Part 1:

Based on the analysis of the CrowdfundingBookAssignment workbook, here are three conclusions we can draw about crowdfunding campaigns:

1. Variation in Success Rates: The outcome column indicates whether a campaign was successful, failed, or is still live. By analyzing the frequency of these outcomes, we can determine the success rate of crowdfunding campaigns. This is important for understanding the factors that contribute to a campaign's success or failure. Using a simple bar chart we can determine that Technology is the best performing category with a 67% success rate. Journalism has a very small sample size and cannot be considered significant.
2. Success Rate and Funding Goals: Using the SuccessRate\_FundingGoal sheet we can see a difference in average goal between successful and unsuccessful campaigns:

|  |  |
| --- | --- |
| Success Rate and Funding Goals |  |
|  |  |
| Average Goal for Successful Campaigns: | $26,405.31 |
|  |  |
| Average Goal for Unsuccessful Campaigns: | $67,122.17 |

Campaigns with lower funding goals might have higher chances of success. This can be inferred by looking at the 'goal', 'pledged', and 'Success' columns. For examples, campaigns with goals that are relatively modest compared to the average might have a higher likelihood of achieving or surpassing their funding targets.

1. Backer Engagement. The number of backers “backers\_count” and the average amount donated “Average Donations” are indicators of campaign success. Campaigns with a higher number of backers and reasonable average donations have a higher success rate. This relationship highlights the importance of not just the funding amount, but also the community and supporter engagement in crowdfunding success. You can see from the data below that both average number of backers and average donation have an impact of whether the campaign is successful or not:

|  |  |
| --- | --- |
| Backer Engagement |  |
|  |  |
| Average Number of Backers for Successful Campaigns | 851.15 |
| Average Number of Backers for Unsuccessful Campaigns | 568.38 |
| Average Donation for Successful Campaigns | $69.43 |
| Average Donation for Unsuccessful Campaigns | $65.41 |

Successful campaigns tend to have a higher number of backers and slightly higher average donations, indicating that both the quantity and the quality of backing are important for success.

The Limitations of this dataset include:

Limited Demographic & Geographic Information: Without detailed demographic information about campaign creators or backers (such as age, gender, or location), it's challenging to draw conclusions about who is most likely to succeed in crowdfunding and why. Also, If the dataset does not contain certain regions, the findings might be too generalized and not offer important analysis.

Missing or Incomplete Data: If the dataset is missing values in key columns or is missing data from amount of time then the results will lead to incomplete analysis and skewed results. Also, If the dataset is not recent, it might not reflect current trends or changes in the crowdfunding, such as new backer preferences or new categories.

Other possible tables and/or graphs that could be created

1. Category Influence: To explore how different categories impact the success of campaigns, we could examine the Top 5 Successful Parent Categories and Top 5 Successful Sub Categories. Observing these success rates could suggest that the type of project significantly influences its likelihood of success.
2. Campaign Timeline Analysis: By calculating the time span between the creation and end date of each campaign, we could construct a comprehensive timeline. This timeline would illustrate the number of campaigns launched over time and their respective success rates. Additionally, it would enable us to analyze whether the most successful campaigns are those that achieve or exceed their pledged funding goals within this timeframe.
3. Geographical Analysis: Assuming data on the locations of campaign creators is available, analyzing this geographically could reveal regional differences in crowdfunding success. Such an analysis would help identify significant geographical variations, uncovering patterns that indicate how the success of crowdfunding campaigns might be influenced by their creators' locations.

Part 2:

Determining Whether the Mean or Median Better Summarizes the Data:

For Successful campaigns, the mean is 880.39 and the median is 203. The mean is substantially higher than the median, suggesting that the data might be skewed with some outliers pulling the mean up.

For Unsuccessful (or Failed) campaigns, the mean is 591 and the median is 118, similarly to the successful campaigns, the mean is higher than the median, indicating a skewed distribution.

In both cases, the median would be a better measure of central tendence since it is less affected by outliers and skewed data. The median gives us a better representation of a “typical” campaign.

Determining Variability Between Successful and Unsuccessful Campaigns:

The standard deviation for successful campaigns is 1293 which is higher than the standard deviation for unsuccessful (or Failed) campaign which is 963. This indicated there is more variability in the number of backers towards successful campaigns compared to unsuccessful.

It does make sense since successful campaign are likely to have a wider range of backers. They can potentially go viral or attract a vast number of backers due to marketing, demographic, and geographic influences. Unsuccessful campaigns, on the other hand, may not gain as much traction, resulting in a more uniform (and lower) number of backers across different projects, hence less variability.