**Analysis Report**

**Written report**

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

* Just over half of the crowd-funding projects are successful.
* Most of the campaigns are for theatre, one-third, followed by music and film & video, which comprise two-thirds of the projects.
* There is a spike in successful campaigns launched in June and July. Summer in the northern hemisphere may have something to do with that.

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

* This is a snapshot, and when it is taken, it will have an effect on the data.
* The dataset is taken over a set period. Over a longer period of time may have seen different results.
* There are only a finite number of countries in the dataset. Not only is there a geographical bias but also would be a cultural bias.
* Crowd-funding platforms change over time, and this does not take into account changes in the platform.
* There may be external factors that affect the data, such as changing legislation, economy, weather and other events that may give more context.

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

* A pie chart of total projects per country may indicate where most data is focused.
* Tables that focus on the failed and the successful campaigns over time. That could have made line graphs over time for the different successful or failed campaigns.
* A table looking at the duration of the campaign and whether it was successful or failed. We could see if longer or shorter campaign length was a factor.
* We could use the spotlight column to see if that played into the success or failure of the campaign. Take two pie charts, successful and failed campaigns, divided by using spotlight or not.
* Box and whisker and a histogram when looking at the Bonus Statistical Analysis. Would show the skew of the data.

**Statistical Analysis**

For these datasets, the median is a much better indication of the central tendency of the data than the mean due to the data being heavily skewed to the left. The mean is much better for a more standard distribution. Most of the campaigns that were successful or failed did with much lower funding than the more expensive campaigns with very high levels of funding. There is a long tail to the right with a lot of outliers.