**Report Module 1 challenge**

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

The 56.6% of the crowdfunding campaigns are successful, the category that has more successful campaigns is theater which represents a 33% of the successful ones. Only 5.7% of the campaigns were canceled.

Thru the years I can´t identify a tendency in the behavior of the successful or the unsuccessful campaigns but in total of each month for all the years, I can conclude that the month with most successful campaigns was July and with most failed January.

In the outcomes based on goal graph we can see that percentage of successful campaigns base on the goal is inversely proportional to the percentage of failed crowdfunding campaigns, in other words, if one percentage increases the other decreases.

1. What are some limitations of this dataset?

The amount of data of each country is not the same and the majority (76.3%) is from the US so it will be skewed towards U.S. results, also I believe it will be better if all the data had the same currency for better understanding.

1. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

A bar graph on average donation depending on the country, that way we could see which country is more interested in crowdfunding campaigns.

Boxplots of backers counts and percent of funding to appreciate the distribution of the data.

1. Use your data to determine whether the mean or the median better summarizes the data.

In this case, the median summarizes better the data mainly because the data is skewed and there are many outliers (as we can appreciate in the boxplots), that could compromise the mean, in other words due to the outliers the mean would get higher, and it would not show the reality of the data. Other point that we need to take in consideration is that we are analyzing the number of backers, which means 50% of the number of backers in each crowdfunding campaign is by far lower than the mean.

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

The successful campaigns have more variability. It does make sense because the number of successful campaigns is larger than the unsuccessful and the range is bigger so the number of backers oscillate in a larger space. Also it shows that the data is more spread form the mean