Samuel Slomowitz

SMU Data Science Bootcamp

12-19-2021

**Delve into Kickstarter**

**Question 1:**

1. The data reveals that theater, music, and film & video are highly successful while publishing, food, and journalism are not.

Table

Description automatically generated

For example, this trend is stable along countries like USA, Great Britain, and New Zeeland.

Graphical user interface, application, table, Excel

Description automatically generatedGraphical user interface, table, Excel

Description automatically generatedGraphical user interface, application, table, Excel

Description automatically generated

Interestingly, the USA, Great Britain, and Canada makes up 92% of the Kickstarter projects locations overall while the USA has a 75.6% success rate when state is broken-down.

Table

Description automatically generated Table

Description automatically generated

1. Secondly, Plays in the category of theater performed well while Rock, Classical Music, Electrical Music, Pop, and Metal performed 100% successfully as a percent of Row Total.

Table

Description automatically generated

Additionally, Video Games, Mobile Games, Places, Nature, People, and Restaurants had 100% failure rate as a percent of Row Total.

Table

Description automatically generated

1. In addition, including only theater, music, and film & video, May and June tend to have a high rate of success while December has the lowest rate of success.

Graphical user interface, table

Description automatically generated

For the most recent year with complete data for the entire year, successful rate was highest for theater, music, and film & video in the months of March, April, and May.

Chart, line chart

Description automatically generated



**Question 2:**

One limitation of this data set is that the data is outdated. The most recent year with complete data is 2016, which is seven years old. A lot of developments and trends could happen to the field and the Kickstarter domain. Additionally, the data did not account for variable like gender, culture, and education level of the Kickstarter project members. These data variables could shed some light on successes and failure rate.

**Questions 3:**

To visualize the Staff picks along the state of the project (successful, failed, etc.) along the columns and values while looking at the percentage of row total, we see that staff picks are more successful (87%) than projects that are not staff picks, which are successful by only 48%.

Table

Description automatically generated

Finally, the history of success (54%) vs. failure (38%) is favorable. With the average donation amount low and little to no financial risk to the project starter, Kickstarter is a great way to raise necessary funds for projects that might never get launched the traditional way. My call to action is to start a theater, music, and film & video project in May and June, especially Plays and Rock, Classical Music, Electrical Music, Pop, or Metal music projects.

**Chart, pie chart

Description automatically generatedChart

Description automatically generated**

Additionally, I would ask for only $10,000 as a goal for funding (Bonus 1).

Chart, line chart

Description automatically generated

**Addendum**

**Bonus Question 2**

The median describes the data better as there are outliers. Since the means are much larger than the median for both successful and failed groups, the data is skewed to the left with large outliers. The variance of successful is much larger and thus, has more variability. This makes sense since popular Kickstarter projects seems to gain a larger variability of backers. A successful Kickstarter project has different qualities that are viewed differently by different types of backers while a failed Kickstarter project has consistent qualities that backers recognize that lead it to a negative outcome. In addition, a two-sample t-test reveals that the means of the outcome are significantly different (p<.0001). Yet, the limitation of conducting a test with the mean is the poor representativeness of the data with the mean compared to the median.

Table

Description automatically generated

**References**

|  |  |  |
| --- | --- | --- |
| **Topic** | **Cells** | **Source** |
| Splitting a cell by delimiter | Column Q and Column R | https://www.howtoexcel.org/tutorials/split-text-by-delimiter/ |
| Converting Timestamp to Date | Column S and Column T | https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html |