Problem: To receive funding for a kickstarter campaign, the campaign project must meet or exceed an initial goal. The client is trying to determine the trick to developing a successful Kickstarter campaign based on past campaign successes and failures. One Thousand campaigns were analyzed via the counts of successful, canceled, live and failed projects. Outcomes based on the number of backers, dates created and deadlines, location of project, and categories of the project were analzed.

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

One Thousand Kickstarter projects were analyzed to determine three conclusions on the trends that lead to successful funding. Approximately 14% of this sample were projects that were canceled or still live and were not taken into consideration for some of these factors.

1. The size of the funding goal amount matters.

The sample size contained projects that ranged from $100 to $199,200. The mean value of the total sample size is $43,983 and the median goal amount is $8300 with a standard deviation of $58,962. However, the overall mean of all successful Kickstarter projects is $26,405.31. Highlightins that projects with smaller project goals on average are more successful than those with increasing project goal sizes.

Projects ranging between $15,000 and $35,000 trend more successful than others. The top 50 highest goal projects of the 1000 failed, with only two of those projects succeeding. And out of the top 100 projects with high goal values, approximately 90% of them failed.

1. The categories of projects have a significant affect on what is successful or what fails.

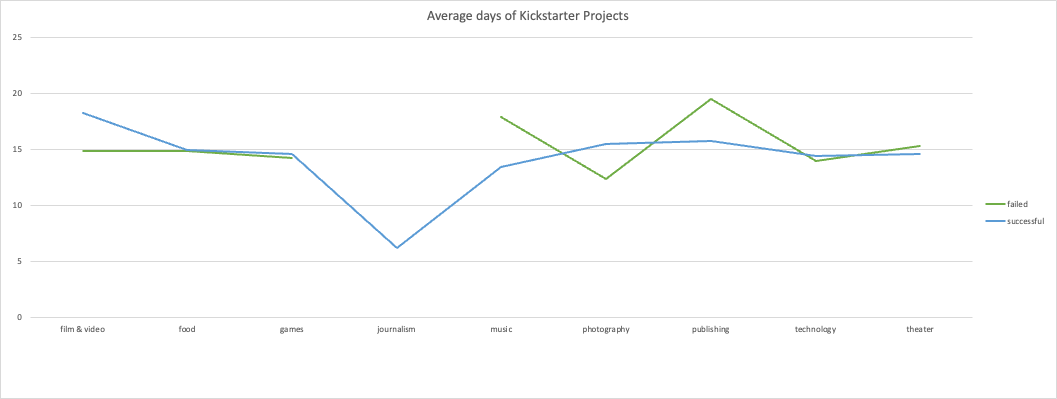
The type of campaign (Parent Category) has an affect on whether it will be successful or not. Theater, film & video and music are popular project categories with more projects being developed in these areas. Journalism (60%), Technology (67%), and Photography (60%) on average have a higher percentage of success. The category ‘Theater’ makes up 34% of the total sample and 54% of all theater kickstarter projects has been successful.

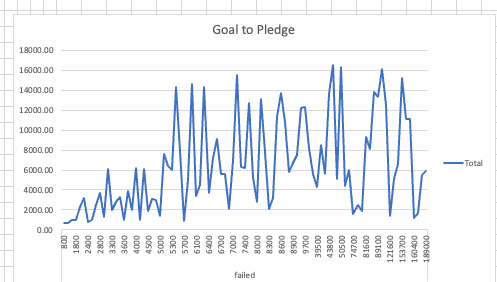
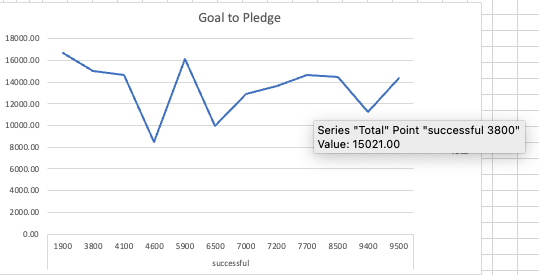
1. Location matters

The majority of the projects in this sample originated in the United States (76%) and 57% of these projects were successful. Projects originating in Great Britain, had a 58% success rate.

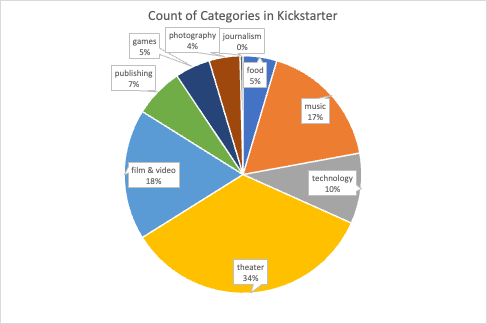
1. Number of Backers

The average number of backers from successful Kickstarter projects is 897 with a standard deviation of 1319. In comparison, the Mean number of backers of failed Kickstarter projects is 585 with a standard deviation of 961. Less backers tend to produce less successful campaigns.

* What are some limitations of this dataset?
  + We do not have all the data that is possible from the Kickstarter data. For example the write ups from the website could be beneficial. If the blurbs were available in the data there would be some benefit to them. Kickstarter campaigns are often judged by the way they are communicated. You can do a lot with the blurbs to check sentiment and word analysis. Using a Python NLTK(Natural Language ToolKit) is a quick way to analyse.
  + Prior to this even a quick sentiment analysis via Grammarly, on a small sample can provide insight to the overall sentiment of the text.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + Length of campaign from the date launched to the deadline could be relevant, but the sample size does not show too much relevancy.On average successful campaigns lasted 14 days, failed campaigns lasted 15 days.
  + Goal of campaign versus pledge amount: This graph would show the amount pledged again the goal. It could show that if the goal is too high, then the chances of reaching the goal may be unsuccessful. It could also show that some successful projects may exceed their goals.



* + Percentage of category- This can give a base understanding of which categories are represented in the data.



* + Name of campaign could be an interesting insight. Is there a pattern with the length of the campaign name.
* A brief and compelling justification of whether the mean or median better summarizes the data (5 points)
  + The median summarizes the data better than the mean because the variance of the numbers is very large. The graph is right skewed. The mean is greater than the median and the mode. The variance is large and measures the average of which each point differs from the mean. A larger variance can help to determine that the mean does not reflect the sample. In addition, since the distribution is skewed, the median is a better indicator for the data.