

Data Analysis Excel Challenge

(Crowd Funding Data)

Submitted by:

Roshni Rana

University of Toronto

Data Science, Analysis, and visualization, 2024

Instructor : Bharat

Presented report includes answer the following questions:

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
2. What are some limitations of this dataset?
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Before creating a report using the data that was provided, let's examine the information and make some deductions:

Conclusions about Crowdfunding Campaigns:

1. **Campaigns That Are Successful:** Most of the dataset's crowdsourcing initiatives are profitable. In a variety of areas, including Theater, Music, and Film & Video, successful campaigns outweigh canceled, unsuccessful, or in-progress ones. This implies that raising money for different initiatives through crowdfunding can be successful.
2. **Campaign Success Rate by Category:** Some categories have higher campaign success rates than others. For instance, compared to journalism, which has very few effective campaigns, theater has many. This suggests that depending on the project's categorization, crowdfunding campaigns may or may not be successful.
3. **Temporal Trends:** The data reveals variations in the monthly total number of campaigns. Crowdfunding campaigns, for example, peak in the summer (June, July) and decline towards the end of the year (November, December). This implies that a crowdfunding campaign's success could be impacted by when it is launched.

Dataset Restrictions: (Limitations of the dataset)

1. **Absence of Context:** The dataset does not provide campaign details, financing targets, or campaign duration, although it does include counts of outcomes for various categories and months. Without this knowledge, it is difficult to comprehend the variables affecting the success or failure of a campaign.
2. **Missing Data:** The analysis as a whole may be impacted by the absence of information for a few categories or months. Results for several parent categories, such "audio" and "world music," are absent from the dataset, for example.
3. **Aggregate Data:** You can't undertake in-depth research or pinpoint specific trends within subcategories since the data is aggregated at a high level (e.g., counts of outcomes by category or month).

Additional Tables and Graphs for Analysis:

1. **Success Rate by Category:** Make a table or bar graph that displays the proportion of campaigns that are successful for each category. This would reveal which categories have the highest rates of success.
2. **Trend Analysis Over Time:** For each outcome category, create line graphs that display the number of campaigns over time (e.g., by month or year). This would make temporal trends easier to see and patterns in campaign activity easier to spot.
3. **Geographical Analysis:** Examine the distribution of results by nation or region and, if possible, incorporate data on the geographic locations of campaigns. This may shed light on geographical variations in the success of crowdsourcing.
4. **Analyze the relationship** between the length of the campaign and its result (success, failure, etc.). Make a box plot or histogram that displays the distribution of campaign durations for campaigns that were successful and unsuccessful.

By incorporating these additional tables and graphs into the analysis, We may obtain a deeper comprehension of the dynamics of crowdfunding and pinpoint the elements that influence a campaign's success or failure