**01 Excel / Homework 1**

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
2. Music projects are the most successful type of campaigns.
3. Plays are the most popular type of project promoted on Kickstarter.
4. Projects started in December have the lowest probability of success.
5. What are some limitations of this dataset?

The dataset doesn’t give us any information on the reward systems/tiers made available to backers; these rewards act as incentives to backers. We have no information about how this data was acquired from Kickstarter: is it truly a random selection? We also have no background information on the companies or project managers: what experiences/resources did they have available that may have influenced their success?

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

We could make comparisons on the effectiveness of Kickstarter support: did “Staff Pick” or “Spotlight” lead to more success? (Count of state for neither staff pick nor spotlight vs staff pick vs spotlight vs both staff pick and spotlight).

We could look into how the length of the campaign influences success (deadline - date created = campaign length). Subcategorize campaign length then compare success rates.

We could make a table to look at how the size of the goal amount compared to success.

Bonus Statistical Analysis

1. In these cases the median is more telling of where the “center” of the data is. Since backer counts vary significantly from 0 to over 25000, if the large numbers are outliers, they will skew the mean much more than the median. In these cases, most of the backer counts for the successful group were under 100, for the failed group, under 10, but the means are large due to large outliers. Median is more meaningful.
2. There is more variability in the “Successful” data set as demonstrated by its larger standard deviation (844 as compared to the “failed” of 61). This make sense because in general we should see the failed campaigns backer count hovering closer to 0 backers (a complete failure) and the successful campaigns stretching from the success of small campaigns (1 backer) towards the largest count.