

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

Although there are many ways to examine the business outlooks, analyzing the data and sorting the information is the first and most fundamental step. One of the lessons from analyzing this campaign is to always keep tracks of categorical information along with the numerical data. This way, we are able to not only check the general trend of the whole business, but also examine from a particular angle. Below are some examples:

- Based on the outcome, there are slightly more than half of the program succeed.
- Most of the programs are from US. Because of this, US market is a better indicator for the general trend; what's most popular/ successful in the US are possibly the same internationally (in this dataset).
- Theater, Technology, and music are the top three most occurred categories.

2. What are some limitations of this dataset?

- Given the large size of this dataset, there are still some limitations. One of the limitations is, for companies that are outside US, the conclusions drawn from this dataset might not be as insightful for them, as majority of the data are from US; it would be better for them select their own home country in this dataset.
- Another limitation is also about the generalizability of this dataset. Because most data here are collected between 2015 to 2017, it is hard to say if the same results are holding truth in the following years. For those who would like to see more "trend" related information, it is better to look at the data that includes at least more years.
- Another limitation of this dataset is, it answers what is it, but not why it is like this. We can check the patterns and make hypothesis about the relationship between the variables, but without further tests it is hard to understand why the data are in this shape.

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

- We can also plot boxplot, which gives us an idea of where the majority are and where is the median.
- Scatterplot would also be a good idea for comparing two numerical variables.
- Bar chart is another good plot to visualize the relationship between categories/sub-categories and outcomes.