**Georgia Institute of Technology – Data Science and Analytics Boot Camp**

**Excel Homework 1: Kickstart my Chart with MS Excel**

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

* The number of successful Kickstarter campaigns accounts for 53% of the overall number of campaigns, according to the Pivot\_Table1 sheet. By far the most successful and well-supported project is theater/plays. Their numbers surpass all other categories in “Category” and “Sub-Category” analyses.
* According to Pivot\_Table2, the subcategory “play” has the most effective campaigns (694).
* According to Pivot\_Table3, the months of May, June, and July had the most successful as well as total number of campaigns out of a total of 4064 campaigns spaced out over ten years.

1. **What are some limitations of this dataset?**

* Because none of these elements can be accurately analyzed, it is unclear whether the company's planning, aim, pledged endeavor, project design, or marketing could have improved the success rate.
* A comment section on each project may have provided more information about how to forecast success or whether it is statistically significant.
* There could have been a link between how long campaigns took to achieve their objectives/goals and how long they took to fail.
* There is no way to predict only category and only sub-category outcomes.

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

* Examine the average donation by category to assist people in determining what their goals should be.
* Create a graph for each country based on successful campaigns and draw inferences to determine which country's campaigns are most effective.
* Analyze the growth of Kickstarter campaigns by category over time.
* It would be interesting to look at the results of Kickstarter campaigns in specific locations (filtered by country, state, and location), which could then be evaluated geographically.

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

* The mean of successful v/s failed campaigns summarized the data more meaningfully compared to median.

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

* Variance is a measure of heterogeneity in each data. Higher the variance, more heterogeneous it is and smaller the variance, more homogeneous it is.
* As the variance of backers with successful campaigns is higher than failed, we can conclude that successful campaigns has more variability.