1. Three Observations about the data

The first observation is that the overall category of music has the highest success for all of the categories. 77.1% of all music Kickstarters were successful during this time range. Additionally, the unsuccessful sub-categories were predominately distributed between the faith and jazz categories. Excluding those two categories, the overall success rate for music is 93.1%

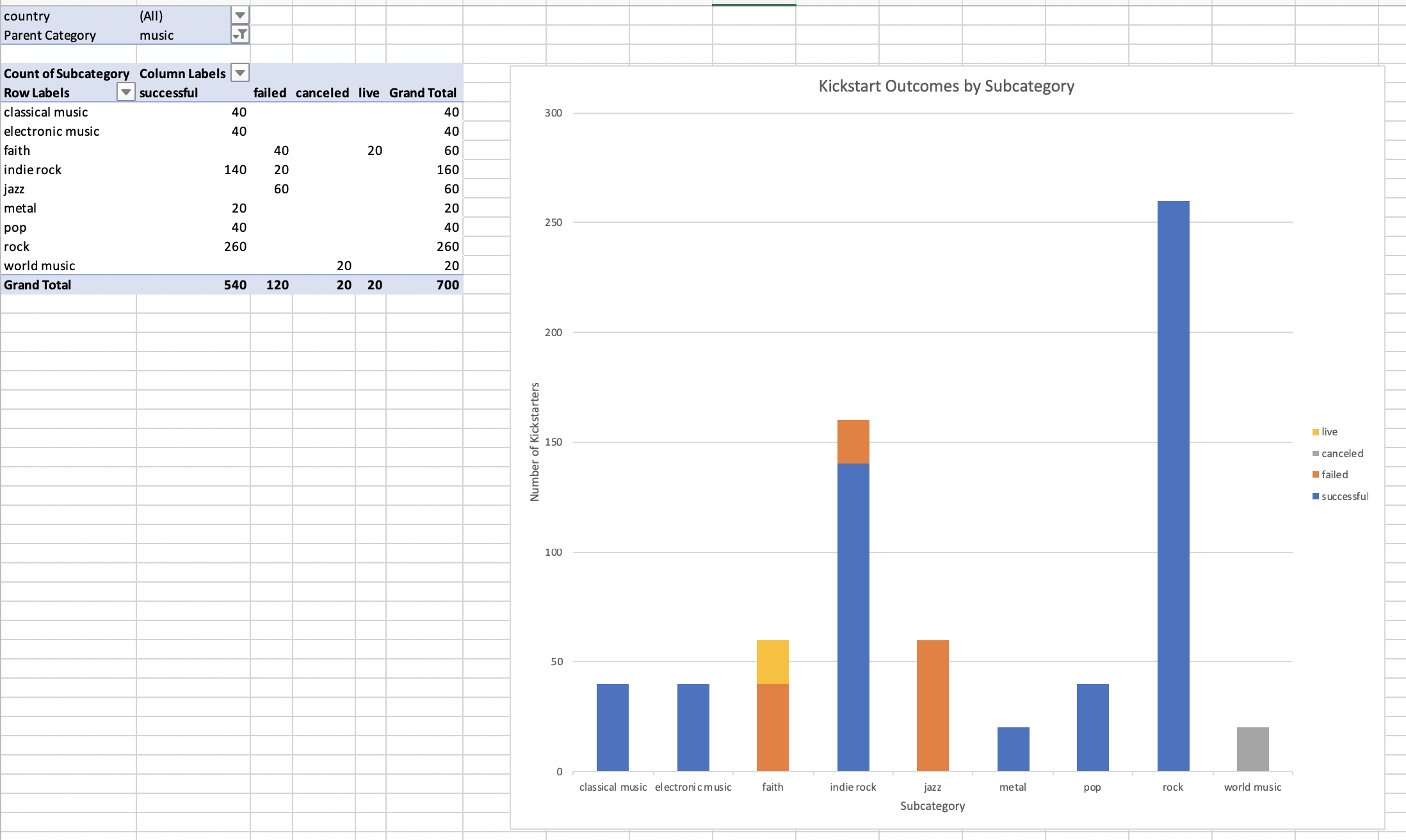


Figure 1. Faith and Jazz were entirely unsuccessful

The top 20 Kickstarters in the amount of money raised are predominantly in the technology category. 18/20 had the technology category. This can be seen by sorting them by the pledged amount. The subcategories in this technology group seemed to differ in their success, however. The Hardware sub-category, however was overwhelmingly successful.

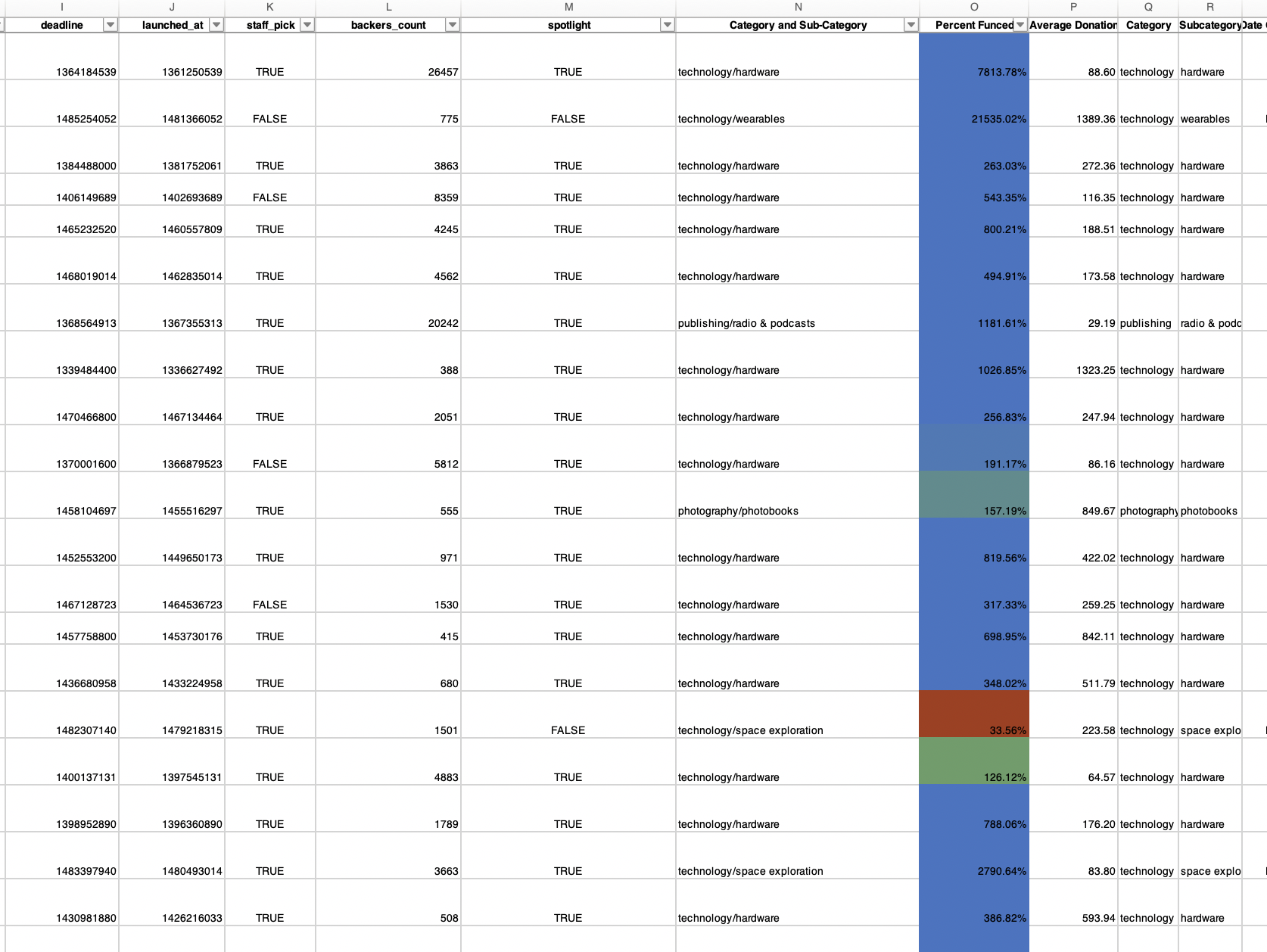
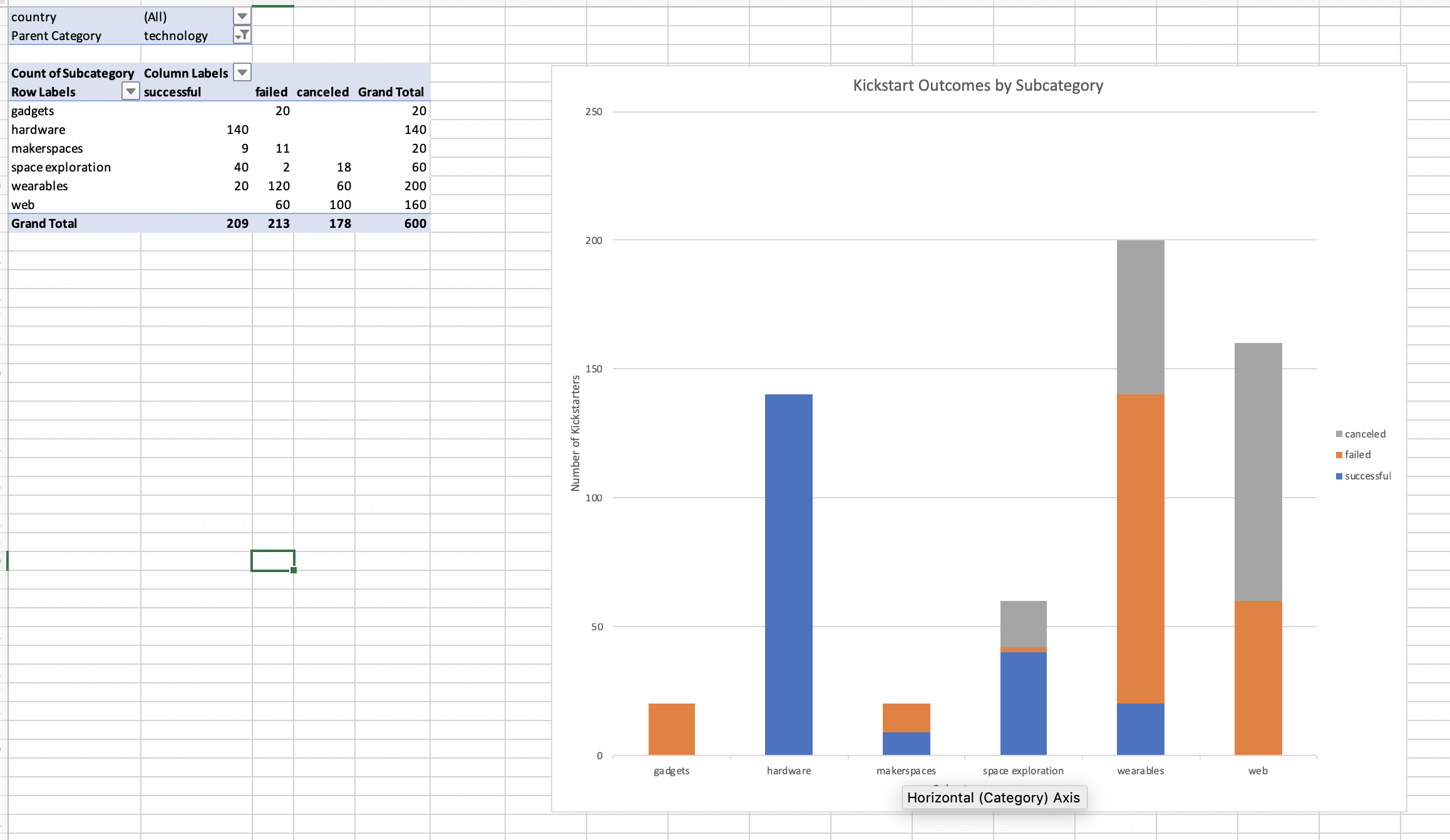


Figure 2: Top 20 highest pledged amounts.

Figure 3. Hardware Sub-Category Successes

The last observation is that there are both fewer Kickstarters and fewer successful Kickstarters in Q4 (Sept through December). The trend for Kickstarter numbers in both overall and successful appeared to be periodic. Most of the Kickstarts seemed to happen in the summer.

1. Limitations of the dataset

The dataset does not include several factors that could have contributed to the success of each individual Kickstarter. This data is lacking the information about product quality. Additionally, there are no product ratings for the good or service that’s being offered. There is no way to determine the how good what they were selling was.

This data also does not include any way to gauge what the entire write up of the Kickstart was. It’s limited to the short blurb and the media on the page, such as videos, pictures, additional writeups, and pledge rewards all contribute to if someone would back the Kickstarter or not. This write up would have also included information about the people who are doing the project, which is critical in determining the success of something that’s art focused, such as music, theater, film, or photography.

Additionally, this dataset does not have information about advertising, nor does it have any information that may track activity or links on social media. These things will contribute to exposure of each individual project, and in turn to their overall success.

1. Additional Tables / Graphs

It would be interesting to see if the staff pick Boolean value had an attributable effect on success or fail rate. To do this, I would create a graph in a pivot table that would have 2 bars. One bar will be for true and One bar would be for false. An alternate way to visualize this is to two separate lines for percent funded and compare how funded each of those.

It would be interesting to see if the average donation size had a bigger effect on success than number of donors. To do this, I would make one graph. First, I would normalize both sets of data so where the largest number of each is at 100 percent. Then, I’d create 10 categories for the lines greater than 0 to 10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, 70-80%, 80-90%, and 90-100%. I would use a total of 6 lines for this graph. There would be 2 sets of colors and 3 sets of patterns. The x axis would represent percentage. The Y axis would be kickstarter numbers. The solid green line would be successful by average donation. The dashed green line would be cancelled by average donation. The dotted green line would be failed by average donation. The solid blue line would be successful by number of donations. The dashed blue line would be cancelled by number of donations. The dotted blue line would be failed by number of donations. This would be created in a similar manner as the Bonus exercise in that all of those categories are created by countifs.

A third graph that could be interesting is a stacked line graph that counts the number of kickstarters on the Y axis and has the length of time on the x axis. The bottom line could be cancelled, the middle line, failed, and the top line successful. The stacked shaded lines will provide a view of their proportionality and since the length of time is contiguous, then we should be able to determine if the length the kickstarter was open had an effect on success.