1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
2. What are some limitations of this dataset?
3. What are some other possible tables and/or graphs that we could create?

Answer 1:

The three conclusions are:

A: As per the Cat Pivot (Category Pivot) tab on the excel file, the table shows that the overall success rate is over 50%.

B: The most popular category is theatre with a success rate of 60.23% yet the most the most successful category is music with a success rate of 77.14%

C: Journalism is not suitable for Kickstarter as it shows 0% success rate.

Answer 2:

The data was collected from 2009 to 2017, therefore it might not include recent factors.

There is no data showing how much funds were raised in each week since the campaign was launched.

Answer 3:

The donation and backers are critical elements to the campaign, and this should be shown on another table/graphs.

To make clear of which category could possibly make the most profit, a table/graph with total income versus the target is needed.

To find out which country uses Kickstarter the most, a table/graph with country in count is needed.

**Bonus Statistical Analysis**

|  |  |  |  |
| --- | --- | --- | --- |
| **Summary** | Successful | Unsuccessful (Failed + Cancelled) | Unsuccessful (Failed only) |
| Mean # of Backer | 220.5874464 | 19.49228313 | 17.7098039 |
| Median # of Backer | 71 | 3 | 4 |
| Min Backer | 1 | 0 | 0 |
| Max Backer | 26457 | 1501 | 1293 |
| Variance of Backers count | 830562.7939 | 5243.674102 | 3773.22167 |
| Standard Deviation | 911.3521788 | 72.41321773 | 61.4265551 |

~ Use your data to determine whether the mean or the median summarises the data more meaningfully.

As shown on the above table, the mean could summarize a more meaningful data.

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

The successful campaigns have shown more variability compared to the unsuccessful campaigns, regardless of whether the cancelled campaign included as unsuccessful category.

It makes sense because the successful campaigns had more backers in general. The highest number of backers in a successful campaign was 26,457, whereas the highest number of backers in a campaign is 26,457 whereas the highest number in an unsuccessful campaign was only 1,501. With such significant difference in the max backers between the successful and unsuccessful campaigns, it is evident that there is more variability with the successful campaigns.