Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

* Given the data provided, the three conclusions that can be drawn from the Kickstart campaigns are the following: First, Kickstarters related to entertainment seem to have more of a success rate than the other categories. Specifically looking at Film & Video, Music, and Theater, these three categories started off great and hit their highest success count in May but had a major decrease in December. Second, All Journalism Kickstarters were cancelled and only appeared in the US and Germany. The US had plenty more cancellations within this category than Germany, with the US having 23 cancellations and Germany only having one. Also, Journalism had more frequent cancellations between 2015 and 2016, this tells us that Journalism may not be the best type of creative project to launch. As per the dataset, it has no rate of success. Lastly, 2015 had the most successful Kickstarters from 2009 through 2017, with theater having the highest success rate of all categories.

What are some limitations of this dataset?

* Although this data set looks at various states of Kickstarter projects, the state with the lowest grand total is those that are live. I believe we can take a deeper look at these projects and analyze their data to determine what their outcome will be.

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

* Another graph that can be used in this data set is a scatter plot, as this can tell us if two items have a relationship and/or correlation. This can be very useful because it could potentially tell us why 2015 had the most success rates with theater, and/or why Journalism did not have a single successful kickstarter.

Bonus Questions: \*Reference Summary Statistics Table Sheet\*

Use your data to determine whether the mean or the median summarizes the data more meaningfully.

* When looking at the Successful Projects bar chart, the median best summaries the data more meaningfully, due to the data being more evenly spread out. As the chart seems to be slightly more symmetrical, calculating all values in the dataset can give us more of an insight to what we are looking for. However, looking at the Failed Projects bar chart, the dataset seems to skew more to the left. This can result in the mean being dragged in the direction of the skew, therefore the median can be considered as the best representative of the dataset, rather than mean, as it is in the most central location of the data. In conclusion, both the mean and median can be used in both datasets and bar charts as they are both similar with not much of a skew. However, if the dataset is more centralized it is best to use the mean for a better significance. If the dataset is skewed, it is best to use the median as the it will be the more centralized value and the mean will be further apart from the median, the bigger the skew.

Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

* Given the data, the successful campaigns had more of a variability. This makes sense because Kickstarter had a lot more backers in successful campaigns. However, this also tells us that there are more and larger dissimilarities among the backers, which can result in more outliers. This can have great affects to the mean, as it can skew more towards the outliers.