**Joshua Marcus**

**Rutgers Data Science Bootcamp**

**Excel Homework**

**April 27, 2019**

**KickStarter My Chart Report**

**Abstract:**

Kickstarter is an internet funding platform for creative projects based out of Brooklyn, New York. The objective of this project is to organize and analyze the Kickstarter database of over four thousand past projects in order to uncover hidden trends in the successful projects.

**Introduction:**

Crowdfunding is the practice of funding a project or venture by usually raising small amounts of money from a large number of people via the internet. Kickstarter is a crowdfunding website similar to GoFundMe and INDIEGOGO, except that it is an all or nothing. In order to launch a Kickstarter project, you enter some basic information then start you campaign, determine your minimum funding goal and deadline. An import breakdown to understand is that each launch on Kickstarter falls into four possible states successful, failed, cancelled, or live upon the deadline.

Over two billion dollars have been raised using Kickstarter. However not every project has found success. Of the over 300,000 projects launched on Kickstarter, only a third have made it through the funding process with a positive outcome. The data analyzed falls into nine (9) categories and forty-one (41) sub-categories. Of the projects launched in the data set 53.11% were successful in reaching their funding goal.

**Three conclusions about Kickstarter campaigns given the provided data?**

1. Kickstarter campaigns with a contribution goal of >= $50,000 have the highest failure rate, 58.11% and highest canceled rate, 22.52%.
2. Kickstarter campaigns with a contribution goal of < $1,000 have the highest success rate, 71.08%.
3. The music category has the highest success rate, 77%.
   1. Per the music sub-categories, 6 out of the 9 music sub-categories have a 100% success rate; pop, metal, rock, classical music, and electronic music.

**What are some of the limitations of this dataset?**

1. The dataset is a sample population (i.e. 4,114 campaigns) and may not accurately represent the over 300,000 kickstarter campaigns. The dataset is only a 1.37% sample of 300,000 campaigns (4,114 / 300,000 = .0137).
2. 73.85% of the kickstarter campaigns in the dataset occurred in the US. The dataset may not give an accurate depiction of the international market. For example, the most successful category (i.e. music) maybe skewed to US inclinations.
3. The anonymity of backers. It’d be helpful to have the name/company and/or occupation of the backer. The identifying information may help identify a trend of backers who support a specific category. The fundraiser can exploit the trend for target marketing.

**What are some other possible tables/graphs that we could create?**

1. Compare the days campaigned to the success rate of a campaign. Put the days campaigned into intervals of 10 (i.e. 1 – 10, 11 – 20, 21 – 30…), and then filter for only the successful campaigns.
   1. Create a line chart that graphs the relationship between the days campaigned and the number of successful campaigns within the day interval. (hint; successful campaigns are = 100% and the intervals are a % of the successful campaigns.)
2. Determine if there’s a correlation between blurb character length and the success rate of a campaign. Compare the blurb character lengths (i.e. =len...) of each state; "successful," "failed," "cancelled," or are currently "live.”
   1. Table and clustered column graph of the average blurb character length per state.
3. Table and clustered column graph of the average number of backers per successful and failed projects.