**Create a report in Microsoft Word and answer the following questions.**

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Smaller goals seem more likely to be more “successful” than bigger goals. For example the successful rate of goals less than $1000 tend to was about 75% on average compared to a successful rate of just 19% on average for goals greater than or equal to $50000.
   2. The number of projects canceled stay pretty constant over the year, while success and failure seem to vary with the season. It seems that there’s more success during the spring than say the winter.
   3. Theater, music and technology are the top three categories of fundraiser projects on the Kickstarter platform, while journalism has the fewest projects.
2. What are some limitations of this dataset?
   1. The data seems to have significant outliers making it challenging to draw meaningful conclusions about the measures of central tendencies.
   2. In addition, we don’t know how much resources - in terms of time and incentives - creators of a project spend in order to reach their funding goals. While the time a project was launched and the deadline to reach its funding goal could be proxies for time spent on raising funds, they don’t capture the actual effort put in reaching backers.
   3. It is hard to tell from the data how close some of the failed goals/projects were to being successful and vice versa.
   4. It’s hard to make any predictions about whether a project will be successful, a failure or canceled.
3. What are some other possible tables and/or graphs that we could create?

Other possible charts that could have been created are weighted or 100% stacked bar graphs, pie charts for primary categories for example, scatter diagrams for example showing the relationship between the number of backers and the size of pledges or rate of success or failure of a project, or more generally the relationship between each state of a goal/projects and key factors such as month, year, location, etc. We could also create a graph/cart that shows the distribution of key variables such as number of backers to illustrate whether the data is normally distributed or skewed.

**Bonus Statistical Analysis**

* The median represents the data more meaningfully than the mean because the data contains significant outliers that skew the mean in a significant way. The fact that the mean and the median are significantly different from each other may imply that the distribution of the data is skewed.
* According to the variance and standard deviation measures, there is more variability in successful campaigns than unsuccessful campaigns. This makes sense because - even though percentage outliers are similar in both successful and failed campaigns, namely, 11% to 13% respectively – a campaign is likely to be more successful the larger the number of backers. Given that projects with larger goals may require a larger number of backers, some specific individual projects are likely to receive a large number of backers and therefore skewing the distribution of the data.