**KICKSTARTER REPORT**

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

From the analyses of the 4,000 projects (raw data), in addition to the pivot table and pivot charts we could conclude the following:

* If we take a look at the first pivot table/chart which analyses the correlation between the different categories and the state, we can see that music and theater have the highest successful rate which account for 25% and 38% respectively for successful projects. The Music and theater categories sum up to half of the total projects. We also notice that the US has the highest numbers of launched projects in comparison to other countries and Singapore has the lowest number of projects.
* The second conclusion we can infer when we look at the 2nd pivot table/chart which analyses the correlation between the state and sub-category is that plays have the highest numbers in terms of success and makes 23% of the data. If we take a further look at the sub-categories, we notice that at least 4 sub-categories in music such as electronic music, pop, rock, are all successful.
* In pivot table/chart 3 which analyses the correlation between the state and the dates of the project, we can observe that there is an increase in successful projects from the months of May to July. This period seems to be favorable to the launching of new project, while December has more failed projects in comparison to the number of successful ones.

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

* From the dataset provided, we notice there is no information about the creators of the projects. The experience of a particular creator could have a proportionate impact on the outcome of a project. Creators with more experience might have more chances to reach their goals.
* The dataset provided for this analysis might not give an accurate description of the Kickstarter. In the background information provided for this homework we are told only a third of all Kickstarter have made it through the funding process with a positive outcome, but we see in our analyses that there are more successful campaigns in comparison to failed or canceled ones.
* Also, we notice that the range of goals differ exponentially from each other in the study. Including campaigns with vey high goals and campaigns with smaller goals might not give us an accurate analysis as it might be easier to reach a smaller goal faster than it would be for a bigger one.
* Additionally, the data does not provide information about backers’ preferences. Certain categories might have an established and solid fanbase, thus making it easy for the campaign to be successful. If we look at theater as a case study here the reason for its success might be attributed to the establish trust the backers have for this Category.

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

* A graph that analyses the correlation between the date created and the date completed for each project would help us understand the probability for a project to be successful or fail depending on the time frame. Projects with a longer durability might have more chances to attain their goals than projects with a shorter duration.
* In addition, a graph that analyses the funding goals based on sub-categories categories would provide a more detailed analyses on why a sub-category like “Plays” is more successful in comparison to others.
* A graph that analyses the different subcategories in relation to the different timelines and help us understand why some subcategories are more successful than others. Comparing each sub-category to a particular timeline would help us understand what season works best for certain sub-categories.
* Graph and display that analyses the correlation between the most successful and failed projects; this can be further structured into the different categories, sub-categories with an analysis of the average success or failure rate for each category.
* A graph of the dataset without the most successful category; this would help us see the success rate of other categories with that category is out of the equation.
* In addition, a graph that analyses the correlation between each country and the sum pledged to see which country has the highest amount of pledges, we can further analyze this in correlation to which category has the highest amount of pledges from each country.

**Statistical Analyses**

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

* If we look at the statistical analysis, we observe that the mean summarizes the data more meaningfully.

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

* There is more variability with successful campaigns than unsuccessful campaigns as we can see in our statistical analysis, the variance shows us that each point differs from the mean and the standard deviation show us how spread the data is from the mean.