Assignment Report

Unit 01 - Excel: Excl-Challenge

Sheri Shojaie, May 16, 2020

* Conclusions about Kickstarter campaigns

1. Between May 2009 and March 2017, Kickstarter campaigns in the categories of theater (~34%) and music (~17%) have been the most common world-wide. Theater and music campaigns have also had the highest success rate at ~60% and ~77%, respectively.
2. Within subcategories of Kickstarter campaigns during this period, plays (theater) account for over ¼ of all campaigns at 1,066 world-wide, with a ~65% rate of success for meeting their fundraising goal. This is more than 4x higher than the next most common sub-category of campaigns, pop (music), at 260.
3. Data looking at the launch timing of campaigns suggests a seasonality trend in campaign success. Analysing historical data of launch months from 2009 to 2017, it seems that campaigns that launched between February and May had the highest rate of success (3-7% higher rate of success than the annual average), and the lowest rate of failure (3-6% lower than the annual average).

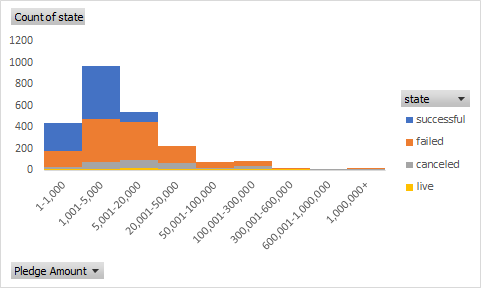
* Dataset limitations

Some limitations of this dataset include:

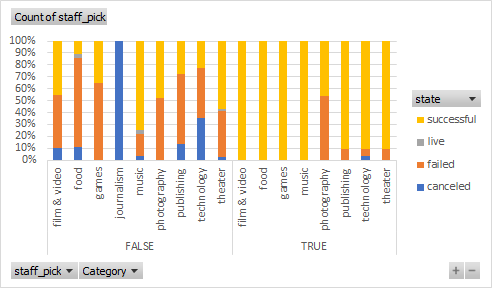
1. A relatively small sample has been provided, as over 300,000 Kickstarter projects have been launched. Is the data on these ~4,000 projects representative of the broader list of Kickstarter campaigns?
2. There seems to be a disproportionate representation of countries. Over 88% of all campaigns in this dataset are based out of the United States (US) and United Kingdom (GB), limiting our ability to confidently analyze the outcome of campaigns in other countries
3. Do not have full representation of all months for the years reported; for example 2017 only includes Jan to Mar
4. There is limited data on project specifics and the people involved in the list of campaigns provided. There may be additional factors that are important to the outcome of campaigns that have not been analyzed here. For example, the level of professional experience of campaign organizers or a track record of launching successful campaigns may play a role in the outcomes observed

* Other possible tables and/or graphs that we could create?

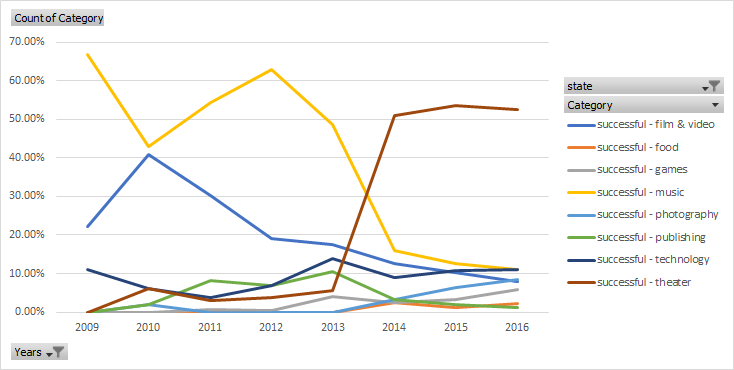
1. Could analyze the size of the campaign’s initial goal and determine whether it correlates with the outcome.

* In a new column categorize the fundraising goal into discrete increments and create overlapping histograms visualizing the campaign outcomes to determine whether there is a trend in campaign size and a successful outcome

1. Do campaigns that have been identified as a “Staff Pick” or given “Spotlight” have a higher likelihood of a positive outcome?

* Create a pivot table that shows how many campaigns were successful, failed, canceled or currently live per category, segmented by whether they were a staff pick or not (True vs False). Visualize trends using a 100% stacked column bar chart.
* Similarly, create a pivot table that shows the above, but instead of staff pick, segment the campaign outcomes by whether they were a Spotlight pick or not (True vs False). Visualize trends using a 100% stacked column bar chart.

1. Could analyze whether there is a trend in category and subcategory success year over year; are certain categories becoming more successful in more recent years? Success can be measured as?
   * + - campaign outcome (state)
       - number of backers
       - average donation by backer

* Using state as the example for campaign success, create a pivot table that shows year along the rows and the share (%) of campaign categories for successful campaigns along the columns. Visualize trends using a line graph to determine whether the count of successful campaigns has changed for the different categories over the years. Exclude 2017, due to limited data. Line chart will quickly show that theatres have more recently contributed to successful campaigns, while this has been dropping for campaigns in the music category.