Data Analysis Conclusions:

Looking at the pivot chart in Sheet2, we see that in absolute terms, the parent categories of Film & Video, Music, and Theater are the three categories with the greatest number of successes. However, they also have more failures than any of the other categories. This may be partially because these three categories are also the largest overall. If more crowdfunding campaigns fall into these three categories than others, then it stands to reason that those categories will have more overall successes and failures (possibly both) than the other categories.

We also see that as a category, Journalism contained the fewest number of crowdfunding campaigns. Although all its campaigns were considered successful, having only four total campaigns makes it difficult to see whether Journalism as a category has as much viability as the other categories.

Looking at Sheet2, Games was the only category that had more failures than successes, although the difference between failures and successes is quite small. We can gain more insight by looking at the pivot table for Sheet3. Video Games has more successes than failures, but only by two. Mobile games has 13 total campaigns, with eight being failures and four being successes. As a subcategory, Mobile Games has more failures, and thus, drags down the success of the Games category.

All the categories outside of Film & Video, Music, Theater, and Journalism had modest but substantial numbers of campaigns in them. By doing some mental calculations with the numbers in the pivot table in Sheet2, one can see that, except for Games, these remaining categories generally have as little as 33% more successes than failures.

Limitations and Other Useful Graphs:

While it is useful to know how many successful and failed campaigns fall into each category or subcategory, it would be helpful to how many successful and failed campaigns there were as a percentage of the whole. Categories can have widely varying numbers of campaigns. The graphs that are used here do show visually the proportion of how many successful and failed campaigns each category and subcategory have, but they do so with bars of varying heights. Having a chart that keeps all the bars the same height and shows only the proportion of successes to failures would be very helpful.

Another limitation is that we are not asked to present a way to see how each category breaks down into its subcategories. Per my third conclusion, we know that a category can have more failures than successes overall, while certain of its subcategories may be doing very well. The fact that a category has more failures than successes may simply be due to one poor-performing category. On the other hand, a category may be doing very well, but this may be due to just one or two subcategories that are performing very well. The other subcategories may be duds. We could have one bar graph for each category where the subcategories are listed along the horizontal axis and the number of successes and failures for each subcategory are shown as pairs of bars, one bar for the successes and one bar for the failures. It would also be useful to show on top of each bar the percentage of the whole subcategory that each bar in each pair represents.

Statistical Analysis Question Answers:

In both the successful and failed campaigns, the median appears to summarize the data best. The median and mean number of backers for the successful campaigns are 201 and approximately 851 respectively. For the failed campaigns, the median and mean number of backers are approximately 115 and 586 respectively. This indicates that there are more campaigns where the number of backers is less than average, as opposed to campaigns where the number of backers is greater than the average. This is regardless of whether the campaign was a success or a failure. This would suggest that the median is a better measure of center, and that the median summarizes the data better than the mean.

The variance for successful and failed campaigns is 1,603,373.732 and 921,574.6817 respectively. I guess this is understandable. In order to be considered a success, a campaign needs to only meet a minimum number of funds. Anything less and the campaign is considered a failure. One would expect more backers to correlate to more funds, thus increasing the odds that a campaign will be successful. Since a failed campaign is below its fund goal, it stands to reason that the number of backers may be limited, so there may not be very much deviation from the average.

For the successful campaigns, the mean is approximately 851, the median is 201, and the maximum is 7,295. For the failed campaigns, the mean is approximately 586, the median is approximately 115, and the maximum is 6,080. By these figures, successful campaigns tend to have more backers than failed campaigns. Because the number of funds that a successful campaign can have is hypothetically limitless, one could indeed expect a greater variability with the number of backers in successful campaigns.