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

Table Monthly Outcome Distribution by Quarter

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Count of outcome** | **Column Labels** |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| **Qtr1** | **17** | **97** | **1** | **142** | **257** |
| **Jan** | **6** | **36** | **1** | **49** | **92** |
| **Feb** | **7** | **28** |  | **44** | **79** |
| **Mar** | **4** | **33** |  | **49** | **86** |
| **Qtr2** | **7** | **93** | **4** | **147** | **251** |
| **Apr** | **1** | **30** | **1** | **46** | **78** |
| **May** | **3** | **35** | **2** | **46** | **86** |
| **Jun** | **3** | **28** | **1** | **55** | **87** |
| **Qtr3** | **17** | **89** | **2** | **144** | **252** |
| **Jul** | **4** | **31** | **1** | **58** | **94** |
| **Aug** | **8** | **35** | **1** | **41** | **85** |
| **Sep** | **5** | **23** |  | **45** | **73** |
| **Qtr4** | **16** | **85** | **7** | **132** | **240** |
| **Oct** | **6** | **26** | **1** | **45** | **78** |
| **Nov** | **3** | **27** | **3** | **45** | **78** |
| **Dec** | **7** | **32** | **3** | **42** | **84** |
| **Grand Total** | **57** | **364** | **14** | **565** | **1000** |

Table Outcome Distribution by Content Type

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Count of outcome** | **Column Labels** |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| animation | 1 | 10 | 2 | 21 | 34 |
| audio |  |  |  | 4 | 4 |
| documentary | 4 | 21 | 1 | 34 | 60 |
| drama | 2 | 12 | 1 | 22 | 37 |
| electric music |  | 8 |  | 10 | 18 |
| fiction | 1 | 7 |  | 9 | 17 |
| food trucks | 4 | 20 |  | 22 | 46 |
| indie rock | 3 | 19 |  | 23 | 45 |
| jazz | 1 | 6 |  | 10 | 17 |
| metal |  | 3 |  | 4 | 7 |
| mobile games |  | 8 | 1 | 4 | 13 |
| nonfiction | 1 | 6 | 1 | 13 | 21 |
| photography books | 4 | 11 | 1 | 26 | 42 |
| plays | 23 | 132 | 2 | 187 | 344 |
| radio & podcasts |  | 4 |  | 4 | 8 |
| rock | 6 | 30 |  | 49 | 85 |
| science fiction |  | 9 |  | 5 | 14 |
| shorts | 1 | 5 | 1 | 9 | 16 |
| television | 3 | 3 |  | 11 | 17 |
| translations |  | 7 |  | 14 | 21 |
| video games | 1 | 15 | 2 | 17 | 35 |
| wearables |  | 16 | 1 | 28 | 45 |
| web | 2 | 12 | 1 | 36 | 51 |
| world music |  |  |  | 3 | 3 |
| **Grand Total** | **57** | **364** | **14** | **565** | **1000** |

Table Outcome Distribution by Parent Category

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Count of outcome** | **Column Labels** |  |  |  |  |
| **Parent Category** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| film & video | 11 | 60 | 5 | 102 | 178 |
| food | 4 | 20 |  | 22 | 46 |
| games | 1 | 23 | 3 | 21 | 48 |
| journalism |  |  |  | 4 | 4 |
| music | 10 | 66 |  | 99 | 175 |
| photography | 4 | 11 | 1 | 26 | 42 |
| publishing | 2 | 24 | 1 | 40 | 67 |
| technology | 2 | 28 | 2 | 64 | 96 |
| theater | 23 | 132 | 2 | 187 | 344 |
| **Grand Total** | **57** | **364** | **14** | **565** | **1000** |

## i. Seasonal Variations:

From the first pivot table, there appears to be some seasonality in crowdfunding outcomes. For instance, the first (Qtr1) and third (Qtr3) quarters have the highest totals of canceled campaigns, while the second (Qtr2) has the highest number of successful campaigns. This might indicate that timing can be a significant factor in the success or failure of a campaign.

## ii. Category Significance:

The second pivot table underscores the importance of campaign categories. Theater campaigns, despite having the highest number of failures, also boast the most significant number of successes. Similarly, music and film & video projects show a favorable success rate. Conversely, journalism projects have the lowest numbers across all outcomes, suggesting they might be less popular or viable on crowdfunding platforms.

## iii. Sub-Category Insights:

Diving deeper into sub-categories, 'plays' within the theater category have a high failure rate, indicating that while the overarching category is successful, specific niches within might face challenges. In contrast, genres like 'indie rock', 'rock', and 'documentary' have commendable success rates, revealing them as potentially lucrative niches for crowdfunders.

# 2. What are some limitations of this dataset?

## i. Lack of Financial Insights:

While the dataset provides counts of various outcomes, it doesn't provide financial details such as the amount of money sought, the amount raised, or the average donation size. These insights can be crucial in understanding the financial viability of projects.

## ii. No Data on Marketing Efforts:

There’s no information on the promotional or marketing efforts behind each campaign. Such data could offer insights into what strategies work best for crowdfunding campaigns.

# 3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

## i. Success Rate by Category and Sub-Category:

A table showcasing the percentage success rate for each category and sub-category can quickly pinpoint the most and least successful niches.

## ii. Month-wise Breakdown of Outcomes:

A line graph plotting the monthly success, failure, and cancellation rates can provide a clearer picture of the seasonal variations and their impact on campaign outcomes.

## iii. Category-wise Distribution of Live Campaigns:

Given that 'live' campaigns are ongoing, a pie chart or bar graph can help stakeholders understand which categories currently dominate the crowdfunding space and where potential investors are focusing.

## iv. Outcome Trends over Time:

A time-series graph showcasing the trend of successful, failed, and canceled projects over months or years can provide insights into the evolving nature of crowdfunding preferences and market dynamics.

# 1. Does the mean or the median better summarize the data?

## Successful Outcome:

* Mean: 480.898
* Median: 87.5

## Failed Outcome:

* Mean: 213.164
* Median: 0

Analysis: When there's a significant difference between the mean and median values, it's often an indication of skewness in the data.

For the successful outcome, the mean is much higher than the median, suggesting a right-skewed distribution. This could mean that there are some campaigns with an exceptionally high number of backers that are increasing the average value.

For the failed outcome, a median of 0 indicates that at least half of these campaigns had no backers at all. Yet, some failed campaigns still had a considerable number of backers, leading to a mean of 213.164.

Conclusion: Given the difference between the mean and median values, the median might be a more representative summary of the central tendency, especially for the failed campaigns, as it is not influenced by extreme values or outliers. Therefore, the median gives a clearer picture: half of the successful campaigns have backers less than 87.5, and at least half of the failed campaigns have no backers. However, the mean provides an average which can be useful to understand the general trend but might be influenced by outliers.

# 2. Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or why not?

## Analysis:

Variability can be assessed using either variance or standard deviation. A higher value indicates more variability:

## Successful Outcome:

* Variance: 1,083,958.116
* Standard Deviation: 1,041.133092

## Failed Outcome:

* Variance: 414,846.4111
* Standard Deviation: 644.0857172

The successful campaigns show a higher variance and standard deviation compared to failed campaigns, indicating that there is more variability among the successful campaigns.

There is more variability in the number of backers for successful campaigns than for unsuccessful ones. This might be expected: successful campaigns could vary widely in their appeal, with some just reaching their funding target, and others becoming viral sensations. Conversely, failed campaigns, by definition, didn't resonate as widely, so they generally have fewer backers, leading to less variability.

The vast range of backers in successful campaigns reflects the unpredictable nature of what might become popular or resonate with the audience, while the consistency in failed campaigns might indicate that they struggled uniformly to capture backers' interest.