**Crowdfunding Analysis**

Jason Pealy

Northwestern Data Science Bootcamp (Spring 2023)

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

* Crowdfunding campaigns appear to be most common among projects in the performing arts (film & video, music and theater), while crowdfunding does not appear to be a popular way to raise money in journalism. This could be because people are more likely to support the projects of their favorite performing artists (directors, actors/actresses and musicians) with donations, while in the case of journalism most publications simply charge a subscription fee which eliminates the need for crowdfunding.
* Overall success rates for most of the parent categories and the different sub-categories were above 50% in terms of meeting their stated funding goals. Only the ‘games’ parent category had an overall success rate below 50% and only 2 of the 24 sub-categories had overall success rates below 50% (mobile games and science fiction). This could be because the creators of different projects will engage in some research before starting their crowdfunding campaigns by polling in order to gauge interest in their campaign. Doing some pre-emptive research would likely ensure that they have a higher chance of success than if they just picked a random project to start a crowdfunding campaign for.
* Overall cancelation rates among projects across all parent categories and sub-categories were very low. Two thoughts come to mind. First, if the projects’ creators are doing a bit of research ahead of time (and have a reasonable expectation of success based on their research), then there shouldn’t be a reason to cancel the campaign along the way. This would make sense given the relatively high success rates across all categories/sub-categories. Second, as a crowdfunding campaign can easily be run through any number of different online platforms these days (Kickstarter, IndieGoGo), it is a relatively low-cost way to fundraise which eliminates most of the overheard costs normally associated with fundraising. This would allow a person running a crowdfunding campaign to do so without the fear that running their fundraising campaign would be too expensive.

1. One limitation of the dataset is its size – 1000 points of data on crowdfunding campaigns is not a large enough sample set to make any definitive conclusions beyond some basics about which categories have more instances of crowdfunding. Additionally, without any additional data (beyond country) of where the projects and the crowdfunding campaigns are based, it is difficult to tell if there are other sampling bias such as big city vs small town preferences or differences in interests by region within a given country. Finally, there is a lot of deviation in terms of the size of fundraising goals across the different campaigns in the dataset. There are several large campaigns but then a lot of very small ones in the dataset. It may paint a more accurate picture if the datasets were separated based on relative size (in terms of fundraising goals).
2. Adding a pivot graph with a regular column chart the success/failure counts side-by-side by category/sub-category would make it easier for the viewer to determine the success and failure rates more quickly than the stacked column chart. A column chart showing the different outcomes by category/sub-category as a percentage of their total (i.e., 50%-win, 40% failure, 10% other) would be another option to make it easier to digest the information. Adding a chart showing success rates by number of backers and/or average donation would be a good way to help potential creators of new crowdfunding campaigns determine if they should be focusing on small donations or large ones.