Given the data set presented we can conclude that Kickstarter campaigns in theater, music, and film and video categories chance of being successful. All these industries have an indie subculture to them and with more ways to share an individual’s ideas, sites like SoundCloud and YouTube are shaking up the industries. Content creators do not have to be in an industry hotspot to get a record deal anymore and the data reflects that but does not mean that there is causation. A second conclusion is that campaigns started in May have a significantly higher success rate than any other month, while December is the only month that the campaign is more likely to fail. The third conclusion some sub-categories are significantly better than others. Of the 180 documentary campaigns all were successful while all animation and drama campaigns failed. If you are going to Kickstart fund a game, it needs to be tabletop and not a mobile or video game.

I think that the data set is fairly complete. It is almost to board of information and for a company to use they would need to break it down into one category before being able to apply it to their product or idea. I also think more focus needs to be spent on looking at the year and cultural trends at the time. I feel a kickstarted in a subcategory of fantasy would have had more support between season 7 and 8 of game of thrones because of the cultural hype it created for that genera.

I think a few interesting graphs to incorporate into our data would be to see the success rate of staff picked items and look at the success rate of spotlight items. Campaigns that fall under these two categories are going to get more web traffic which will draw more potential backers. A quick glance at the first 120 cells shows that successful TV shows were all shown to be a spotlight pick while the canceled shows were not. Starting at cells 242 we see that those successful items were both staff picks and spotlight items.

Bonus

I believe the median summarizes the data more meaningfully. The difference between the minimum and max number of backers is significant and gives outliers, where the median gives a more realistic representation. There is more variability with successful campaigns judging by the larger standard deviation.