# Barron 2023-02-17 Crowdfunding Report

## Three Conclusions

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + The majority of Crowdfunding projects tend to involve theater (34.4%), film & video (17.8%), music (17.5%), and technology (9.6%). At the subcategory level, the most prominent Crowdfunding projects include plays (34.4%), rock music (8.5%), documentaries (6.0%), and web (5.1%).
  + Investigating outcomes grouped by month started, there is a tendency for Crowdfunding projects that started in June and July to be successful compared to other months. This tendency does not occur in August.
  + On average--across all categories, goals, and years—any given Crowdfunding project within this dataset has a 56.5% chance of being successful. On average, Crowdfunding projects within the technology, publishing, and photography categories tend to be slightly more successful. Crowdfunding projects within the food and games categories tend to be slightly less successful. Note there is not enough data on the journalism category to make informed inferences.
* What are some limitations of this dataset?
  + This dataset seems to include Crowdfunding projects across the globe. Accordingly, the goal and pledge variables appear to be in different currencies depending on the country of origin. They cannot be directly compared (see ‘unknown denominator problem’) until they are converted into a comparable metric, which necessitates using conversion rates based on the year the project took place.
  + Presumably, one prominent confounding variable in Crowdfunding success is the state of a country’s economy when the Crowdfund occurred. Inferences about these statistics, particularly if comparing across years, will likely be inaccurate without controlling for that statistically.
  + It is likely one would want to use this dataset for applications within their particular context. For instance, to determine the probability their Crowdfunding project would be successful. It’s possible that this data may not be a representative sample for their context (i.e., might not generalize to their context). For instance, cultural variables may cause country to moderate the relationship between Crowdfunding project characteristics and Crowdfunding success, so one may only want to use data from their country to ensure a representative sample.
  + In a similar vein to the above point, the Crowdfunding landscape has likely changed over the last few years, particularly with the pandemic. I suspect it would be best to remove data before 2017 or so, as data before then is likely outdated.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + **Using the amount pledged as the dependent variable:** All of the analyses thus far have operationalized the outcome as a nominal variable based on whether the funding goal was met. It would be intriguing to look at the pledged amount as a dependent variable once it is converted into a comparable metric (e.g., inflation-controlled US currency). It would be interesting to look at how the amount of money pledged varies across categories and varies across months. This would provide information about the amount of money rather than the success of the Crowdfund. For instance, we could create a bar plot with Category as the categorical x variable and amount pledged as the continuous y variable.
  + **Incorporating Staff Pick and Spotlight as predictors:** It would be interesting to investigate whether Crowdfunding that was spotlighted or a staff pick differ in success rates from Crowdfunding that was not spotlighted or not a staff pick. Causal claims would be difficult, but it would still help elucidate the relationship between these and Crowdfunding success. For instance, we could create a grouped bar chart with count of outcome as the y variable, and spotlighted and staff pick as x grouping variables.