Will Copeland

Data Bootcamp

Questions:

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

First off - campaigns are typically less successful in the month of December. In the chart “State per Date Created” we can see that there is a dip in this month. Second – in the “State per Category Chart” we see that theater Kickstarter campaigns are the most common by far (at least within this sample), so any conclusions that we find about the data may be disproportionately related to theater campaigns. And lastly in the same chart we can see that Music Kickstarters have the highest success rate (at least in this sample).

1. What are some limitations of this dataset?

First of all the size of the dataset is limited – if we had a dataset of all kickstarter campaigns ever created we would be able to make a lot more decisive conclusions – but alas – that would be an absolutely massive dataset. Another limitation of this dataset is that it contains some outliers – specifically some campaigns that have over-ambitious or under-ambitious goals. For example some of the campaigns have goals of $1 which is obviously an easy goal to reach so these campaigns disproportionately raise the success rate.

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

Since the dataset includes a column of “staff picks” we could create a success vs. staff pick chart or pivot table to measure how well the staff’s intuition is. We could also see if there was a correlation between project title and blurb lengths and success rate (I.e. do simple/concisely articulated projects succeed more?)