Excel Homework 1

By

Pratima Gokhale

Data Analysis and Visualization Nov 2019-May 2020

Contents

[Homework repository Link 2](#_Toc24460447)

[Conclusions 2](#_Toc24460448)

[Limitations 3](#_Toc24460449)

[Other possible tables/graphs 4](#_Toc24460450)

[Figure 1Trend of Successful Projects Based on Average Donations 4](#_Toc24460452)

[Figure 2Pareto chart showing the successful projects verses Backer count 5](#_Toc24460453)

[Figure 3 Trends in popularity of categories 6](#_Toc24460454)

# Homework repository Link

<https://github.com/PGGokhale/Excel_HW1.git>

There are total 3 files in it.

1. README.md
2. PG\_ExcelHW\_1\_Report.doc
3. StarterBook.xls

# Conclusions

Based on the data provided we can draw following conclusions.

1. For a project to be successful it must have funding 100% or more than its goal.
   * We can conclude this by observing the Kickstarter Table by filtering on the “State” column. It is observed that all the Successful projects have 100% or more Percent Funded.
   * Majority of the Failed or Canceled projects are less funded than their Goal.
   * If we filter on the Percent Funded and select all the projects which are funded >=90% then we can see total 242 projects. Out of these 842 projects 827 are successful, 8 are live, 6 are canceled and 1 is failed. We can also see that out of 6 canceled projects 4 are from technology/wearable category. Hence the reason of their cancelation could be something other than funding.
2. The Categories “Music”, Theater” and “Film&Video” are found to be the top three categories of successful projects. There are certain sub-categories which are majority successful and certain categories which are majority unsuccessful. This could be an indication of the popularity of certain sub-category.
   * There are certain sub-categories in which all of the projects are failed e.g. animation, drama, faith, jazz.
   * There are certain sub-categories in which all of the projects are successful e.g. classical music, electronic music, documentary, metal, pop, rock, shorts, television, indie rock
3. The plot based on the launch date verses state suggests that the success drops as the project launched date is increased. That indicates that the likeliness of the project success decreases as the project start is delayed.
4. It is observed from the Outcome based on Goal that the success rate is inversely proportional to the Goal. That is, the projects with low Goal the percentage of successful projects is higher and as the amount of Goal increases the percentage of successful projects decreases.

# Limitations

What are some limitations of this dataset?

* + There is two-thirds of the data from US and hence the conclusions are more appropriate to US than to the rest of the world.
  + There are more data samples from the categories “Music”, Theater” and “Film&Video” and coincidentally these are the categories for which more successful projects are found. Equal samples from the other categories could provide a fair comparison of the data to draw conclusions.

# Other possible tables/graphs

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

1. As seen in the conclusions, it is very important to have full (atleast 100%) funding for a project to be successful. Hence some analysis into the number of backers and average donations might help in deciding how the funds should be raised.
   * Table showing the count of projects in each state for a given range of average donation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| State -> | Successful | canceled | failed | live | Total |
| Average Donation |  |  |  |  |  |
| 1 to $10 | 8 | 21 | 211 | 1 | 241 |
| 10 to 100 | 1583 | 147 | 819 | 29 | 2578 |
| 100 to 500 | 570 | 56 | 190 | 13 | 829 |
| 500 to 1000 | 19 | 5 | 16 | 1 | 41 |
| 1000 to 1500 | 3 | 2 | 3 | 0 | 8 |
| 1500 to 2000 | 1 | 0 | 1 | 0 | 2 |
| 2500 to 3000 | 1 | 0 | 2 | 0 | 3 |
| 3000 to 3500 | 0 | 1 | 1 | 0 | 2 |

Table Statistics of project state based on Average Donations

Figure Trend of Successful Projects Based on Average Donations

The above plot shows that majority of the successful projects had average donation somewhere between $10 to $500.

Table showing the count of projects in each state for a given range of backer count.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| State -> | Successful | canceled | failed | live | Total |
| Backer Count |  |  |  |  |  |
| 1 to 10 | 68 | 144 | 794 | 20 | 1026 |
| 10 to 25 | 338 | 32 | 229 | 12 | 611 |
| 25 to 50 | 518 | 20 | 112 | 6 | 656 |
| 50 to 100 | 528 | 15 | 53 | 3 | 599 |
| 100 to 1000 | 670 | 20 | 54 | 3 | 747 |
| 1000 to 10000 | 61 | 1 | 1 | 0 | 63 |
| 10000 to 20000 | 0 | 0 | 0 | 0 | 0 |
| 20000 to 30000 | 2 | 0 | 0 | 0 | 2 |

Table 1Statistics of project state based on Backer Count

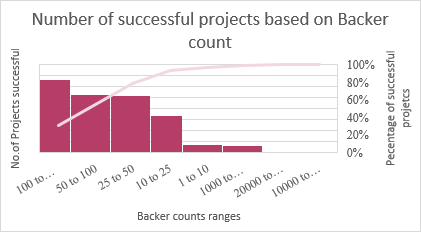


Figure Pareto chart showing the successful projects verses Backer count

The above plot shows that the 80% of the successful projects had backer count between 25 to 100.

1. From the conclusions it is also clear that the projects of certain categories are more successful than the other. To understand this better we could analyze by checking the percentages instead of absolute numbers. We could also check if there is any shift in the trend of popularity of a category based on its year of launch.
   * Percentage of successful, canceled, failed projects based on categories. This will help in identification of the top 3 categories which have highest percentage of successful, canceled and failed projects.
   * Percentage of successful, canceled, failed projects based on sub categories. This will help in identification of top 3 successful sub categories.
   * Plot of number of successful projects verses their year of launch for each of the categories will help in identification of the popularity of certain category as the time changes. This is important when deciding the category of the new project to be launched.

Figure Trends in popularity of categories

The above plot indicates that around

* “Music” is overall popular with a spike at 2012
* Film & Video has a uniform popularity from 2010 till 2016,
* Theater was suddenly popular in 2015 this could be an outlier.
* Technology, games, photography started gaining popularity after 2014