**CROWND FUNDING CAMPAIGNS REPORT**

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

. People prefer to back projects in entertainment industry (film & video, music, theatre) rather than in other areas

. People are more interested in backing plays and rock categories

. July and August are the most 2 months of the year that people are willing to going out more and spending more money in investment

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

This dataset only represents some certain projects in some industries. It didn’t tell the whole story that investing in entertainment industry is a key to success.

This dataset didn’t say how many invitations that each project’s owner sent out so it can tell how many people turned down/ not interested in an invitation.

There is no data value showing the number of backers coming from which country. It only mentioned the campaign was held in which country.

There was no age group information of backers.

There was no mention of these 1,000 sample projects was from 1 source or multiple platforms (e.g.: Kickstarter or Indiegogo or Fundrazr…). If this dataset came from multiple sources, the outcome can’t be justified.

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

Average donation and percent funded per outcome and backers so it can provide how much percentage each projects received and any correlated to the outcome of the project.

The length of each project can also provide an additional information of how long each project took to success or fail or cancel.

**Bonus Statistic Analysis.**

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

The median is better to summarise the data because the mean only calculates the average of the whole data set, it doesn’t take into account whether there is any significant changes or the gap between these data values. As a result, the mean in this situation has distorted the results and visualisation (due to some significant outlier values)

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

There are lots of variability with the successful and unsuccessful campaigns. For example, the project Jackson-Lewis only had 16 backers to attend but it was successful because it can raised more than the goal target. Whereas, in the unsuccessful campaign, for the project owner Whitehead, Bell and Hughes, it even had 5681 backers to attend and pledged the support but it couldn’t meet the target. As a result, it doesn’t make any sense and it can only be concluded that the number of backers attended at the event was not important as the amount of fund they pledged to support a project. If a project can attract an attention of the number of high key value backers, that project would easily achieve or surpass the target.

Besides, based on the standard deviation results, it can also be concluded that the data point is more spread out and above the mean value.