Excel Challenge \_ Kickstarter My Chart

Vijitha Samarasinghe

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

Category analysis:

Based on the category statistical data of 4,000 past Kickstarter projects for the given countries, the highest campaign rate, 33.85% represented by the Theater category with a total of 1,393 campaigns. The Music has shown the second highest campaign rate of 17%. Technology and Film & video had 14.6% and 12.64% campaign percentages respectively. However, over 50% success rate was shown by three categories which were Film & Video, Music and Theater (See: Figure 1). The highest cancellation, 29.67% showed by the category Technology.

Photography and Games showed 0% of canceled projects, however, the campaigns of Photography represented 5.34% of total campaigns. The least successful category out of all nine categories was Journalism, which represented by 0% of successful campaigns and 100% of canceled campaigns.

The success of the Theater category was varied by the geographical location. For example, Denmark conducted more campaigns for Technology than the Theater (See. Figure. 2). The Theater has shown 0% success in Austria, Netherlands, and Mexico. Film & Video category campaigns were not conducted by many countries, and while it had a few successful campaigns in Spain, France, and New Zealand, the majority of successful campaigns were in the U.S.

Subcategory analysis:

Based on the subcategory statistical data available for all the countries, the highest campaign rate of 25.91%, showed by the Plays. The subcategories Rock, Documentary, and Hardware had 0% canceled or failed projects for all the countries. The Science Fiction and Art Books had shown a 0% success rate due to the cancellation of all the campaigns. Animations and Video game campaigns had shown 100% failure rate. The geographical sensitivity affected the successes of subcatergorical campaigns, as well. For example, Australia represented over 50% of the failure projects in Plays. The highest cancellation rate, 6.16% were represented by Web campaigns subcatergory.

The relationship between the launch date and success of a campaign for all nine categories, between 2009 and 2017, represented the number of successful campaigns and decrease in the number of failed and canceled campaigns for campaigns launched in February and April. The highest number of the successful campaigns launched in May. However, the number of the failed campaigns among those launched during May, increased as well. The least successful month to launch a campaign was December. The number of the successful campaigns is lower than the number of the failed campaigns in December. (Figure.4).

Kickstarter data analysis for all the given countries for nine years, concluded that success of campaigned affected by the type of campaign, geographical location and the launch date.

1. What are some limitations of this dataset?

In the given data set, 2009 (included Aug, Jul, May, Nov, Oct, Sep) and 2017 (included Feb, Jan, Mar) were not included the complete year data. Therefore, this is not a true reflection of the complete nine years of data. 2018 and 2019 data were not included in the dataset; therefore, analysis of recent patterns cannot be identified through this dataset.

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

Relationship between campaign goal and success rate

Number of campaigns by the country

Geographical sensitivity (continent) by success/ failure and cancellation rate

The success rate of categories by the country

Relationship between a campaign goal and the amount pledged

Relationship between season (using Lauch\_at) and the campaign success

Analysis of campaign length and its success rate

Bonus section:

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

Explanation:

|  |  |  |
| --- | --- | --- |
| **Backers statistics** | **Outcome\_Successful** | **Outcome\_Unsuccesful** |
|  |
| Mean | 194.43 | 17.71 |  |
| Median | 62.00 | 4.00 |  |
| Minimum | 1 | 4 |  |
| Maximum | 26457.00 | 1293.00 |  |
| Variance | 712840.99 | 3773.22 |  |
| Standard deviation | 844.30 | 61.43 |  |

Mean values of both Outcome\_successful and OutCome\_unsuccessful do not truly summarize the data set due to the larger number of outliers associate with the dataset. But median summarizes data more meaningfully than the mean. (Refer: Scatter plots and Box plots\_ Sheet: Bonus\_Statistical\_Analysis of Excel File: Assignment\_1\_StarterBook)

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

Based on the variance data, more variability showed in the successful campaigns than the unsuccessful campaigns. It is concluded in successful campaigns; the backer count is higher than the unsuccessful campaigns. The failure was due to having a lower count of backers for the campaigns. Success campaigns were pledged way over the goal amount due to a higher number of backers. This is because more realistic goal amount and the larger number of backers.