#### Kickstarter report requirements:

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

What are some limitations of this dataset?

What are some other possible tables and/or graphs that we could create?

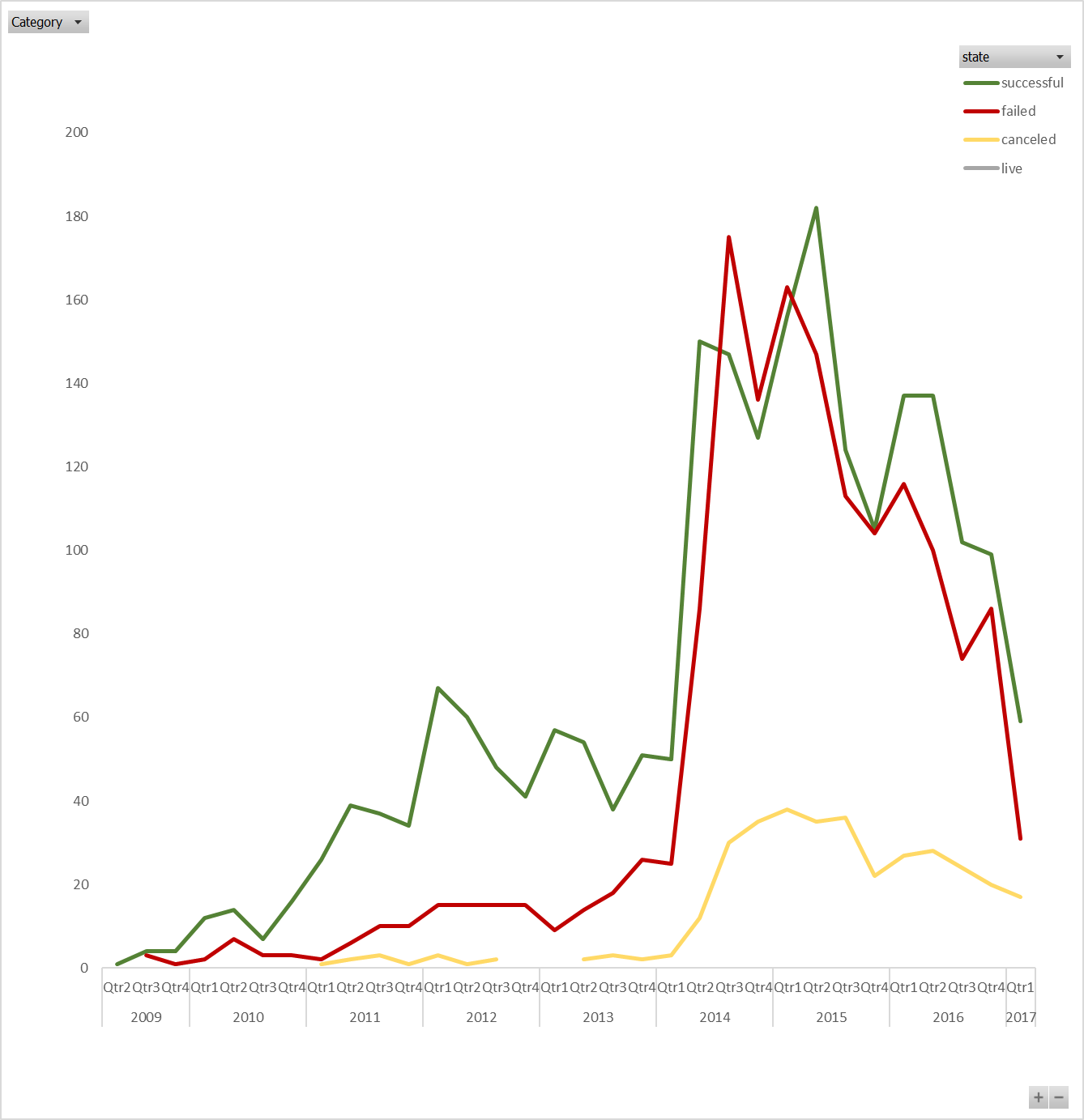
# Kickstarter Analysis based on Study Data Conclusions:

Data clearly indicates that campaigns related to entertainment were more often than not successful at meeting their campaign goals. In particular Music, Theater and Film/Video categories had far greater successes than other categories.

|  |  |
| --- | --- |
| **Row Labels** | **successful** |
| theater | 60.23% |
| music | 77.14% |
| technology | 34.83% |
| film & video | 57.69% |
| publishing | 33.76% |
| games | 36.36% |
| photography | 46.82% |
| food | 17.00% |
| journalism | 0.00% |
| **Grand Total** | **53.11%** |

The timing of the introduction of the projects had little correlation to the outcome of the campaign. If it did one would expect to see successful campaign spikes for campaigns on some quarterly basis. No such relationship exists.

The relationship between failures and successes seem to track in similar relative proportions across the timeline.



# Campaigns with higher monetary goals run a greater risk of not meeting their goals. This stands to reason since it is hard to convince sufficient people to support a project to raise the funds for larger cost projects. The chart below shows the rate of success vs campaign financial goals.

# Further investigation into the blip at $40k - $45k would be of interest. It is likely due to a quirk in the sampling data since the detailed view shows projects across a few categories and in technology as well as entertainment. Previous analysis demonstrated that ‘Entertainment’ campaigns have a high degree of success. If that sample range was more heavily weighted to ‘Entertainment’ that would drive a hypothesis that the category swayed the spike.

# 

## Limitations:

Statistically speaking our sample size represents only 1.3% of the total population of Kickstart campaigns. In addition, this sample is heavily weighted to the US with 3005 of 4064 entries. In contrast Singapore shows 1, Luxemburg and Belgium 2 each. One cannot determine from this data whether that is proportional to the total population of Kickstarter campaigns.   
A full third of the campaigns in the sample fall into “theater” category. That should be compared against the total population to see if our results are skewed.

Statistical analysis was done to look at the quantity of backers per campaign for successful vs failed campaigns. It became obvious that success was more likely with more backers. What we cannot determine from this data is WHY donors chose some campaigns over others.

There are no web marketing, advertising or quantity of ‘hits’ data to see if these factors matter.

There are no fields to show us if there are certain entrepreneurs who have greater success. Also, the entrepreneur’s quantity of positive or negative reviews.

Does expected time to market make a difference. People can be impatient.

## Other possible Tables/Graphs:

Average backers per category and subcategory

Country percentage successful vs per capita wealth – to see if that mattered, I used world bank data on per capita GDP for the countries in the sample. The relationship was inconclusive. Mexica and Hong Kong had zero successful projects but only Mexico is poor and neither had many total campaigns.

Days open vs state to see if longer time to gather pledges had any effect

Goal vs Pledged to identify outliers

# Statistical Analysis

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Backers Per Campaign** | | | |
|  |  |  |  |
|  | **Overall** | **Successful** | **Failed** |
| Mean | 112.6 | 194.4 | 17.7 |
| Median | 25.0 | 62.0 | 4.0 |
| Min | 0 | 1 | 0 |
| Max | 26457 | 26457 | 1293 |
| Variance | 388,861.9 | 713,167.4 | 3775.7 |
| Standard Deviation | 623.6 | 844.5 | 61.4 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Use your data to determine whether the mean or the median summarizes the data more meaningfully. | | | |
|  |  |  |  |
| Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not? | | | |

The data above was taken from the workbook. As you can see both the mean and the median show an order of magnitude difference in the number of backers per campaign. The stark contrast of the Median indicates a severe lack of backers in failed projects since half of them have 4 or fewer backers. The likelihood of being able to raise money is surely limited when there is little to no interest in the project. Even a small goal would be unattainable.

There is far greater variability with the successful campaigns; again, by order(s) of magnitude. The failed projects are limited in terms of quantity of backers so it makes sense that you will not see the variability. Successful projects could raise far more in pledged than needed. For example: ‘3Doodler: The World's First 3D Printing Pen’ raised $2.3M for a $30K project! That’s a whopping 7814% funded!