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

  All these conclusions are worldwide:

  1. Theatre-related projects were the most published. 54% of these projects were successfull.
  2. The most successful category (apart from journalism is technology, with a 67% success rate, followed by photography (62%) and publishing (60%).
  3. The games category is the less successful category with 44% success and 48% failure rate (remaining 6% falls either into canceled or live). The food category has a similar story, with only an 48% success rate and 43% failure rate.
  4. Failure of the games category is due mainly to mobile games campaigns, as it has the lowest success rate of all sub categories and is 18 percentile points below video games (31% vs 49% respectively).
  5. The success rate of the film and video category is lowered by the science fiction proposals (-21pp vs category average).
  6. June appears to be the best month to launch a campaign as it has the highest success rate.
- What are some limitations of this dataset?
  1. There is a duplicity in the theater/plays category/subcategory combination. This is not inherently wrong but it makes the plays subcategory appear overrepresented.
  2. I excluded journalism from my conclusions because the sample size was too small.
- What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. I would arrange the data in a pivot that shows the % of successful campaigns that were staff picks to determine if there is a correlation between the two variables.
  2. I would add a pivot table with campaign duration to determine if that is a factor rather than month created or category (e.g. maybe plays are usually posted for a shorter period of time and therefore they don't reach their goal because of lack of time rather than interest).
  3. Another interesting chart would be by country and having category and subcategory as filters to compare the success rate of similar campaigns around the world.

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

  For this data set, the median summarizes the data better than the mean because both data sets are skewed right (positive skew). This can be shown by the very high SD, maximum and minimum values and the position of the mean and the median in relation to max and min.

- 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 have a bigger variability. This makes sense because if a campaign fails, it is because it didn't reach the minimum amount of money needed. If a campaign is successful, then the sky is the limit in terms of the number of backers it can get. Moreover, successful campaigns can be further accelerated thanks to a "snowball effect" as the more backers a campaign has, the more traction it gets and the more it is promoted.