Pivot Table Analysis

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  1. The most campaigns took place within the “theater” parent category, followed by film & video, and then music. Journalism had the least campaigns.
  2. The subcategory with the most campaign was plays. World music was the subcategory with the least campaigns.
  3. The most campaigns took place during July, and the least campaigns took place during September.
* What are some limitations of this dataset?
  1. A limitation of this data could be that we don’t know the definition of success. It only shows us that something is marked as failed if they did not receive all the backers that they were hoping for, but that does not mean that the platform was not successful. Success is not only determined on backers or funding but on qualify of product and there is no one to determine the quality of product with the data that we have here.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. I believe a table or graph showing what time of the year do most projects start. This would be to show that if in a certain time of the year you can find more success starting a business
  2. There are a lot of different categories of business’s that were started. I think a table showing what type of business usually finds success compared to oens that don’t could also be a helpful visual tool.

Statistical Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Successful Campaigns** | |  | **Failed Campaigns** | |
| Mean | 851.1469027 |  | Mean | 585.615385 |
| Median | 201 |  | Median | 114.5 |
| Minimum | 16 |  | Minimum | 0 |
| Maximum | 7295 |  | Maximum | 6080 |
| Variance | 1606216.594 |  | Variance | 924113.455 |
| Standard Deviation | 1267.366006 |  | Standard Deviation | 961.3082 |

1. The mean better summarizes the data in this dataset– the median of both gets pulled down by outliers, but the mean shows the higher on average backers of successful campaigns than failed campaigns.
2. There is more variability with successful campaigns– this is shown both in the variance and the standard deviation. The standard deviation is larger on successful campaigns, which shows that there is more distance in the data. This makes some sense– it shows that successful campaigns are not completely dependent on the number of backers, so it can range from a lower to a higher number. It also makes sense when looking at the minimum and maximum– the successful campaigns overall have a larger range of backers than failed campaigns.