Module 1 Challenge | Crowdfunding

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

Given the provided data, we have been able to determine the success rate by category, sub-category, and time of year. By analyzing the relationship between outcome and categories/sub-categories, we can determine which categories/sub-categories have a higher success rate. This can help identify which types of projects are more likely to be successful and which ones may require more strategic planning. Also, by examining the outcomes by the time of year, we can identify trends in campaign success and inform future campaign launch strategies.

* **What are some limitations of this dataset?**

This dataset has several limitations. For instance, it does not account for external factors that could influence campaign success, such as economic conditions, market trends, or competition. Additionally, the dataset lacks information on marketing efforts, which can also significantly impact crowdfunding success.

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

The assignment tasked us with creating a new column to display the percentage funded, but delving deeper into the relationship between the campaign's goal and the amount pledged could provide valuable insights into the impact of goal setting on campaign success. Perhaps a scatter plot could be used to visually represent this information, enhancing our understanding of the data and inform future goal setting.

**Statistical Analysis** (Also Included in Excel workbook)

|  |  |  |
| --- | --- | --- |
|  | successful | failed |
| **The mean number of backers** | 851 | 586 |
| **The median number of backers** | 201 | 115 |
| **The minimum number of backers** | 16 | 0 |
| **The maximum number of backers** | 7295 | 6080 |
| **The variance of the number of backers** | 1603374 | 921575 |
| **The standard deviation of the number of backers** | 1266 | 960 |

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

The median is a more effective summary of the data as it is moderately influenced by skewness, whereas the mean is heavily influenced by the maximum number of backers.

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

The standard deviation and the variance for successful campaigns is greater than those of the unsuccessful campaigns. Therefore, the successful campaigns have more variability in the number of backers compared to unsuccessful campaigns. Yes, this makes sense because the successful campaigns have a wider range of backers, leading to a higher variability.