Joseph Moravitz

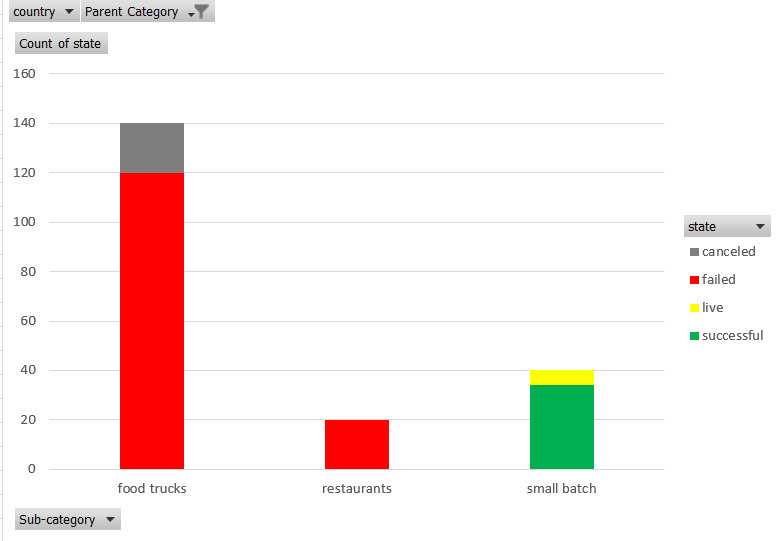
8/23/18

GT Data Bootcamp

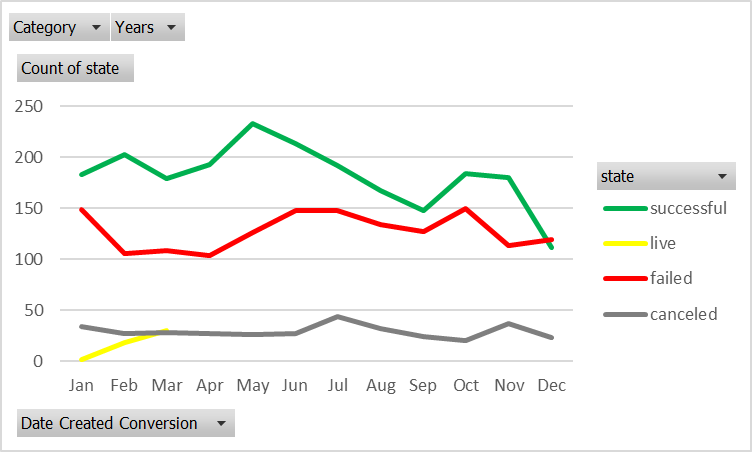
Homework 01 – “Kickstart my Chart”

In this assignment we were given a spreadsheet with data gathered about many projects that were started in Kickstarter. The data included things such as the date it was started, when the deadline was, what the goal in terms of money raised was, how much was actually raised, etc… From this, we were tasked with making simple calculations such as finding what percentage of the goal was met etc… We then continued and made various graphs and tables detailing various relationships between things such as the relationships between what category/subcategory of Kickstarter campaign versus whether or not the campaign succeeded.

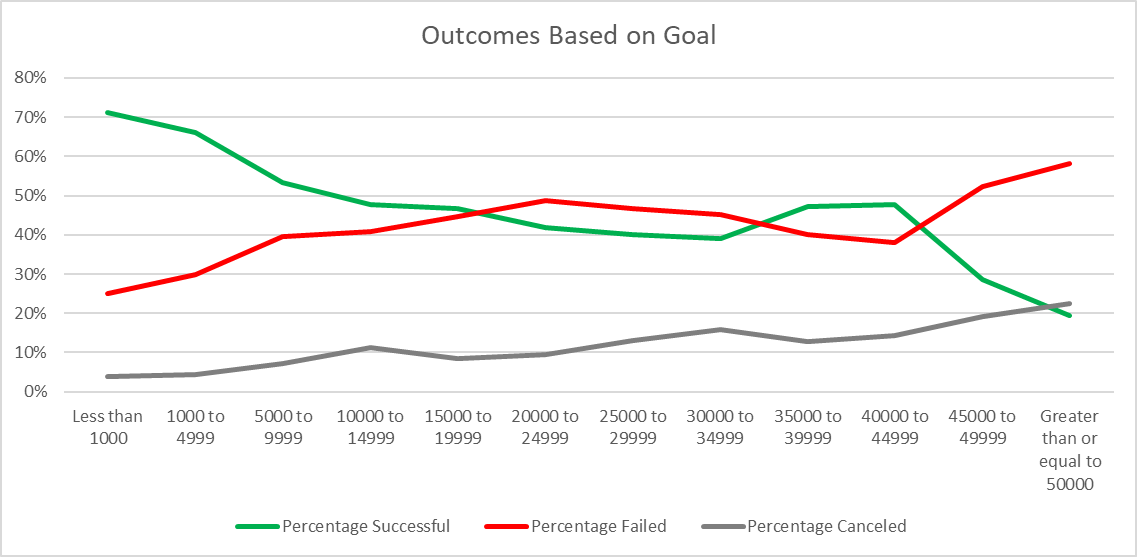
Having played with the data, there are a number of conclusions that we can draw. For one, it appears that it is notoriously difficult to begin a restaurant business, particularly through the use of Kickstarter. Of all 200 Kickstarters in the food category across the entire world, the *only* ones which succeeded were in the subcategory “small batch” which has a perfect track record so far. On the other hand, *every* food truck and restaurant attempt has failed. Perhaps users of Kickstarter aren’t interested in helping support a restaurant half-way across the country where they won’t have any chance of enjoying the food.



Another interesting bit of data is that there are significantly fewer Kickstarter campaigns created during December. Perhaps people are tighter with their time and money as the holidays approach, opting to instead spend them with and for their family or for some rest.



Finally, of particular interest is how the requested goal impacts the success rate. The more money that is requested, the more often the project fails. This shouldn’t be particularly surprising. It is difficult to raise money for a project, regardless of platform and Kickstarter should be no exception.



The data set which we were provided with has a number of limitations. Surely, one important factor in the success of a Kickstarter campaign is in consumer interest and how well it is advertised. The example of Zack Danger Brown asking for a few dollars to make a batch of potato salad going viral and instead raising tens of thousands of dollars is a famous one. One should never underestimate the value of viral trends. The data provided contains no information on any of this apart from the total number of backers. It also contains no information on what return backers are expected to get if helping fund the project. A project which promises that backers will get a copy of the project (e.g. a copy of the video game they are helping fund) should expect to do better than one which provides nothing in return for financial assistance.

We could always do more if we wanted or had time. Using the data provided, I would be interested in also creating charts based on whether the projects received a staff pick or spotlight recommendation (data which we have). I would expect to see that these, being given more exposure, should perform much better on average. This should support the claim that if you want your Kickstarter to have a better chance at succeeding, you should try to make it mimic those projects which received such positive attention (though this begs the question of who we are marketing to, the public, or the Kickstarter staff).

Completing this assignment was not difficult overall, but there was a bit of difficulty in getting my code on the final page to work correctly. In using COUNTIFS(), I was running into difficulty with the program expecting to see quotation marks enclosing the condition while trying to use quotation marks within the condition itself. Once recognizing that it was interpreting the quotation marks as opening and closing a string and chopping things incorrectly I was able to fix the issue. I opted to run with a more complicated expression in COUNTIFS() because although it could easily be brute forced by manually entering each cell with the appropriate ranges, if there were several more or several hundred more lines that we wished to refer to this would no longer be feasible. I wanted a more easily generalizable approach.