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

* If you start a campaign you’re more likely to succeed than fail, so given a rational and valid reason to crowdfund it would lead to a higher chance of success in funding your project.
* Those who did succeed in their campaign, got more money than they asked for allowing for a surplus of money to be spent on improving the quality of the product/project allowing for more resources to be available.

**What are some limitations of this dataset?**

The primary limitation of the dataset is the relatively small total number of projects, which becomes even more noticeable when we examine specific categories. In any analytical context, a larger quantity of data points is needed to establish a robust data distribution that accurately represents the system in question.

With smaller sample sizes, outliers can unduly skew the data, causing significant changes in key statistical measures such as mean, median, mode, and standard deviation. This skewing effect decreases as the sample size increases and the distribution becomes more stable.

For a truly reliable statistical analysis, we would prefer to have a large, evenly distributed sample size across all categories. The current dataset may not be sufficient for drawing firm conclusions about all aspects of crowdfunding campaign success.

**What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

There are several additional tables and/or graphs that we could create to provide more in-depth insights:

* Relative Success Table for Different Goal Brackets: We could create a table that displays the average amount by which successful campaigns exceeded their goals, broken down by goal range. This would provide a clearer understanding of not just the number of successful campaigns in each bracket, but also the degree to which they were successful.
* Comparison of Average Overfunding by Category: Similarly, we could create a table or graph that breaks down the average overfunding (the amount by which funding exceeded the goal) by category. This would allow us to see if certain types of projects tend to generate more excess funding than others.
* Scatter Plot of Goal Amount vs. Overfunding: Creating a scatter plot of the goal amount versus the overfunding amount could show us if there's a correlation between setting a higher goal and achieving overfunding.

**Statistical Analysis**

1. **Mean vs Median:**
   1. The mean is the average value of a dataset, while the median is the middle value of a dataset. When there's a significant difference between the mean and the median, it usually indicates that the data is skewed. In both successful and failed categories, the mean is substantially greater than the median. This suggests that there are a few campaigns with a very high number of backers that are raising the average, indicating a right-skewed distribution.
   2. In this case, the median would be a better measure to summarize the data because it is less affected by outliers and provides a more accurate representation of a "typical" campaign in terms of backers.
2. **Variability in Successful vs Unsuccessful Campaigns:**
   1. Variability in a dataset is typically represented by its variance or standard deviation (the square root of variance), which measures how spread out the numbers in the data set are. The successful campaigns have a higher variance and standard deviation than the unsuccessful campaigns, indicating that there is more variability in the number of backers for successful campaigns.
   2. This result makes sense and could suggest that while there is a base level of support for most campaigns, the highly successful ones manage to mobilize an exceptionally large number of backers. It might also indicate that successful campaigns employ a wider range of strategies or have a wider range of appeal, leading to more variability in the results.
   3. Additionally, the presence of a few successful campaigns with an extraordinarily high number of backers would also increase the variance and standard deviation, reinforcing the idea that the median is a more reliable measure in this context due to the skewness of the data.