Written Analysis:

Three conclusions that be drawn from the Crowdfunding Campaign dataset include: The most successful campaigns were film & video, music, and theater campaigns; plays was the most common sub-category and was also the most successful sub-category; and the most successful campaigns were created in June and July.

However, there are some rather important limitations to the dataset. The first limitation is the number of campaigns per country. All countries, with the exception of the US, had fewer than 50 total campaigns. As a result, the campaigns are less likely to be representative of potential campaigns outside of the US. This is exacerbated since the campaigns were tracked over a 10-year period, so while they hit the minimum threshold of N=30 to be a representative sample, with them being spread over a 10-year period likely decreases how representative they are. Additionally, all campaigns were started between January 2010 and January 2020. While there was a 10-year period of data, given the context of the COVID-19 Pandemic, the data is outdated and likely not representative of how campaigns would fare in 2024.

Other tables and/or graphs that could be created include: average donation per campaign outcome by category and sub-category and backers count per campaign outcome by category and sub-category. While the average donation per campaign outcome table and graph would require a column to convert all donations to USD, it could provide some additional context to the backers count table and graph. These 2 tables and graphs could show differences in how popular certain categories or subcategories are as well as providing information on whether campaigns are being funded by a few people pledging a lot of money or by a large group of people each giving a smaller amount. This could provide much more information on how popular certain categories and subcategories actually are amongst the population. For example, if most of the theater campaigns that were successfully funded were only funded by a few people, then it would indicate that theater campaigns are likely popular within a small, but dedicated, portion of the country. However, if that were the case, it could indicate that theater campaigns might not hold broad appeal across the country. On the other hand, if the theater campaigns were largely funded by a lot people donating small amounts of money, assumptions could be drawn that theater campaigns are widely popular in that country.

**Statistical Analysis:**

Source: From the “Crowdfunding Summary Statistics” worksheet in “CrowdfundingBook\_challenge.xlsx

|  |  |  |
| --- | --- | --- |
| **Summary Statistics** | **Successful Campaign** | **Failed Campaign** |
| Mean | 851.1469027 | 585.6153846 |
| Median | 201 | 114.5 |
| Minimum | 16 | 0 |
| Maximum | 7295 | 6080 |
| Variance | 1603373.732 | 921574.6817 |
| Standard Deviation | 1266.243947 | 959.9868133 |

For both successful and failed campaigns, the median better summarizes the data than the mean. The data appears to be heavily right skewed because, for both outcomes, the standard deviation is greater than the mean and the mean is roughly 4 times the value of the median. Using the mean and standard deviation, it looks like the successful campaigns were more variable. This makes sense for multiple reasons including: the variance for successful campaigns is almost double that of failed campaigns, the standard deviation for successful campaigns was over 300 more than failed campaigns, and the failed campaign with the most backers had 1,200 fewer backers than the maximum backers for a successful campaign.