Crowdfunding Questions:

1. Given the provided data, what are the three conclusions that we can draw about crowdfunding campaigns?
   1. **Conclusion 1**: The most campaigns were executed in the theater category. This category also had the most successful campaigns, the most failed campaigns and the most canceled campaigns compared with the other categories. This suggests that theatre-based crowdfunding is overall the most lucrative category but also the most competitive.
   2. **Conclusion 2**: Crowd-funding in the theatre business is motivated by plays (a single sub-category) only; whereas successful crowdfunding in publishing for example is distributed across a portfolio of 4 sub-categories (products), with healthy ratios of success to failure across all four sub-categories (50% - 67% success rates compared to failures). This suggests that if a single publishing sub-category experiences more failures in crowdfunding, it will not have as big of an impact on crowdfunding for the parent category (publishing) as a decline in crowdfunding for plays would have on theatre. Thus, portfolio diversification may provide some insurance for the publishing parent category of crowdfunding.
   3. **Conclusion 3**: Campaigns overall saw a surge in successful crowdfunding in July followed by a decline in successful funding but an incline in failed funding in August, suggesting that August is a poor time to crowdsource funding.
2. What are some limitations of this dataset?
   1. A large majority of crowdfunding occurred in the US (76%) with the remaining 6 countries contributing less than 10% of the crowdfunding data. Therefore, one would want to be cautious about making conclusions about crowdfunding patterns in one country compared to the US as the other countries may not include a representative population of crowdfunding companies.
   2. We only have how many backers contributed to the total pledged amounts. Therefore the only measure of central tendency we can measure is average. This could be skewed if certain categories of crowdfunding tend to have a single large donor that drives the average falsely upwards of the true average.
   3. We can analyze the data to draw conclusions about what predicts a successful or failed campaign based on category, sub-category, time of year of conversion and percent funded. However, we cannot dive deeper into other variables that may be contributing to success or failure of campaigns, such as number of volunteers associated with the Kickstarter. Additionally, these unrepresented variable could be influencing the true effects of the variables represented in the data set.
3. What are some other possible tables and/or graphs that we could create and what additional value would they provide?
   1. I would create a pivot table that would compare percent funded to outcome to see if campaigns that met or exceed their goal were more or less likely to succeed.
   2. I create a column that a calculated the total time between launch date and deadline and then create a pivot table with outcomes to see if longer or shorter campaigns was a predictor of success.

Backers Questions:

1. Does the mean or median better summarize the data?
   1. The median appears to be better because the data are skewed to the low end with high standard deviation on the high end (suggesting higher variance from the mean). This is according to the box and whisker plot.
2. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this makes sense, why or why not? There appears to be more variability with the successful campaigns than failed campaigns. There are more data points (more overall successful campaigns) which could introduce variability. This doesn’t necessarily make sense though that there would be more variability in successful campaigns as you would think having a consistent number of backers (less variability) would be an indicator of success rather than failure. Unless having fewer backers in general (and less variability toward a lower mean) is a predictor of failure. Then it would make sense.