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

Seasonal Patterns in Campaigns: There seems to be a consistent pattern of higher campaign activity during certain months. The months of March, May, June, and July consistently have higher numbers of successful campaigns, indicating that these months might be more favorable for launching and achieving success in crowdfunding campaigns. This pattern could be due to seasonal trends, holidays, or other external factors.

Cancellation Rates Remain Relatively Steady: The number of canceled campaigns appears to be relatively consistent across different months. While the numbers fluctuate slightly, there isn't a strong trend indicating a specific month with significantly higher or lower cancellations. This could suggest that the reasons for cancellations are spread across the year rather than being concentrated in specific months.

Higher Failure Rates in Winter Months: The number of failed campaigns seems to be relatively higher during the winter months of December, January, and February. This might indicate that campaigns launched during this period face challenges or competition that result in a higher likelihood of failure. It could be attributed to factors such as decreased consumer spending post-holiday season or increased competition from other campaigns launched during this time.

1. What are some limitations of this dataset?

Lack of Contextual Information: The chart might not provide context about the campaigns themselves – their nature, goals, target audience, etc. Without this information, it's challenging to understand why certain months had more failures, successes, or cancellations.

No Insights into Campaign Performance: The chart might not show any insights into why campaigns failed, succeeded, or were canceled. Factors such as campaign content, marketing strategies, and funding goals aren't captured in the chart.

Inability to Analyze Campaign Trends: The chart might only display the counts of outcomes by month. It doesn't show trends in campaign performance over time, which might be essential to understanding seasonal patterns or long-term shifts.

Missing Granularity: If you're interested in understanding specific aspects contributing to the outcomes, like categories of campaigns, funding goals, or geographical regions, a simple pivot chart might not offer enough granularity.

Variability in Campaign Duration: If campaigns have different durations, comparing monthly outcomes might not provide accurate insights. A successful campaign that ran for two weeks isn't directly comparable to one that ran for an entire month.

Inability to Analyze Causes: The chart doesn't explain why campaigns failed, succeeded, or were canceled. Factors like project quality, communication with backers, or economic factors aren't covered.

Complex Relationships: If there are interactions between variables affecting campaign outcomes (e.g., high funding goals leading to more cancellations), a pivot chart might not effectively capture these complex relationships.

Limited to Categorical Data: Pivot charts are designed for categorical data. If you want to incorporate continuous variables like funding amounts, you might need additional visualization techniques.

No Interactivity: Pivot charts usually lack interactivity, which means users can't explore the data further, such as by clicking on a specific month to see details about the campaigns for that month.

Difficulty in Detecting Anomalies: Pivot charts might not readily highlight outliers or anomalies in the data, which could be significant for identifying unusual campaign performance.

Visual Clutter for Long Time Periods: If you have multiple years of data, displaying all months in a single chart might lead to visual clutter and difficulty in discerning trends.

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

Time Series Line Chart: Create a line chart with time on the x-axis (months) and the count of successful, failed, and canceled campaigns on the y-axis. This will provide a clearer visualization of trends over time and seasonal patterns.

Stacked Bar Chart: Instead of separate columns for each outcome, use a stacked bar chart to show the total number of campaigns each month and how they are distributed among successful, failed, and canceled categories. This can highlight the relative proportions of each outcome.

Success Rate Bar Chart: Create a bar chart where each bar represents a month, and the height of the bar represents the success rate (percentage of successful campaigns). This can help understand the overall trend of campaign success.

Heatmap: Use a heatmap to visualize the distribution of outcomes over months. This can help identify months with consistently high or low campaign success rates.

Box Plot: If someone wants to understand the distribution of funding amounts for different outcomes, create box plots for each month, showing median, quartiles, and potential outliers for successful, failed, and canceled campaigns.

Regression Analysis: there are additional variables that might influence campaign success (like funding goal or campaign duration), someone could perform a regression analysis to understand which factors have a significant impact.

Each of these options can provide different perspectives on your crowdfunding campaign data, allowing you to uncover trends, patterns, and insights that might not be immediately evident from a single pivot chart.

1. Use your data to determine whether the mean or the median better summarizes the data

The mean and median for successful campaigns' backing counts are quite different (561 vs. 220). This suggests that the distribution might be skewed or influenced by outliers.

The mean and median for unsuccessful campaigns' backing counts are also notably different (97 vs. 44), again hinting at a potential skew or the presence of outliers.

Given the differences between the means and medians in both cases, it's likely that the data distributions are skewed or contain outliers. In such cases, the median could be a more robust measure of central tendency, as it is less affected by extreme values. However, it's important to analyze the underlying data distribution and the nature of the data before making a definitive conclusion.

1. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Variance and standard deviation are measures of variability or dispersion in a dataset. Higher values of variance and standard deviation indicate greater variability.

In this case: The variance of successful campaigns' backing counts is 360496, with a standard deviation of 600.

The variance of unsuccessful campaigns' backing counts is 21128, with a standard deviation of 145.

The variance and standard deviation for successful campaigns are significantly higher than those for unsuccessful campaigns. This indicates that there is more variability in the backing counts of successful campaigns compared to unsuccessful campaigns.

This finding does make sense. When a campaign is successful, there can be a wide range of factors that contribute to the amount of backing it receives. These factors might include the popularity of the campaign's concept, effective marketing strategies, appealing rewards, and broader community support. These variables can lead to a greater spread of backers and backing amounts. On the other hand, unsuccessful campaigns might have less variation because they didn't attract sufficient interest or lacked effective elements that draw backers in.