Report for Crowdfunding

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

The conclusions from the Crowdfunding campaigns are as below:

* Almost all the campaigns ran for 60 days or less.
* All campaigns for Journalism category are successful. Next highest successful categories are Technology, Photography and Publishing. However, the maximum failed campaigns were in Games, Food, Music and Theater respectively.
* Campaigns that ran for sub-categories Audio, World music were a 100% successful. However, for the sub-category’s science fiction, Mobile games and Radio & Podcast it was difficult to find the backers.
* During January, the Success rate was highest whilst the least during the month of August. On the other hand, most campaigns that were unsuccessful are during the month May.
* Amongst all 7 countries, most of the campaigns were conducted in USA.
* The highest amounts collected were staff picks.

2)What are some limitations of this dataset?

The limitations of this dataset are:

* If the reasons for outcome of the campaigns were provided, statistical analysis would be more specific and helpful for future campaigns.
* We could only deduce the Average donations, if there is information on donations or at least minimum and maximum donations, the forecast can be predicted.

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

* The probability of the successful, failed, and cancelled campaigns can be generated using pivot-charts.
* The staff picks verses outcome tables can be created for a better understanding of the staff-picked campaigns.
* The spotlight verses outcome tables would provide better understanding of the campaigns.
* The country verses outcome table gives the broader picture of the outcomes as shown below:

Table

Description automatically generated

**Statistical Analysis:**

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

* The median summarizes the data in a better way because the data does not have a symmetrical distribution.
* The method of separating the higher sample with the lower value, usually from a probability distribution is termed as the median, hence from the data, the outcomes form various countries can be determined accurately using median as it takes the middle value rather than taking the average of the values.

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

* There is more variability with successful campaigns. This makes sense because, when we calculated the first, second and third quartiles for the successful outcomes, the data distribution can be accurately calculated. The below table depicts the statistical analysis for the successful outcomes:

Table

Description automatically generated

The box plot below depicts the graphical distribution of the successful outcomes :

Chart, box and whisker chart

Description automatically generated