**Dataset Narrative**

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
* What are some limitations of this dataset?
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Based on the provided data, there are several conclusions that can be made regarding crowdfunding campaigns.

* Based on the categories provided in the dataset, crowdfunding campaigns seem to be used primarily for the arts, namely theater, film and video, and music. Crowdfunding might seem to be a natural funding source for technology efforts, but the technology category is the fourth most popular category. The categories generally are consistent across countries, however film and video is the most popular category in GB and CZ.
* When the data is broken down by sub-category, plays clearly is the most popular sub-category; plays are more than three times more popular than any other sub-category. The chart also highlights a few sub-categories that have a higher success rate, particularly the photography books and web sub-categories.
* Crowdfunding activity seems to be higher in the summer months, particularly successful campaigns. June and July account for the highest number of successful campaigns, while August was a popular month for failures.

One of the biggest limitations of this dataset is size. We need to be sure that the sample dataset sufficiently represents all crowdfunding campaigns and is not limited to certain categories, campaign sizes, or geographic areas.

Other areas that could be examined include the size of the crowdfunding campaign and the length of the campaign to determine if either of these attributes impact the rate of success. The question “Are smaller crowdfunding campaigns more successful than larger ones?” could easily be addressed with the data provided.

**Statistical Analysis**

* Use your data to determine whether the mean or the median better summarizes the data.
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

In this case, the median is a better summary of the data than the mean. A histogram reveals that the data is skewed to the right. Most of the crowdfunding campaigns had a smaller number of backers. In many cases where the data is skewed, the median is a more useful statistic, due to the number of outliers.

The data indicate that there is more variability with successful campaigns. That would make sense because successful campaigns might have more outlier backers who make very large contributions.