**Question 1:** the three main conclusions that we can draw about crowdfunding campaigns:

1. Successful campaigns are 1.6 times failed ones and generally have a greater number of backers than failed ones. Cancelled campaigns have the least number of backers
2. Theatres campaigns are the most numerous, journalism is the least in number but with 100% success rate. This result cannot be relied on much because of the small number of campaigns.
3. Campaigns are most successful in June and July, followed by a period of least successful number of campaigns in August.

**Question 2:** some limitations of the crowdfunding dataset

* The study did not separate the parent groups from their respective sub-categories. As a result, each parent category and its sub-category display the same results.
* Additionally, the data does not provide explanation for the success or failure of campaigns aside from the backers’ number, date, or categories and sub-categories.

**Question 3:** Other possible tables and/or graphs that we could generate, their additional value

* We could also have a table and a graph that displays the relation between number of backers and the amount the campaign raised regardless of other factors like categories. This will help establish the relationship between the amount raised and the number of backers. The more backers a campaign has the more it generates. This will help spot any outliers to this rule.
* We could also display the number of campaigns per year to determine any bigger trends that could be playing a role.

**Question 4:** Determining if the mean or the median better summarizes the data.

* The mean better summarizes the data than the median because the data follows a symmetric distribution and is closer to the average number of campaign backers as per the trend line in the scatter plots below. It represents the average value of the number of backers whereas the median is the value of the middle number of backers.

**Question 5:** Determining if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Since the variance and standard deviation for successful campaigns are greater than for failed ones, there is more variability in number of campaigns backers in successful campaigns more than with failed ones. This can also be visualized from the following scatter plots.

This makes sense because of the larger number of successful campaign 565, versus 364 failed ones and because their standard deviation is higher 1267 compared to 961 for failed campaigns.