**Unit 1 | Assignment - KickStart My Chart Analysis (Victor Yamaykin)**

* What are three conclusions we can make about Kickstarter campaigns given the provided data?
* What are some of the limitations of this dataset?
* What are some other possible tables/graphs that we could create?

Conclusions:

1. Theater and plays were by far the most successful category and sub-category across countries and years.
2. There appears to be some seasonality to the success of projects launched in May being one of the months with the highest number of successful projects over the years.
3. There is a negative relationship between goal amount and number of successful projects. Roughly 74% of the outcomes can be explained by a linear regression.

Limitations:

The Kickstarter dataset has a couple limitations. Firstly, how do we define a successful or failed project? Sometimes the product advertised does not get delivered or otherwise takes many months after reaching their goal(s). Secondly, there is a lack of data from social media. We could collect data on which projects had YouTube commercials and the projects’ popularity on Twitter and other social media. There was a team in Switzerland that built a prediction model that uses this social media data and other variables like the rate of donations (<https://www.popsci.com/g00/article/science/tool-accurately-predicts-whether-kickstarter-project-will-bomb>). They claim to have 76% accuracy predicting which projects would succeed within four hours of launch.

Other factors to consider include the types of stretch goals offered and what kind of marketing that each project had before and during their pledge drive. We could take a closer look at the types of advertising outside of spotlight and staff picks.

Other tables:

We have the option with this data to make a comparison using two or more pivot tables. We might use this method to determine which country had more successful projects in a particular category. I found that projects with staff picks and spotlights were more likely to succeed. They had a 12% success rate in this dataset.

We could also gather more data by identifying the genre of each project. For example, we could research the genre of each of the plays and television shows to determine which genres had the greatest number of successful projects. That variable may help to predict whether or not a certain project will meet their goal.