**Report:**

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

**Conclusion1:**

When we analyze the categories of Kick Starter through their states, we can surely conclude that most successful kick starter belongs to Entertainment industry “Music”, “Theater” and “film industry”. These are the one which has highest number of successful state. When analyzed over the years, music industry is the only industry which has lowest number of failed/canceled kick starter.  
Further if we analyze Country wise, biggest contribution of Entertainment Industry belongs to the US.

If we sub categories the Music industry, Rock Music has highest Successful count.

Using the above statistical measures of dataset, we can conclude that if the Kick Start belongs to Rock Music in Music industry, it has the highest chances to get successful.

**Conclusion2:**

Analyzing the state count of the each category, we can see Journalism is one type of kick starter which has Maximum Canceling count. So having a kick starter related to journalism will not have many takers.

**Conclusion3:**

If we further dive into sub-categories other then music industry, there are few subcategories which have been always successful. For example: Under theater, plays are always successful. In Film and Industry, we have documentary sub-category which would be successful for sure. In games, tabletop games are at the top and finally in Publishing we have non-fiction, radio and podcast.

**Conclusion 4:**

Overall, we have maximum number of entries successful in months of May, June and July.

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

For analyzing, each data in dataset has to be consistence, but the column currency is not consistence in our dataset. The column “currency” has multiple values like USD, GBP and NSD etc. This means analyzing the columns like goal, pledged and average donation will have some unexpected patterns in the charts and tables.

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

We could analyse the data by creating the tables and charts relating to columns Average-Donation, Backers count and percent funded.

Using average-Donation data, we could analyse what kind of kick-starters attract big money from backers and are successful.

Percent funded can be used to analyse the data of successful kick-starters have highest rate of getting funded.

**Bonus Statistical Analysis:**

1. Mean of successful kick-starter is 190 and for unsuccessful kick-starter it is 19. This indicates to make your kick-starter successful you need on an average 190 count of backers put money on your kick-starter campaign.
2. Now if we Analyse the Mean and standard deviation successful and unsuccessful kick-starter backers count dataset, the distribution for both the datasets is positively skewed distribution, which means their mean is greater than median or mode, which we can see in calculations also. Therefore to understand the variability in positively skewed dataset, we could use median as the centre of the dataset and further use the IQR as measure of variability. The IQR for the Successful kick-starter is 108 and IQR for unsuccessful ones is 18, this means there is more variability in number of backers in successful kick-starter.  
   One of the reasons for more variability is number of successful cases is more as compared to the unsuccessful ones. Also, Most of the unsuccessful cases have number of backers count equal to 0-1.