Ishicka Forrest

Challenge #1

**Three Conclusions:**

* Category does have an influence on the outcome. The analysis of campaigns by category reveals that most of the categories have similar outcomes except for plays which have the highest number of both successes and failures. When comparing the fails against successes for each category, Plays has the highest successes compared to all other categories. This data can help make informed decisions by the campaign about the type of projects they should pursue based on trends.
* Time of the year does not adversely affect the success of the outcome based on the data. There are no major spikes in the campaign success or failure. Over the period the different categories had a relatively close number of successes and failures. The data does show that June and July have the highest success rate hence this could be taken into consideration for campaign dates.
* Analyzing the outcome, the percentage of successful campaigns are slightly higher than failures, however data reflects that not all crowdfunding campaigns are successful.

**Limitations of the Dataset:**

* The dataset may not include all relevant variables, making it challenging to perform a comprehensive analysis.
* The dataset does not include comprehensive factors that could affect campaign success; hence data analysis may be skewed with additional factors such as economic conditions, financing capacity and marketing efforts.
* No demographic information is available, this limits the analysis as knowing about the different characteristics could help determine which demographics is more likely to support a specific campaign so that a target market can be created in order to drive the success rate.

Additional Tables and Graphs:

Pledge Amount Over Time: A line graph showing how the pledge amount changes over the period can determine whether they are willing to pledge more in the categories that are successful.

Geographic Analysis: Tables that analyze how campaign outcomes vary by country, which will provide insights into the impact of geographic location on crowdfunding success.

**Data Analysis:**

|  |  |
| --- | --- |
|  |  |
| **Statistical Analysis for Successful Backers Campaigns** | |
| **Mean** | 851 |
| **Median** | 201 |
| **Minimum** | 16 |
| **Maximum** | 7295 |
| **Variance** | 1606217 |
| **standard deviation** | 1267 |
|  |  |
|  |  |
|  |  |
| **Statistical Analysis for failed Backers Campaigns** | |
| **Mean** | 586 |
| **Median** | 115 |
| **Minimum** | 0 |
| **Maximum** | 6080 |
| **Variance** | 924113 |
| **standard deviation** | 961 |
|  |  |

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

In determining whether the mean or median better summarizes the data we take into consideration whether the data are normally distributed sample. If the data has a symmetrical distribution the mean and the median would be equal and based on the calculations, they do not. The data therefore consist of outliers and due to these outliers, the median may be better to summarize the data.