Kickstarter: Data Exploration and Conclusions

Kickstarter is a perfect place to present your idea to World and raised money to bring this idea to life. Every day new startupers explore Kickstarter to understand what the best case is to start they campaigns. In this project, I analyzed more than 4 000 campaigns try to uncover hidden trends.

Fig. 1: This chart shows how many campaigns were successful, canceled, failed or still live by category.

Fig. 2: This chart shows how many campaigns were successful, canceled, failed or still live by sub-category

Fig. 3 this chart shows how many campaigns were successful, failed or canceled by month

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. If you look at the chart depicted in Fig.1 you’ll notice that Theater category most popular on Kickstarter. It has the biggest number of successful campaigns as well as failed campaigns. On the second place is Music category which has the most successful campaigns than failed.
   2. If you look at chart depicted in Fig.2, you’ll see that Plays which is sub-category of Theater is more successful than other sub-category, but Rock campaign (sub-category of Music) is always successful.
   3. If you look at chart depicted in Fig.3, you’ll see that May is the best month to start new campaign because it has more successful campaigns than other months.
2. What are some limitations of this dataset?
   1. Data set has a big gap in amount of campaign between each category.
   2. Category Theater looks like outliers
3. What are some other possible tables and/or graphs that we could create?
   1. We get better understanding if we compare categories using percentage instead of amount.
   2. Compare successful, canceled and failed campaign between different countries

## Bonus part: mean vs. median

|  |  |  |
| --- | --- | --- |
| **Column1** | **number of unsuccessful backers** | **number of successful backers** |
| **mean** | 17.71 | 194.43 |
| **median** | 4 | 62 |
| **minimum** | 0 | 1 |
| **maximum** | 1293 | 26457 |
| **variance** | 3773.22 | 712840.99 |
| **st. deviation** | 61.43 | 844.3 |

There is a big difference between mean and median number of backers as well as between minimum and maximum. Maximum numbers of backers show that the data has outliers. In this case median is more meaningful.

