Analysis Report on StarterBook

**Introduction**

StarterBook is an Excel file, which contains a set of data on 4114 projects. These projects come from 21 countries over the worldwide. Each project has information on its goal, pledged, status, main timelines, category, and so on.

This report is going to introduce the strategies which have been used to determine the data. Then it presents some key findings based on analyzing the provided data. It will use relevant tables and graphs to help demonstrate what the data tell. It would be extremely valuable for future project investment and management.

It’s also pointing out the limitations of the dataset and trying to present some possible and meaningful tables and graphs from different perspectives.

**Strategies**

First task is to trim the StarterBook table. This step aims to check up the completeness of the table and the accuracy of data.

The next important work is to convert some data into readable formats in Ms. Excel. For example, the date “deadline” and the date “launched\_at” are timestamp style. A formula =(((A1/60)/60)/24)+DATE(1970,1,1) is used to change them into common dates.

Then, each cell of “state” column is colored according to its status, “successful” as Green, “failed” as Red, and so on. Grouping the data helps understand better. The column of “Category and Sub-Category” is split into two columns as “Category” and “Sub-Category”. This action will make the analysis more detailed and specific.

The next step is to examine the information and look for trends, patterns, and groupings of themes. This phase tries to envision arranged data and present them fully and clearly.

The whole analysis is completely based on the original provided table to approach to the findings, using different formulas and functions of Ms. Excel.

**Three important findings**

**The first finding** is that in the past eight years, from the year of 2009 to 2017, the top field of projects was “Theater/Plays”. “Theater” and “Plays” stood out separately from Categories and Sub-Category, with 1393 and 1066 out of 4114. This result is showed in the Graph 1 and the Graph 2.

Graph 1 The Amounts of Projects by the Categories

Graph 2 The Amounts of Projects by the Sub-Categories

**The second finding** shows project success chance was over 50% whenever goal amount was less than 10000. Surprisingly, when the goal was less than 1000, the successful percentage went beyond 70%. Further, when goal amount increased, the success chance decreased. In another word, bigger goal amount projects tend more likely to fail, as the table 1 presents. Obviously, the canceling possibility of a project increases along with the higher goal amount.

|  |  |  |  |
| --- | --- | --- | --- |
| Goal | Percentage Successful | Percentage Failed | Percentage Canceled |
| Less than 1000 | 71% | 25% | 4% |
| 1000 to 4999 | 66% | 30% | 4% |
| 5000 to 9999 | 53% | 40% | 7% |
| 10000 to 14999 | 48% | 41% | 11% |
| 15000 to 19999 | 47% | 45% | 8% |
| 20000 to 24999 | 42% | 49% | 9% |
| 25000 to 29999 | 40% | 47% | 13% |
| 30000 to 34999 | 39% | 45% | 16% |
| 35000 to 39999 | 47% | 40% | 13% |
| 40000 to 44999 | 49% | 37% | 14% |
| 45000 to 49999 | 29% | 52% | 19% |
| Greater than or equal to 50000 | 19% | 58% | 23% |

Table 1 Project Success Chance on Goal Amount

Therefore, a conclusion could be drawn that a project having smaller amount goal tended to have higher opportunity to success. Oppositely, a bigger goal indicated a higher risk of failure or cancelation. This tendency is illustrated in the Graph 3.

Graph 3 Project Success Chances on Goals

**The third finding** is that most of the projects happened between the year of 2013 and 2017 and it hits the peak in 2015, as the Graph 4 shows. Accordingly, the amounts of successful, failed, and canceled project were much higher than the other time.

It could come to a conclusion that the society provided a suitable environment for hatching projects in these four years. However, more internal and external factors of the projects need to be discovered.

Graph 4 Projects Happening in Years

**Some limitations of this dataset**

* Limited information on the project (except of Goal) restrains from determining more factors which contribute to the success and which cause the failure.
* The majority of the projects are from US (3038 out of 4114, as the Graph 5 displays), which could not fully represent the life of the worldwide projects.

Graph 5 Projects Contribution

**Some other possible tables and/or graphs**

Although this dataset has its own limitations, it provides an amount of information. When combining related categories, more possible tables and graphs can still be produced. Such as the Graph 4 (analysis by year) and the Graph 6:

Graph 6 Success chances on goal amount