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

1) Crowdfunding campaigns do not present a high rate of success or always the most encouraging results. Based on the data set provided, just over half of the pledged amounts equaled or exceeded their corresponding goal amounts. At least one-third of pledged amounts failed to reach their corresponding goals. Expectations, then, about results to come from crowdfunding should be tempered accordingly.

2) Among those campaigns that were successful, pledged amounts were concentrated in three areas overall: theater (plays), film and video, and music. Those areas represented more than two-thirds of successful campaigns and nearly 40% of all campaigns. Australia, Canada, and Italy each presented higher concentrations in the technology category than other nations (and Denmark likewise in publishing), but those nations comprised a relatively small representation of all campaigns in the data set. The high majority of campaigns were ones in the U.S. where the concentrations were clearly in theater (plays), film & video, music. Prospects, then, for projects outside of those areas of concentration will likely be less favorable.

3) Considering all the complete years for which data was provided (2010-2019), successful campaigns show peak outcomes overall in June and July each year, with June the best month for plays, and July for music projects. Film and video projects enjoy higher points in February and August with a stable, strong showing also from May through July. By contrast, failed campaigns are most prevalent, even on a rising slope, in the October-December timeframe. Those patterns should inform planning future projects and projecting outcomes from them.

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

For one, the data set is dated. Full-year results run through 2019 (only two outcomes are available for 2020), making the data by now to be four years old. Accordingly, it would be better to have full data for 2020 as well as 2021, if not also at least for the first half of 2022.

Another limitation is that the data is highly concentrated in U.S. data. The high majority of outcomes were for those in the U.S. And outside the U.S., data appears to be missing for key European countries. Although Denmark, Great Britain, and Italy are included, no data is presented, e.g., for France, Spain, and Germany. Neither is data included – or at least a qualifier that it is not available – for East European or Middle Eastern countries. And what about still other countries on other continents – Africa, South America, and others in the Pacific arena, e.g., Japan and Korea?

The presentation of goal and pledged amounts also begs the question whether data exists and would be available for the actual amount collected against what was pledged. If so, that could be even more telling about the success vs failure experiences for campaigns.

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

Histograms in general would illustrate the distribution and shape of the data set. And more specifically, Pareto charts would reveal the predominant categories of outcomes for crowdfunding projects.

Box plots would also depict in one graphic the measures of central tendency and position along with indications of variation. A box plot could be constructed for successful and unsuccessful campaigns each and compared and contrasted to the other.

**Use your data to determine whether the mean or the median better summarizes the data.**

For this data set, the median is a better summary measure about the data. Considering the respective minimum and maximum values for successful and unsuccessful campaigns each, along with respective measures of variance and standard deviation, one would expect to see in charts depicting distributions wide spreads of data points and possibly skewed ones as well. In contrast to the mean, the median would not be overly influenced by those extreme values and would prove to be a resistant measure of center.

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

More variability appears with successful campaigns than with unsuccessful ones as is evidenced by the much higher variance and standard deviation calculations for successful campaign data.

This makes sense, considering that the number of outcomes/projects for successful campaigns is much higher than those for unsuccessful ones (more than 50% higher) and the count of backers themselves for successful campaigns more than double those for unsuccessful campaigns. With the data set for successful campaigns being that much greater, then, it would be understandable to see more variability in it.

Jerry Youngblood