1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   * When considering the whole time period, December launches is the only time of the year where overall, projects ‘failed’ more often than they were ‘successful’. When data is separated into individual years, projects consistently ‘failed’ more often than they were ‘successful’ for December launches for the last 3 years.
   * Overall, more theatre projects are launched than any other category. On a per country basis though, this is only the case for the UK, USA, Ireland, Australia, Canada, Denmark and Mexico.
   * Overwhelmingly, theatre plays are the most launched sub category overall, followed by rock music.
2. What are some limitations of this dataset?
   * In some countries, the number of projects is very limited.
   * Failed vs cancelled?
   * There is no way to tell what kind of marketing was done (or marketing budget) for an individual project that may also be a contributor to whether or not it was successful.
   * Some categories are very limited (e.g. journalism) and the sub categories within each category really need to be considered as often the sub categories have very different trends to the category as a whole.
3. What are some other possible tables and/or graphs that we could create?
   * Using percentages rather than total numbers so that each category can be better compared in terms of success rate.
   * Looking at the relationship between the goal and the number of backers (how much on average is an individual backer likely to pledge? How many backers are required on average to reach a minimum goal?)
   * Looking more closely at sub categories (for example within tech, hardware is always successful based on this data, while web, wearables and gadgets are rarely (if ever) successful.
   * Looking at the relationship between funding goal and % of goal reached / state. Could use a line chart.
   * Success rate of each subcategory by country (bar chart).
   * Pie chart or stacked bar chart showing % by state overall or of a subcategory.

BONUS

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

Because the data is skewed (not symmetrical) the median is more meaningful.

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

The standard deviation of successful campaigns is higher which makes sense as there are many more backers. In terms of coefficient of variation though, failed campaigns are more variable which wasn’t what I expected. This may be because the mode is 0, which is also one end of the distribution.