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

* The number of backers alone does not guarantee the success of a campaign in raising funds. Even with 6080 backers, as shown in the Statistical Analysis tab of the Excel workbook, some campaigns still failed.
* Analyzing the Outcome by Month tab, campaigns launched in June and July appear to have a higher success rate compared to those launched in other months. Conversely, August shows a lower success rate and a higher rate of failure. However, further investigation is needed to determine the underlying reasons
* Small, niche projects such as Journalist and World Music projects have a 100% success rate, as indicated in the Outcome by Category and Outcome by Sub-Category tabs. There are 4 Journalist projects and 3 World Music projects, which represent a smaller portion of the sample data compared to other project types. The goal range for Journalist projects is between $2,400 - $9,700, while for World Music Projects it is between $4,500 - $88,400.
* Among the project types, theater is the most popular, followed by film & video, and music. Over 50% of projects in these categories have been successful in crowdfunding.

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

* The sample data spans a period of 10 years, from 2010 to 202. Data from five to ten years ago may not accurately reflect the current landscape, given the heightened public awareness and participation in internet fundraising for startup business today.
* With 763 campaigns originating from the US and 237 from other countries, there is a potential limitation in drawing conclusions about trends specific to the US. For a more precise analysis of the US market, focusing solely on US-based samples would be more appropriate. Conversely, if the aim is to assess global internet funding projects, the dataset should ideally balance samples from each country, rather than being predominantly US-centric. As it stands, the dataset is biased towards, studying the US trend, accounting for over 70% of the samples.
* The Statistical Analysis tab in the workbook reveals significant variances in both successful and failed projects, indicating that the dataset encompasses a wide range of samples. Such disparities in data size may hinder the accuracy of the study’s results.
* There is a lack of information regarding the specific internet platforms used for these campaigns within the sample. The popularity and credibility of crowdfunding platforms could potentially have substantial impact on the success or failure of these campaigns
* Crucially, there is no available information about the credibility of the campaign owners. It is conceivable that some of the successful campaigns were launched by well-known individuals or established companies, which could be a significant factor in their success.

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

* Another potential visual representation for this analysis could be a table illustrating the duration of campaign launches and the corresponding numbers of success and failures. This aims to investigate if there exists significant correlation between the campaigns launch duration and its success or failure rate. It’s plausible that campaigns with more time allocated for fundraising may yield higher success rates, while those with limited time frames may experience lower success rates. Additionally, longer campaign durations may enhance project visibility and awareness.
* To construct this table we can categorize campaign launch durations into ranges in the first column. Subsequent columns can display the counts of successful and failed projects, as well as the corresponding percentages. These results can then be visually represented in a graph for a clearer presentation of the findings.

**Statistical Analysis**

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

* In the sample of 1,000 campaigns, there were 565 successful campaigns and 364 failures. The average number of backers for successful campaigns was 851, whereas for failed campaigns, it was 586. Looking at the medians, successful campaigns had a median of 201 backers, while failed campaigns had a median of 114.5 backers.
* In this dataset, utilizing the mean provides more comprehensive summary since it offers the average value of the sample. On the other hand, the median merely identifies the middle value in the sample by arranging the data from lowest to highest. It provides a snapshot of the midpoint without considering the weight of each data point. Consequently, relying solely on the median may lead to less accurate results.

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

* The successful campaigns exhibit greater variability compared to the unsuccessful ones. However, it’s important to note that both categories display a significant degree of variability. This wide variance indicates that the data points are widely dispersed and note closely clustered around the mean. Consequently, the dataset’s lack of consistency makes it challenging to make accurate predictions about the broader population.
* In this dataset, it’s logical for both successful and unsuccessful campaigns to demonstrate substantial variance. This steams from the considerable disparity between the smallest and largest goal amounts. The smallest goal is set at $100, with potentially zero or just one backer, while the largest goal exceeds $199,200, garnering anywhere from two to four thousand backers. This immense gap between a zero-dollar goal and an almost $200,000 goal significantly impact the numbers of the backers for the lower and higher dollar campaigns.
* Given the extreme range of backers in the dataset, where goals vary widely, the mean number of backers for successful campaign stands at 851, while for unsuccessful campaigns, it’s 586. However, it’s important to acknowledge that this may not accurately represent the true population due to the extreme difference between the lower and highest number of backers in the dataset. For more accurate research outcomes, it would be beneficial to consider campaigns with more closely aligned goal ranges, resulting in smaller variances and standard deviations.