ASSIGNMENT 1

**CROWD-FUNDING**

QUESTION-ANSWERS

SUBMITTED BY:

LOVEPREET SINGH

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

Ans:

*Conclusion 1*: **Parent Categories Vs Outcomes**-Most crowd funding campaigns were held under ‘Theatre’ parent category, and minimum were held under Journalism.

From all campaigns, 76.3% of total campaigns were held only in US.

*Conclusion 2*: **Sub\_categories Vs Outcomes-** Most campaigns in sub\_categories were held under plays (34.4%).

Most successful and unsuccessful campaigns were also held under plays itself.

Least campaigns were held under World music subcategory which is only 3 and all of them held in US.

*Conclusions 3:* **Outcomes Vs Months-** Most successful campaigns were held in June & July month.

Month of December depict least successful campaigns and rise in failed and cancelled campaigns.

Month of August depicts plunge in successful campaigns and same time rise in failed campaigns.

1. What are some limitations of this dataset?

Ans:

Limitations of datasets is the nature of campaigns, if it was online, or in person and if it was widely held in specific geographical areas within the country.

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

Ans:

We can create table:

Name Vs Outcomes: including possibilities calculating the success rate of each campaign holder name. That can show if different categories campaign were held under same name.

Name Vs Percent\_Funded: which can show the relationship between each name and the funds received.

Outcomes Vs Year: This can relate if there were any relations between outcomes and years they were held in.

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

|  |  |  |
| --- | --- | --- |
|  | Successful | failed |
| Mean\_backers | 851 | 586 |
| Median\_backers | 201 | 114.5 |
| Min\_backers | 16 | 0 |
| Max\_backes | 7295 | 6080 |
| Variance\_backers | 1603374 | 921575 |
| Std\_deviattion\_backers | 1266 | 960 |

Ans: As mean is greater than median, the distribution is positively skewed. Median best summarizes the data. The mean could be affected by outliers.

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

Ans: The successful campaigns show more variability than unsuccessful campaigns.

|  |  |  |
| --- | --- | --- |
|  | Successful | failed |
| Mean\_backers | 851 | 586 |
| Median\_backers | 201 | 114.5 |
| Min\_backers | 16 | 0 |
| Max\_backes | 7295 | 6080 |
| Variance\_backers | 1603374 | 921575 |
| Std\_deviattion\_backers | 1266 | 960 |
|  |  |  |

The standard deviation values depict the variations in each campaign. If we have a look at the graph below. It proves the huge difference between values.