Claire Fraser

UC Berkeley Data Analyst BC

10/23/2023

Module 1 Challenge

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   1. The most popular crowdfunding campaign categories are theater, film & video, and music.
   2. Plays are the most popular sub-category of all crowdfunding campaigns.
   3. About half of the crowdfunding campaigns have been successful, and about 30% have failed.
2. What are some limitations of this dataset?
   1. The amounts pledged are listed in different currencies, and therefore are difficult to compare across countries.
   2. Most crowdfunding campaigns included in this list fall under the plays sub-category, which may result in bias (data only represents play-related crowdfunding campaigns).
   3. Very few journalism crowdfunding campaigns (underrepresented).
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
   1. A clustered bar graph with columns representing the “FALSE” and “TRUE” staff pick counts for each parent category, and Sum of backers counts along the y-axis. This would show which crowdfunding campaign categories have the highest number of staff picks and backers and is useful if you need to decide which category is best to create a new crowdfunding campaign in. \*See “staff-pick” Sheet in .xlsx file\*
   2. A bar graph showing average or total number of backers counts for each country. An average would provide information about how many backers are involved in each project, while the total numbers would tell you which country has the most backers overall.

Statistical Analysis

* The median summarizes the data better than the mean because this dataset is skewed right.
* There is more variability in backers counts with successful campaigns, compared to unsuccessful campaigns. This makes sense, as the number of backer counts is likely to be higher for projects that were successful.