories by total count —Theater, Music, and Film & Video—, except for Technology (78.3% of category count), more than 90 percent of the campaigns were executed either using a USD or GBP currency denomination.

Table 2: Pledge breakdown table by total amount (filtered by USD currency)

* + - Even though the Technology category ranked 3rd. (among nine categories) in terms of total number of campaigns, it received about 49 percent of the total dollar amount that was pledged by all backers (investors).
    - All non-technology categories accounted for less than 10 percent of the total dollar amount pledged, except for Film & video (13.9% of total dollars) and Theater (12.5% of total dollars).

Table 3: Average pledge amount by total backers (filtered by USD currency)

* + - The average pledge amount per backer/investor was less than $100 for all categories, except for Technology (average of $133 per backer).
    - Failed, live and successful campaigns had almost the same average pledge amount of about $92-$93 per backer/investor; in contrast, failed campaigns presented a much higher average pledge amount (about $273 per backer/investor).

Some examples of additional graphs/charts relevant to this analysis are:

Chart 1: Average campaign duration (Column chart)

* + - The average campaign duration had min and max values of 31 days to 35 days, respectively.

Chart 2: Campaigns picked vs. not picked by staff (Pie chart)

* + - The average success campaign rate increased significantly from about 48 percent to 87 percent if it was picked by the Kickstarter staff versus non-picked staff campaigns; also, it is important to note than less than 14 percent of all campaigns were staff picked.
    - Except for Photography, all other categories had a success rate above 90 percent for staff picked campaigns; whereas the success rate for non-staff picked campaigns was below 50 percent across all categories with the exception of Music (75% success rate) and Theater (57% success rate.