**Crowdfunding Campaigns: Success or Failure?**

***Given the provided data, one can safely draw the following three conclusions, out of plenty of more that can be drawn.***

1. Although not all crowdfunding campaigns are successful, the overarching result of the outcome of the campaigns in this dataset shows that crowdfunding campaigns are generally successful over 50% of the time. For instance, looking at pivot table one, 565 campaigns (56.5%) were successful while 364 campaigns (36.4%) failed. Not counting any would-be successes out of the 14 campaigns which were still live, it is evident that there are more successes than failures and cancellations combined. This conclusion can be made by looking at the provided data as a whole.
2. The provided data also shows that most crowdfunding pledges happened in the U.S. within the provided time frame. The U.S. has had more than double the amount of the remaining six countries combined in the provided sample data.
3. Out of the seven countries represented in the data, Theatre is the highest category in four of them– the U.S., Canada, Denmark, and Italy. On the other hand, there were no campaigns for the category ‘Journalism’ in all the countries except the U.S.; although the data shows that there were only four campaigns for this category, it was successful all four times, making it the only category with a 100% success rate.

***Limitations of the dataset***

One of the more prominent limitations of the provided dataset is that it lacks valuable demographic information of the campaign backers. According to a survey conducted to establish the demographics for crowdfunding participants in the U.S. (startups.com, 2024), age, gender, income, and the devices used by prospective crowdfunding participants were found to have correlations with the success of campaigns. For example, the survey showed that ages 24-35 were much more likely to participate in crowdfunding campaigns, while those over 45 were less likely. Similarly, men were found to be participating in unknown startups than women. The income levels of campaign backers also play a pivotal role in the amount of funds that can be raised. If these information were present in the provided sample dataset, a more nuanced analysis of the success and failure of the campaigns would be possible.

Another limitation that comes to mind is the type of crowdfunding employed in the campaigns in the provided sample dataset. According to crowdfunding.com, crowdfunding can be categorized into different types. These include donation-based, reward-based, equity, and debt (crowdfunding.com, 2024). If such information were present in the provided dataset, a detailed analysis of which form of crowdfunding works better vis-a-vis the categories would be possible.

***Additional tables and graphs recommended***

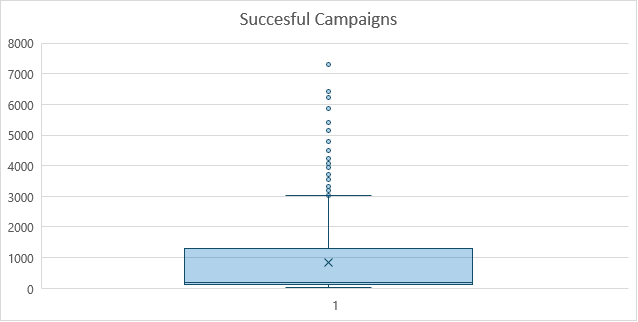
If the above information listed in the limitations were present, additional tables that show campaigns by demographics such as age, gender, and level of income could be created. These tables could also be analyzed using more bar charts and possibly pie charts to show which backers, based on demographic information, account to what amount/degree of backing to a given category.

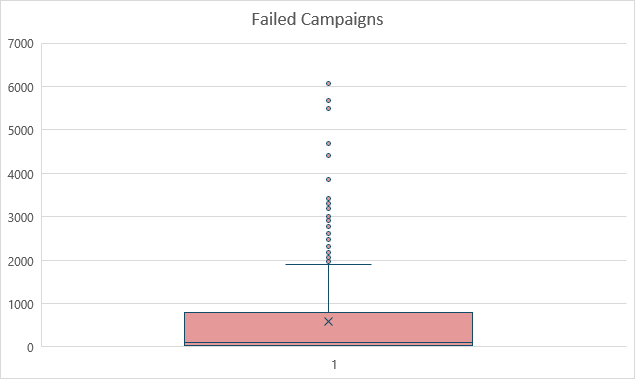
The type of crowdfunding strategy used could also generate a valuable insight into the provided data. It would be interesting to see which form of crowdfunding generates more traction in relation to the given categories.

Another table that can summarize the provided dataset can be a table that shows the goals vs the percent funded that can be filtered by category. For this table, a clustered column chart can be used to compare the values across the categories.

***Better Summary for the Data: Mean vs Median***

To gain an insight into this topic, Box and Whisker charts were created for both successful and failed campaigns. The charts show that the data is skewed and entail that the median better summarizes the provided data.





In another attempt to identify between the mean or the median when it comes to better summarizing the provided data, the excel SKEW formula was employed for both the successful and failed campaigns. The results are as follows:

*Successful Campaigns: =SKEW (B2:B566) = 2.18*

*Faile Campaigns: =SKEW (E2:E365) = 2.70*

Since the values for both the successful and failed campaigns have a skewness of over 2, thereby indicating a highly skewed distribution, the median is a better measure for the central tendency of the provided data. In other words, the median is more reliable since it is not drastically affected by outliers.

***Variability: Successful vs Unsuccessful Campaigns***

The standard deviation for the successful campaigns is 1267.37, whereas the standard deviation for the failed campaigns is 961.31. Hence, there is more variability in the successful campaigns. This makes sense since there are more outliers in the number of backers for the successful campaigns. Additionally, one can see that there is a wider range of value in the number of backers for the successful campaigns with a maximum of 7295 as opposed to the maximum number of backers for the failed campaigns, which is 6080.