Kickstarter Excel Project

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

* Theater category seems to be the most prevalent and successful category
* May, June and July seem to be the most active months for campaigns
* Goals less than $10k have a better than average probability to be successful

Based on the data given if one would want a successful campaign via Kickstarter, having a theater project for less than $10k during the late spring would give you the best odds of success.

# What are some limitations of this dataset?

Theater, music and film/video have the highest success rates of all categories with theater having as many campaigns as music, film & video combined. The introduction mentions that only a third of the projects reach a positive outcome but within the data set given, the success rate is over 50%. This is a big difference and maybe we would need a larger sample. This dataset is heavily weighted with the theater category but don’t know if that is the case for the whole population.

Other data that could be helpful is the type of backer and why they tend to favor the arts category.

* Small sample
* Theater heavily weighted in this sample
* What type of backers (bias towards the arts?)

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

* Campaigns per country to see the weight of each country and if outcomes follow pattern
* Success trends per category per year. Are there any trends we see over time
* Outcomes per average donation
* Percentage of total projects by goal

# Bonus:

## Use your data to determine whether the mean or the median summarizes the data more meaningfully.

Based on the data I would say that the mean summarizes the data more meaningfully. The reason for this determination is that there are more than 4 times the number of backers over the median than there are below it, therefore closer to the mean.

## Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

According to the statistical metrics the successful campaigns have more variability. This makes sense to me as successful campaigns obviously need backers and there would be a variety of goals and projects.

With the failed campaigns, there are large numbers without any backers and the maximum amount of backers for a campaign is rather low compared to the successful campaigns. So in essence we have less of a sample for failed campaigns giving it less probability of variance.