Starter Book

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

**-** Music campaigns are the most successful ones. Specifically classical, electronic, metal, pop and rock music.

- All the Journalism campaigns were canceled.

- Publishing campaigns can work very well if they are about nonfiction and Radio & podcast. Otherwise they tend to fail

- While games campaigns have a very low successful rate overall, tabletop focused campaigns have all been successful.

- The Launch date does not seem to have a heavy impact on the outcome of the campaign. Just a small advantage for the month of May in Q2

- The data also shows that the lower the goal of the campaign the higher the chances of it being successful.

**2. What are some limitations of this dataset?**

- This is a sample of a much bigger set of data so it might not represent accurately the full set of data.

- We are lacking information about the cities where these campaigns took place.

- It’s not clear where were these crowdfunding campaigns promoted.

**3. What are some other possible tables and/or graphs that we could create?**

- I looked at the length to try to understand if that has any kind of impact on the outcome. Even though most of the campaigns longer than 9 weeks were successful, the volume is too small to be able to assume a correlation.

- Another interesting table is the table below. Looks like (in general) the successful campaigns were more cautious and moderate than the unsuccessful ones. Both the average and the Median of the failed campaigns have values much higher than the successful ones.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| State | Avg of goal | Max of goal | Min of goal | StdDev of goal | Varp of goal | Median |
| successful | 9867 | 400000 | 1 | 21725 | 471760796 | 3500 |
| failed | 60556 | 10000000 | 2 | 418365 | 174915208432 | 8000 |
| canceled | 517985 | 100000000 | 50 | 5735843 | 32805622883469 | 15000 |

- The Country could have been an interesting analysis but the US has a much higher volume than all the other countries.

**BONUS**

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

- The Median summarizes it better since the data has some very extreme data points. It will give a better representation of the data

**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. Yes, it makes sense. We can see a positive correlation between the number of backers and the success of the campaign. The reason the variance is higher on successful campaign is because there is no cap/limit on how many backers a campaign can have. Hence some of the successful campaigns have as many as 20,000 backers.

On the other side, the lowest number of backers a campaign can have is 0 and we see several of them with unsuccessful campaigns.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Values** | **successful** | **failed** | **canceled** | **live** |
| AVG | 194.4 | 17.7 | 27.3 | 36.0 |
| MAX | 26457.0 | 1293.0 | 1501.0 | 884.0 |
| MIN | 1.0 | 0.0 | 0.0 | 0.0 |
| StdDeviation | 844.3 | 61.4 | 107.8 | 123.8 |
| Variance | 712841.0 | 3773.2 | 11615.1 | 15335.7 |
| Median | 62 | 4 | 2 | 8 |