Conclusions:

Crowdfunding campaigns in June and July are the most successful; campaigns in August and December are the least. When we apply standard deviation to the results of campaigns by month, we see that only numbers for June and July exceed one std dev, while numbers for August and December are lower by more than one std dev. All other months lie within one std dev.

Crowdfunding campaigns for technology are the most successful worldwide (excluding journalism, which had a statistically insignificant number of projects); campaigns for food and games are the least. Applying standard deviation to parent categories, we see that only technology rises above one standard deviation, while food and games fall below.

Interestingly, crowdfunding campaigns in the U.S. follow a different breakdown by parent category. While games fall below one standard deviation still, technology lies within one standard deviation; however, a new champion emerges – photography is the most popular crowdfunding campaign type in the U.S. See limitations for other countries.

Limitations:

The entire dataset is skewed heavily towards the U.S., as 76.3% of all projects counted were performed in the U.S. The remaining 6 countries only account for 23.7% of all projects, and as such it is impossible to determine which categories of projects might be most popular in individual countries. Because the dataset is skewed towards the U.S., seasonal changes that impact the U.S. might skew the results as well. Consider: December contains the Christmas holiday, and people’s budgets are often heavily impacted by the buying of presents for this holiday when they celebrate it. China is a country listed in this data, and only about 1% of the population there celebrates Christmas, meaning that their December projects could be considerably different. Further, seasons (especially Winter and Spring) might impact the data, but Australia is listed and has opposite seasons to those of the other nations in the list.

Regarding funding, we might be tempted to conclude that campaigns should always have a goal between 10000 and 35000, as for 3 of the 4 ranges, campaigns are always successful, and for the range 25000-29999 they are still successful 79% of the time. However, the number of campaigns in the range 10000 to 49999 is only 84 total campaigns, representing just 8.5% of the total campaigns run during this time. Compare this to the range 5000 to 9999, which had 315 total campaigns, or 31.9% of the total. As such, one limitation of the analysis is that there isn’t enough data in the individual monetary ranges to draw reliable conclusions. Additionally, campaigns with funding greater than 50000 are all grouped together, but the largest goal in the study was 199200, significantly higher than the 50000 floor for its group. Extremely large projects might succeed or fail at rates that skew the data. Extremely small projects might as well – 20 projects had goals of only 100, but had at most one backer and an average donation of 5 or less.

Those monetary values are also in local currencies (rather than dollars). Because these results are not equal in currency type, exchange rates might skew the data.

Future study:

One graph that would be useful would be a breakdown of goals by percentile instead of by range. This would give us a standard number of projects in each division, which could make for a better breakdown for determining how project goals relate to project success.

Another useful breakdown would be making a pivot table combining parent categories and sub-categories, so that we may look at the breakdown of sub-categories within a parent category. For instance, seeing electric music, indie rock, jazz, metal, and rock together in a single table would be more useful than seeing them alongside wearables and video games. This would also help us to see what sub-categories make up a parent category, as “games” could theoretically include board games and sports, and we would be remiss if we made conclusions that included those types of games when our conclusions are based on only mobile and video games.

One thing that could be easily analyzed given the data we already have is how the length of the project impacts its success. We know both the launch and end dates for the projects, and we could subtract these two to determine the length of each project, and apply a similar percentile breakdown to them to see their success rates.

Finally, it might be useful to see a breakdown of funding averages and medians by month. We might see that projects within certain monetary ranges happen more frequently in certain months, and this might affect how often they succeed or fail. For instance – projects with a goal of less than 1000 fail 39% of the time, but do these projects also happen more frequently in August and December than they do in other months? Is December the “Big ticket” month, as crowdfunding projects are placed there in the hopes of benefiting from people’s good will during Christmas, and this skews the data for either December or those more expensive projects? This breakdown would help answer that question.

Summary Statistics:

In the Summary Statistics tab, Median appears to be a better summary of the data than the Mean, because the standard deviation is quite large. Among both the successful and failed projects, the standard deviation would include values that fall below Zero. This is largely due to Right-skewing of the data - there are a small number of projects with a very large number of backers.