**Kickstarter Data: Homework #1**

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Theater had the most successful campaigns than any other category.
   2. The more backers supported campaigns, the higher the probability the campaign was successful.
   3. May had the highest rate of successful campaigns.
2. What are some limitations of this dataset?
   1. This dataset can be subjective based on the viewers and location. Natural biases can cause a viewer to favor a genre over another.
   2. I would love to get more a clear definition of what it means for a campaign to be canceled vs failed.
   3. Huge biases that favors the US. The US is a very popular country therefore the US voice can easily persuade other viewers to watch shows coming from the US. Majority of the campaigns came out of the US as well.
3. What are some other possible tables and/or graphs that we could create?
   1. I would like to compare successful and failed percentages by country and genre to see if countries prefer a genre over another.
   2. Currency – The value of money is weight differently between countries.
   3. How the fiscal season work based on countries. Most non-profits budgets conclude in June. I wonder if this is the reason why May had the highest number of successful campaigns.

**Bonus Questions**

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.
   1. Median - Successful Data: The median is a be way to summarize successful campaigns due to large amounts of backers. When you take into account outliers that may lie in a large dataset, the median would be meaningful to get more information regarding successful campaigns.
   2. Mean – Failed Data: The mean would summarize the data more meaningfully because it has a much smaller dataset. From here you could gather the less backers’ campaigns have, the greater the chance it will failed.

In summary, the more backers you have, the greater the amount of support you would receive.

1. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
   1. Successful campaigns have more variability while failed campaigns has less variability. This makes sense because the successful campaigns had a very large data set versus the failed campaigns. Since the mean for failed campaigns is 18, this tells you that the number of backers in each campaign is fairly low.