**Reports**

**Q1**. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

Taking the above graph as an example,

1. **Successful Outcomes Dominated**: The outcome represented by light green bars (successful campaigns) consistently has the highest frequency across most dates compared to other outcomes (failed, live, and canceled campaigns).
2. **Steady Failed Outcomes**: The outcome represented by red bars (failed campaigns) appears consistently throughout the time span but with less frequency than successful outcomes. However, it is relatively stable over time.

From Goal Analysis,

1. **Goal Value**: Keeping in mind the no of projects, we can clearly conclude, that the campaigns which had 1000 to 4999 as their Goal has been the most successful with 83% success rate and similarly, the campaigns with more than 50000 Goal value are the least successful campaigns with 37% success rate.

**Q2**. What are some limitations of this dataset?

Limitations:

1. This dataset, because of its granularity, restricts the completeness and accuracy. As we see, from the above graph, majorly only 2 categories are visible, Green (successful campaigns) and Red (failed campaigns). But to derive any trends from this would be incomplete.
2. Although we can identify the most and least successful campaigns using goal analysis, this does not assist us in creating any trends or patterns because the dataset contains numerous anomalies. For example, there are numerous Goal divisions where the success rate is 100% (example: Goal value 15000 to 19999 and Goal value 30000 to 34999), but there are relatively few projects in these categories.

**Q3**. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

We could have added a graph wherein we have No of Backers as one dimension and Country as another dimension in the graph. This would give us a valuable insight in understanding the pattern as to why a certain country is more successful than the other. I believe, the no of backers is a very key component in a crowdfunding campaign.

**Statistical Analysis**

**Q1.** Use your data to determine whether the mean or the median better summarizes the data.

In my opinion Mean better summarizes the data, as it gives us Average rather than median which only gives us the middle point of a certain dataset.

**Q2**. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Successful campaigns see more variability throughout the dataset. Be it no of backers, or goal values, successful campaigns see a higher variability. This can be because our data is left aligned towards successful campaigns. Even if we analyse the data based on country, we have 565 successful campaigns out of 1000 against 364 failed campaigns out of 1000.