**Crowdfunding Dynamics: An Analytical Perspective**

# Introduction

# This report provides an analysis of the crowdfunding data, identifying key patterns across categories, understanding dataset limitations, and assessing statistical attributes. The objective is to offer a perspective on the dynamics of crowdfunding success.

# Main Trends and Observations

-Notable categories such as Film and Video, Theater, and Music see a high volume of crowdfunding projects. Possible reasons could be the greater audience interest or inherent appeal in these project types.

-Categories like Theater and Film and Video don't just have a high number of projects; they also have impressive success rates. This refers to effective project strategies or strong supporter-to-supporter engagement.

-Activity in different months show general stability with slight increases in January and July. This can show the effect of particular events or seasonal factors on crowdfunding.

# Data Limitations

*Contextual Gaps*:

We can only look at surfaces without understanding the reasons behind the results. For example, increased cancellations may be due to external factors such as a recession or internal factors such as poor policy. This lack of depth prevents us from developing strategies. It also means that any conclusions we draw can be based on correlation rather than causation.

*Limited Categorization:*

Provider categories may not cover all of the services available on the platform. Some projects may span multiple categories or not fit well into one. This can erode our understanding of popular or successful niches. Furthermore, detailed segmentation can reveal hidden features or opportunities within the platform.

*Sample Size Discrepancies:*

Categories with fewer projects might not provide a representative sample. For example, high success rates in a category with very few jobs may be due to coincidence rather than realism. Drawing conclusions from these categories can provide insight if it deceives. When assessing the importance of any attribute or pattern, it is important to consider the sample size.

# Further Analytical Possibilities

* Utilize a pie chart to show the number of projects in different outcome categories.
* Employ Box plots to visualize the amount of funds allocated to various project outcomes.
* Assess success rates variations over different months or years for individual category and subcategory.
* Pinpoint elements closely with project success.

# Statistical Analysis

To determine whether the mean or the median better summarizes the data, it's important to consider the presence of outliers, skewness, and the spread of data. Here are the observations:

For Successful projects, the mean number of backers (851.15) is considerably higher than the median number of backers (201). For Failed projects, the mean number of backers (585.62) is also notably higher than the median (114.5). This significant difference between the mean and median in both categories suggests that the data is skewed.

The maximum number of backers for both Successful and Failed projects are much higher than their respective means. Specifically, the maximum for Successful projects is 7295, contrasting with its mean of 851.15. For Failed projects, the maximum is 6080, much higher than its mean of 585.62. Moreover, the minimum value of 0 for failed projects suggests a positive skew, where the majority of data points might be clustered at the lower end with few values at the higher end.

The standard deviation, which measures the dispersion of data, is high for both categories, suggesting that the number of backers varies widely from the mean.

The significant difference between the mean and median suggests the data might be skewed. In such cases, outliers can significantly influence the mean. Given the substantial difference between the maximum values and the mean for both categories, it's likely that there are outliers that are pulling the mean higher. The median, on the other hand, is more resistant to outliers and provides a more central measure of the data's distribution.

Given the presence of outliers and the skewness suggested by the substantial gap between the mean and median, **the median might be a better summary statistic for this data set as it is less influenced by extreme values.**

To determine the variability, we can look at measures like the variance or the standard deviation.

From the provided data:

|  |  |  |
| --- | --- | --- |
|  | Successful | Failed |
| **Mean Number of Backers** | 851.15 | 585.62 |
| **Median Number of Backers** | 201 | 114.5 |
| **Minimum Number of Backers** | 16 | 0 |
| **Maximum Number of Backers** | 7,295 | 6,080 |
| **Variance of Number of Backers** | 1,603,373.73 | 921,574.68 |
| **Standard Deviation of Number of Backers** | 1,266.24 | 959.99 |

Higher variance or standard deviation indicates more variability in the data.

Successful projects have a higher variance and standard deviation compared to Failed projects. This suggests that the number of backers for successful projects varies more widely from the mean than that for failed projects.

The mean number of backers for successful projects is higher than that for failed projects, and this is also reflected in the wider spread (or variability) of the number of backers.

The observed trend aligns with the expectations for the following rationales:

* Successful projects might have a broad range of appeal, where some projects become viral hits with a massive number of backers, while others achieve their goals with a modest number of backers.
* Successful projects might have vastly different financial and supporter goals. Some projects might aim for a large backing for ambitious projects, while others might have smaller, niche targets.
* Failed projects might have a more consistent lack of appeal or challenges that prevent them from getting a wide range of backers. The number of backers for failed projects might consistently fall below a certain threshold, leading to less variability.

# Conclusion

It's logical for successful projects to have more variability because the success threshold can be achieved in various ways - some projects might just surpass their goal with a few backers, while others might massively exceed their goals. In contrast, failed projects, regardless of their goal, did not resonate enough with the audience, leading to a more consistent pattern of fewer backers.

In conclusion, the higher variability in the number of backers for successful projects suggests a broader range of outcomes within the realm of "success," while failed projects tend to be more consistently below their desired backing.