**Crowdfunding conclusion**

1. The data presents the outcome of crowdfunds over a period of 10 years. From the data, it can be seen that the success of a crowdfund is not determined by the duration of the crowdfund is run. It can be seen that some successful crowdfund is achieved in 3 days and some failed after a period of a month. From here we can conclude that how successful a crowdfund is determined by the interest backers.

For example, zero interest in the crowdfund will lead to a failed crowdfund. High interest means backers more willing to donate more and crowdfund more likely to be success.

1. From the chart we can see that the category theatre and sub-category play has the most crowdfunding compared to the other categories. This indicates that most backers have higher interest in theatre plays, regardless of how successful the crowdfunding will be. Theatre category has the highest success and failed count.
2. From the date conversion chart, we can conclude that most backers will crowdfund a category mostly at the start of the year between February – May. This further can be seen that when the successful drops around July and August, the failed crowdfund increase. This could indicate that at that July and August would not be the best time to start a crowdfund.
3. The charts do not indicate why some crowdfunds are cancelled. Unless there measures to indicate reasons, relativity of cancelled crowdfunds to success or failed crowdfunds can’t be determine. From the data, we can see that some cancelled crowdfunds last for 3 days with over 50% funded. And the chart, it shows that the increase of successful crowdfunds also see an increase of cancelled crowdfunds. This could indicate external factors that determines if a crowdfund get cancelled and not the numbers of backers.

**Statistical Analysis -Bonus**

These statistical data help give you an insight on your data. Mean is the average of the data. backers. Median however is the middle value of a sorted list of values. This is a measure of central tendency that is not influence by extreme values.

Variance is the spread of values around the mean. High variance indicates the values in the data are spread widely and low suggest otherwise. Standard deviation measures the dispersion of data points around. Hight standard deviation suggests data points are spread widely from the mean and low suggest data points are clustered around the mean.

For example. The mean for successful is 851.15 and the variance and standard deviation indicates that the value and the data points are widely spread. The median tells me the middle point of the data is 201.