## Excel Homework: Kickstart My Chart

Kickstarter's mission is to bring creative projects to life through its global crowdfunding service. The service allows creators to have complete control of their work throughout the entire process by eliminating the need of grants, donor or private investor. Although starting a campaign is relatively easy, not all projects come to life. The following analysis is based on data from 4,114 Kickstarter campaigns created between May 2009 and March 2017.

Based on the data provided, the following three conclusions can be drawn about Kickstarter campaigns.

- The sample data provided shows that the Theater Category had the largest number of total Kickstarter campaigns (1,393) as well as the largest number of successful campaigns (839). However, on a percentage basis, Music had the best success rate within their campaigns (77%).

Category	Successful	Live	Failed	Canceled	Grand Total	Percent Successful
theater	839	24	493	37	1393	60%
music	540	20	120	20	700	77%
technology	209		213	178	600	35%
film & video	300		180	40	520	58%
publishing	80		127	30	237	34%
photography	103		117		220	47%
games	80		140		220	36%
food	34	6	140	20	200	17%
journalism				24	24	0%
Grand Total	2185	50	1530	349	4114	

Campaign State by Category
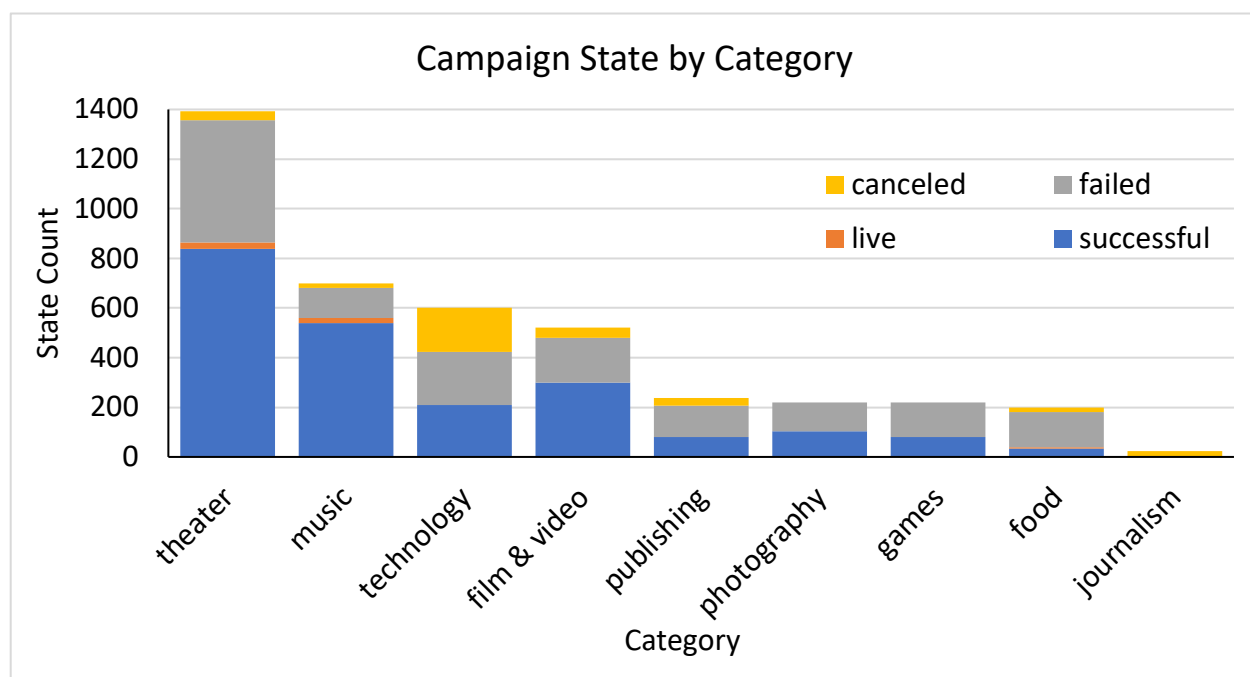
- The sample data set indicates that, on average, campaigns have a 53% chance of being successful. When we analyze the data by the month in which the campaign was launched, those that began in April and May have success rates well above the average (60% and 61% respectively), while those launched in December have the worst success rate, 8% below the average. This conclusion includes the values for campaigns that are labeled as live.
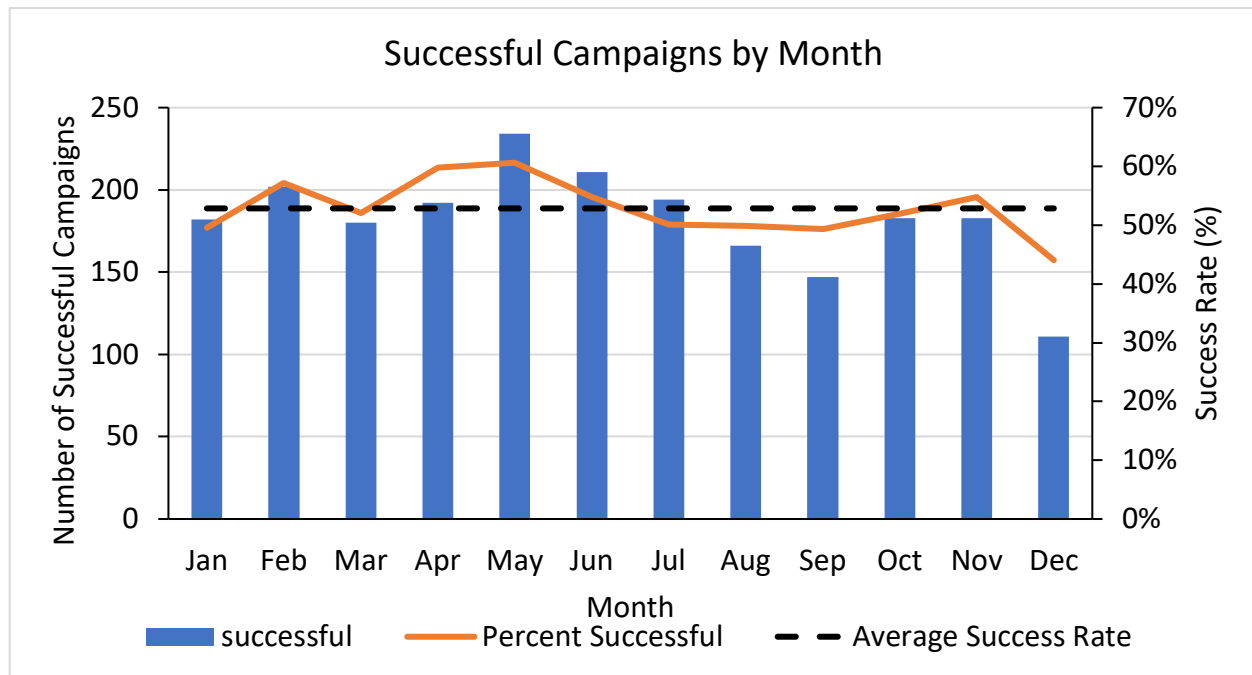
- The larger the goal, the higher likelihood that a campaign fails or is canceled. By converting the line chart from the Bonus work into a stacked line chart, the trend is a bit clearer.
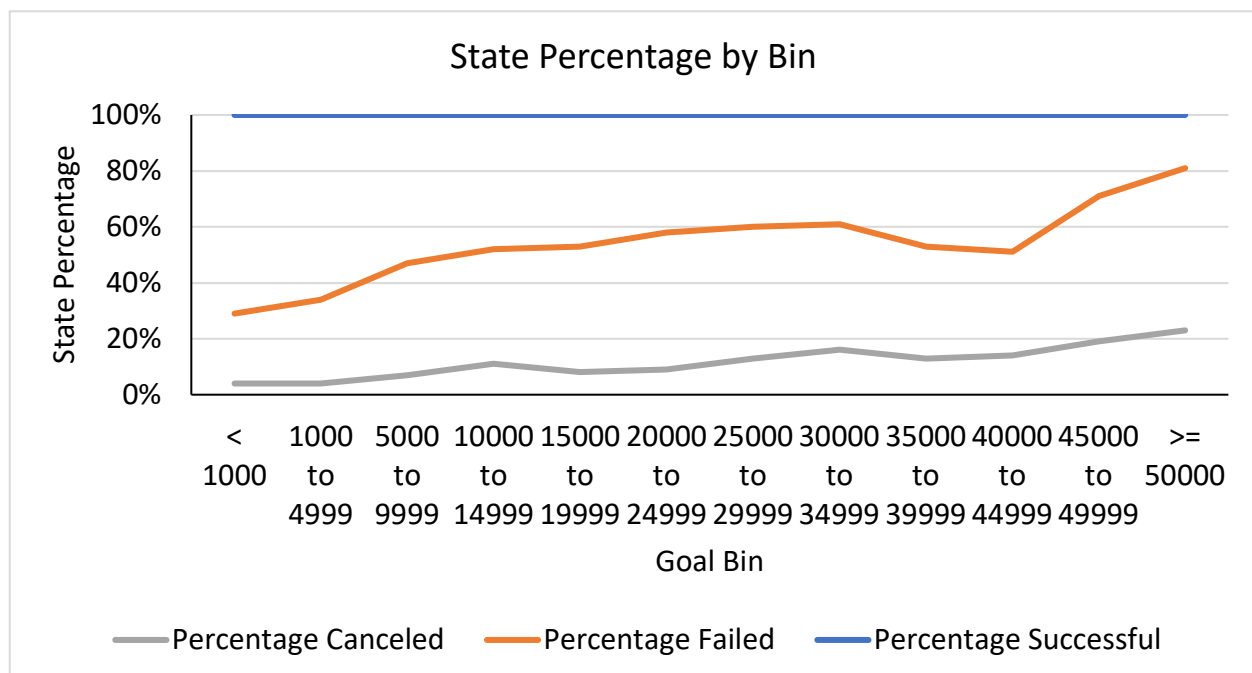
Some limitations of the data set provided include the following:

- The 4,114 campaigns are only a small sample of all of the Kickstarter campaigns ever created, making these conclusions based on only a sample of the population.
- The data set only includes campaigns created through 2017, there may have been a substantial change in the data trends since that time.
- Over 70% of the campaigns within the data set originated in the US, leaving only a small fraction of the sample as a global representation.

Some examples of additional charts and tables that would aid in the analysis of the data are:

- A Pie chart showing the breakout of successful, failed, canceled and live campaigns as a good visual representation of the sample data set.
- A histogram of the number of backers for successful campaigns as well as a similar chart for unsuccessful campaigns. A box and whisker plot would summarize this data concisely as well.
- A Pivot Table of the state of a campaign on the columns and country of origin on the rows.
- A line graph showing how average donations have changed overtime in order to pinpoint specific instances where greater economic factors may have influenced the donation amounts and its impact to a campaign's success or failure.

Additional statistical analysis was done on the number of backers for both successful and failed campaigns in order to determine variability within each data set.

Below is the summary for various statistical calculations:

Statistical Measures of Number of Backers	Successful Campaigns	Failed Campaigns
Mean	194	18
Median	62	4
Minimum	1	0
Maximum	26457	1293
Variance	712841	3773
Standard Deviation	844	61
Range	26456	1293
Campaigns with Less Backers than Mean	1789	1247
Number of Campaigns	2185	1530
Percent of Campaigns Below Mean	82%	82%

Based on the calculated mean and median for both the successful and the failed campaigns, it appears as though the median is a better measure of central tendency because of the variability within the data. In both campaign states, the calculated mean value is greater than 82% of the number of backers for each campaign, while the mean lies in the exact center of the data sets and disregards major outliers on either end.

When looking at the standard deviation calculations for both data sets, there is a larger spread, or more variability, within successful campaigns. The Standard Deviation is a measure of spread of the data around the mean, which is skewed by outliers within the sample. It makes sense that there is more variability with successful campaigns as a campaign can be considered successful by merely meeting their goal or exceeding it by an order of magnitude. Meaning that successful campaigns have a lower bound described by their goal but are unbound on the higher end. However, failed campaigns are bound by a pledge amount of zero up to the goal, limiting the range of the data. For this reason, the number of backers within the successful campaigns has a larger degree of variability.