1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. Stage arts, music and theater, are the Kickstarter categories with the most total campaigns with a high ratio of success in comparison to other categories. Technology ranks third for total number of campaigns and is fairly even between successful and failed campaigns.
   2. About 53 percent of the campaigns in the dataset were successful but half of all sub-categories did not have any successful campaigns. Initial analysis does not indicate any correlation between the sub-categories with and without successful campaigns. For example, rock music campaigns are were 100 percent successful but jazz music campaigns 100 percent failed.
   3. Using the Launch Date Outcomes graph, the majority of successful Kickstarter campaigns began in May. Filtering each year individually, the time of year is much more variable indicating more a correlation between May and successful campaigns than a causation.
2. What are some limitations of this dataset?
   1. If you wanted to analyze how to make a successful Kickstarter, this dataset can’t help you. Is marketing, disorganization, lack of time/resources the reason for failure? It’s unclear.
   2. The currencies are noted in the table but because they are not all converted to be the same makes it unclear how people of different countries respond to each campaign.
3. What are some other possible tables and/or graphs that we could create?
   1. Successful/failed campaigns by Staff pick using country as a filter.
   2. Successful/failed campaigns by goal

Bonus Statistical Analysis:

**Mean vs. Median:**

For this data set, the mean is more useful than the median to summarize the data. The data is graduated, by sorting the data in descending order there are no huge gaps or jumps. The median vastly under represents the potential for a successful campaign.

The successful campaigns column has two outliers (two campaigns received 20,000 backers and the next highest number of backers in 8000), however, when leaving the outliers out of the equation, the mean is very similar (173 versus 194), indicating that the mean is still an appropriate summary of the data.

**Variability:**

The variability for the successful campaigns is greater. The span of the data is from 1-26,000 versus the unsuccessful campaigns which has a span of 1500-0.

An additional check for the interquartile ranges, the range for the successful campaigns is 112 and the range for the unsuccessful campaigns is 11. Successful campaigns have a larger variability.