**Crowdfunding observations**

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  1. Although plays had a lot of failed campaigns, they were very successful in more than half, bringing in a lot for the crowdfunding.
  2. June & July had a good success to failure ratio (with limited canceled), making those months the best months for crowdfunding.
  3. When the goals were in the range of 15000 to 35000, they had a success rate of 95% combined.
* What are some limitations of this dataset?

One limitation is the cost of creating each activity. Although plays were successful, they also have a high cost to implement (construction, materials, etc.). We don’t know if the goal was after costs or before. Did the activities actually make money.

Also, the goal information covers multiply countries. There is no indication about exchange rates and if the values reflect the exchange rate of the individual country or their own.

We don’t know the specific location and population size for the selected data. This would be important to know when tracking how to set the goals for a crowdfunding. There is a big difference being in a big city verses a small town when it comes to raising funds.

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

I would create a graph to show donations per country/city. This would be useful to know if you get better results in a small, medium, or big city to see that correlated with a successful campaign if there are certain places have better success than other locations.

Another table/graph would be donations by donor per category. This would help to determine if a handful of donors provided the success of all the crowdfunding or did a lot of donations from multiple people provide success for the goals.

**Statistics**

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

The median would be best to summarize the data. This reflects more of an accurate value than the mean would. There are a few larges values that are increasing the mean and making that data point less useful in this situation.

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

While it looks like both have a lot of variability, the successful campaigns almost double the variability of the failed campaigns. This does make sense to me because the success if based on the goal which is typically a number that randomly picked by a few individuals. We don’t have any indication that the goals were set from previous years donations, changes in the economy, etc.

Failed

|  |  |
| --- | --- |
| Mean= | 585.6154 |
| Median= | 114.5 |
| Min= | 0 |
| Max= | 6080 |
| variance= | 921574.7 |
| std deviation= | 959.9868 |

Successful

|  |  |
| --- | --- |
| Mean= | 851.1469 |
| Median= | 201 |
| Min= | 16 |
| Max= | 7295 |
| variance= | 1603374 |
| std deviation= | 1266.244 |