**EXCEL CHALLENGE CONCLUSIONS – UCI Data Science HW #1 – Kevin Avery**

1. Given the provided data, three of the conclusions that I can draw about Kickstarter campaigns are as follows:
   1. The best months to launch a Kickstarter campaign in order to give yourself the best chance of success are between the months of February and June. Since the year 2009, any campaign that was launched during those 5 months has had a 58.5% chance of success. More specifically, if you launch in May, there has been a success rate of 60.6% which is the highest of any specific month. On the flipside of that fact, campaigns that launch between the months of July and December have only had a 50% chance of success. Lastly, December proved to be the worst month to launch a campaign as it had the lowest success rate of any specific month at 44%.
   2. With regards to the kind of projects that have success, music and theater projects proved to have the most success over all other categories. The success rate for music projects was 77.14% and the success rate for theater projects was 60.23%. This means that over 77% of the music projects and 60% of the theater projects that were launched found success as opposed to failing or being cancelled. Conversely to that fact, the worst kind of projects proved to be journalism and food. There has still yet to be a successful journalism project as it currently has a success rate of 0%. While food projects had the second worst success rate of only 17% of all projects launched.
   3. The third conclusion that I came to was that your chances for success on Kickstarter start to go down significantly if your project goal requires a large sum of money. Projects that asked for less than $5000 to complete their goal had a combined success rate of 69%. While projects that asked for more than $45,000 had a combined success rate of only 24%.
2. The limitations that I found with this dataset are as follows:
   * 1. It is my understanding that a major part of getting donations on Kickstarter campaigns is based on the marketing of that campaign. The more people that see the campaign…the more donations that are usually given. So, one thing that is unclear by looking just at this dataset is how this affects the success and failure rates of the campaigns. How many “online views” did the successful campaigns have versus the failed or canceled campaigns? Did the failed campaigns fail because of the amount of money that they were asking for? Or did they just fail because they simply weren’t shared online enough and they didn’t get enough people to see what the project was all about.
     2. Another limitation of the dataset is that it is hard to put a quantitative value on how “good” the projects actually are. It is more likely that the success rate of a project is based more on how “good” the project is as opposed to what month it is launched in, how much money they are asking for, or even what the category is. For example, I am sure a musical band that already has an online presence and following would have a high chance of success if they were asking for a donation to make a music video. On the flipside if some random musical band that was less talented and didn’t have a following was asking for money to go on the road and take a trip, I am sure they would have a lower chance of success. However, in both of these examples it is based more on talent and how good the idea is, and less about the measurable quantitative data that our original data set provides. Even the details provided in the “blurb” section of the dataset would be hard to quantify into any kind of chart or table that we could analyze.
     3. A third limitation of the dataset is that it doesn’t tell us the exact reason why certain projects were canceled. It is possible that some of the canceled projects were actually canceled based on something that had nothing to do with donations or money. They may have even been well on their way to obtaining the necessary goal donations but then were canceled for some other reason. This would essentially lower the success rate of a specific category in the dataset and it could possibly skew the data to be misleading.
3. The other tables and/or graphs that we could create to help analyze the data would be as follows:
   * 1. One chart that we could create would be to compare the amount of “average donation” to how successful each project is. For example, do projects with higher average donations have a higher chance of succeeding? Or is the better path to success through a lot more of smaller or medium sized donations?
     2. Another chart that would help is actually one that I created separately from this worksheet…and it was to measure how successful each project was based on its “category” as opposed to the goal donation level. On the bonus assignment we compared the success rate to the amount of money that was required to meet the goal. However, I also though it was helpful to see what the success rate was based on the type of project that was being funded.
     3. One more chart that would be helpful is to compare how many projects found “success” based on how long their donation time period was. For example, is it more likely for a project to be successful if it has a longer period of time before the donation deadline? Or did some of the projects “fail” because they didn’t have a long enough time to obtain donations?
4. Bonus Statistical Analysis Question – Is there more variance with successful or unsuccessful campaigns?
   1. There is a much larger variance with campaigns that were successful as opposed to unsuccessful. The reason for this is that the path to success on any given campaign can find many different variations. Some of them are successful with a very large number of backers, some were based on really large donations, some were based on a combination of both. However, the path to a failed campaign had a much more common route. The failed campaigns were usually a product of very few backers and little to no donations…Meaning that most of the “failed” campaigns looked very similar to each other which caused the variance to be lower.