Written Analysis Kickstarter Projects (Excel Homework – 01)

1. Given the data set, a few points stick out from initial review. First, campaigns under the “theater” category clearly have the most started campaigns, and within this category, “plays” are far and away the largest subcategory of all, with nearly 4 times the amount of campaigns started than the closest competitor. Next, although campaigns from the “theater” category are the most in overall volume, a greater percentage of campaigns in the “music” category end up being successful; campaigns from within this category having the greatest chance of success from all the categories. And finally, when looking at all Kickstarter campaigns within this data set over time, the chances of success, failure, and being cancelled are fairly static over the period of the year, with the primary exception being that the possibility of success drops drastically from November to December to close out the year, where it is actually more likely for campaigns started in December to fail than succeed, with the opposite being true for the rest of the year.
2. Some of the limitations from the data set and making predictions based on it are that the data does not have any indicator or measure of the overall state of economic measures showing of the general health of the overall economy. Since Kickstarter is essentially asking for donations from the general public, some indication of general economic health would be useful to further evaluate the categories, subcategories and relative success of each campaign and type. Another limitation is that, while this is a relatively significant amount of data, it is still showing small numbers in relation to the general population, as well as having less than 10 years of data overall. This limits the ability to analyze the historical data and to attempt to make statements about the future or make statements predicting the behavior of the full population.
3. There are several other tables that we could use to organize and visualize the data. We could make a scatter plot and attempt to fit a regression of the number of overall backers over time in order to see trends over time of the amount of people donating to the campaigns. We could do the same thing analyzing the amount of money overall donated to these campaigns over time. We have analysis of the amount of campaigns that were successful within each category and with the bonus challenge, we analyzed any effect of goal amount on the probability of success, but it would also be useful to analyze if different categories were more likely to have a particular goal level. This would enable us to see if there were any trends within the categories regarding goal level and if the goal or the category had a greater relation to whether or not the campaign would be successful.