**excel-challenge**

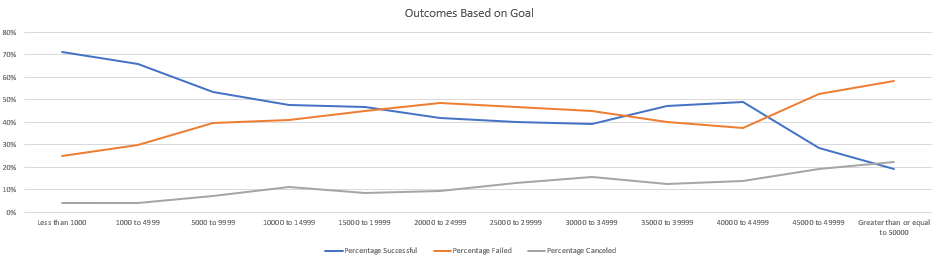
**Questions**

* Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

1. The data suggests there is a high likelihood that consumer psychology plays a role in the Kickstarter business model.

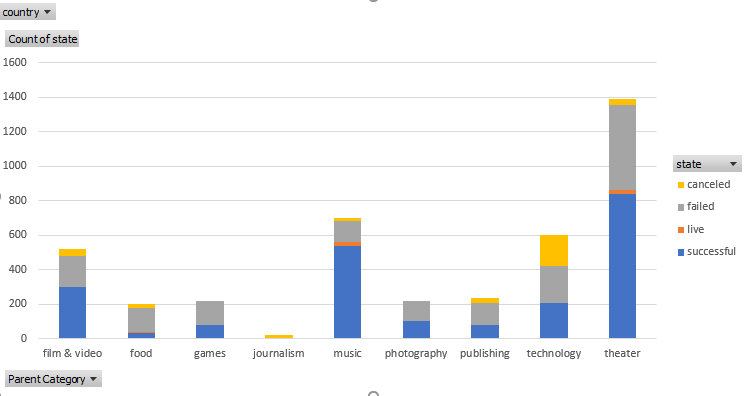
i.a. Community. Everyone wants to be supportive of a good creative cause. The lower the Kickstarter goal, the more likely it is to be meet its goal (see **Figure 1** below). Participation is more likely the lower the cost.

**Figure 1. Relationship between percentage of outcome and goal amount.**



1. There is a possible relationship between project feasibility and excitement with investment. If the backer views the project as attainable then they will risk possibly using their own money to finance the effort. Furthermore, more exciting creative category are more likely to be funded. See Figure 2 below. The categories of film, music and theater have the possibility of having high interest and low start up costs (capital) than other categories like technology (high start up cost), journalism (lower degree of excitement).

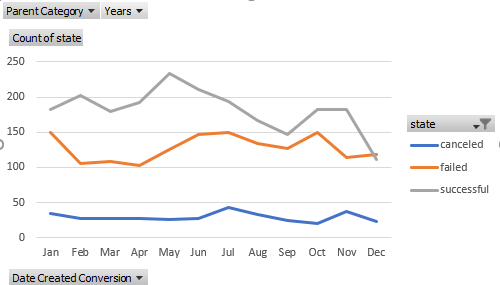
**Figure 2. Stacked column states (canceled, failed, live and successful) for categories of campaigns.**



**Questions (continued)**

1. Annual trends in spending. In most elective retail environments money is spent heavier towards the end of the year (fourth quarter, Q4) due to gift giving and other reasons with less output in the first quarter, Q1. **Figure 3** shows a possible relationship between successful campaigns and launch date.

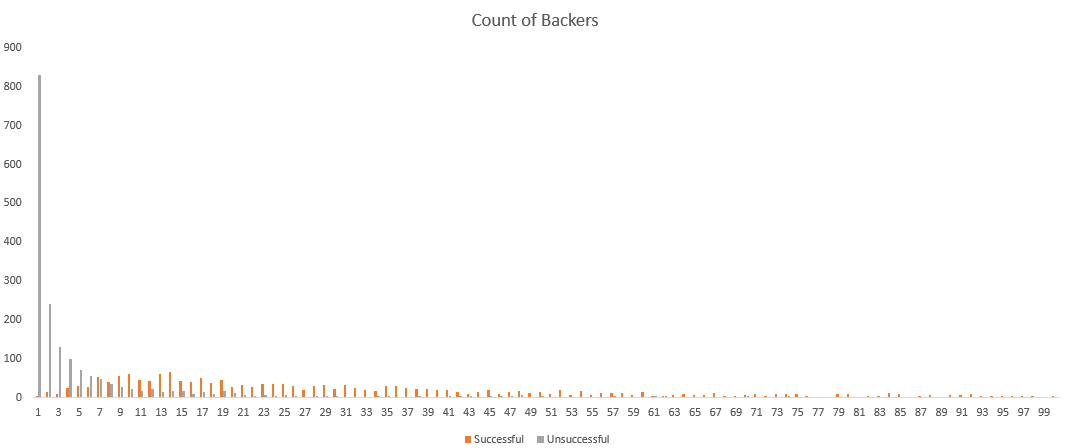
**Figure 3. Launch date of campaign and project state.**



* What are some limitations of this dataset?
* Current trends, the dataset stops in early 2017
* Consumer ‘Backer’ demographics (location, age, race, income, etc.)
* Rewards or benefits given to investors (value or return on investment by the backer upon successful funding)
* Relation of backer to campaign
* What are some other possible tables and/or graphs that we could create?
* Count of Backers by range. For example how many campaigns have (0-50, 50-100 investors, etc.). See **Figure 4** below.

**Questions (continued)**

**Figure 4. Count of Backers by groups of 2 (i.e. 0 – 2, 2 – 4, etc.).**



* Graph that integrates information from a Kickstarter database or data scrapes the Kickstarter website at a specified latency period.

**Statistics Bonus**

* Does the mean or median summarize the data more meaningfully?

Both are useful in determining skewedness. In a perfectly normally distributed dataset, of substantial sample size (n), the median equals the mean. If Mean > Median that means the data is skewed to the right and vice versa. We expect to see most of the data for count of backers to reside around lower numbers with few infrequent campaigns to have blockbuster outlier counts (causing right leaning skewedness).

* Is there more variability with successful or unsuccessful campaigns?

There is more variability, higher standard deviation and variance, with successful campaigns as expected. We would expect the data for unsuccessful campaigns to be grouped together with fewer backers. This is displayed with the mean of successful and unsuccessful campaigns (successful = 194.43, unsuccessful = 19.49) and the graph in **Figure 3**.