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

In this dataset, the most common category of crowdfunding campaigns is theatre, with the majority of campaigns in every country being plays. Around 57% of all campaigns are successful. The probability of a campaign being successful increases when the pledged amount greatly exceeds the goal. The most common goals for a crowdfunding campaign fall between the 1000-10,000 range, however there is no significant trend in the relationship between the goal amount and the % of successful campaigns. This suggests other factors influence a campaigns chance of success.

1. What are some limitations of this dataset?

While this dataset contains useful information about crowdfunding campaigns (number of backers, funding goals, amount pledged and category), it is highly likely that there are additional variables that may influence a campaign for which data has not been recorded. Examples of this could be more specific geographical information (cities/counties/regions within each country), demographic information about backers (age, gender, income etc.) including the amount funded by each backer, as well as any information about how the campaign was marketed (social media sites used, engagement numbers, events etc.) It is not clear which factors have been used to determine whether a campaign is successful or unsuccessful and so it is difficult to draw concrete conclusions about correlation and causation. There are also fields in the data which do not have a clear purpose (e.g. “staff pick” and “spotlight”).

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

It may be useful to calculate the length of the campaigns (using date started & date ended) and plot this against outcomes to determine whether there is a relationship between campaign length and chance of success. A table or visualisation for the % of the goal funded in relation to outcomes and/or other variables may also provide insight into correlation with success rate.