Prepared by Leslie Vazquez

Penn Data Analysis & Visualization Bootcamp

kickstarter projects

Data Analysis Report

**Data Analysis: Kickstarter Projects**

**Background**

Over $2 billion has been raised using the massively successful crowdfunding service, Kickstarter, but not every project has found success. Of the more than 300,000 projects launched on Kickstarter, only a third have made it through the funding process with a positive outcome.

**Purpose**

Getting funded on Kickstarter requires meeting or exceeding the project's initial goal. Therefore, the purpose of this data analysis is to understand the trends of 4,000 past campaigns by analyzing these using Microsoft Office Excel.

**Analysis & Conclusions**

*Success Rate*

The first table of statistics analyzes the states of nine categories: Film and Video, Food, Games, Journalism, Music, Photography, Publishing, Technology, and Theater. The states are defined as being: live, canceled, failed, or successful. Based on the graph (Figure 1), there are three categories with over 50% success rate: Music, Film and Video and Theater. Theater has the highest count of successful campaigns (839). However, Music’s success rate is 79% excluding the live campaigns, which exceeds Theater (60%). When analyzing the grand total of success rate of the nine categories, 25% is represented by Music. This category is the second highest with a total count of 700 campaigns. Theater has the highest grand total count of all the campaigns (1393), which represents 34% of the grand total of success rate of the nine categories. However, it is the second highest category in success rate. The rate of canceled campaigns of Theater is 2% resulting in the lowest rate compared to the other categories. Film and Video is the third category with the highest success rate (58%), excluding the live campaigns. The cancel rate for Film and Video is significantly low at 8%. This category had a total count of 520 campaigns placing it in fourth place. Technology is the third highest with a total count of 600 campaigns. Journalism was the least successful of all categories. It had a grand total of 24 campaigns that were 100% canceled.

Figure 1. Category and state

*Country*

The analysis of the success of the nine categories within different countries displays geographical sensitivity and variation. In the United States (Figure 2), Theater has the highest count of successful campaigns (525). However, Music has the highest success rate with 30%, excluding the live campaigns. The campaigns of Theater in the Netherlands and Mexico only reported data for live (36%) and failed (64 %) campaigns. The count of campaigns and success rate of Film and Video is predominant in the U.S. (Figure 3). This category reported success in Australia, Canada, France, Germany, New Zealand, Spain, and United Kingdom. However, the count of the successful campaigns and grand total of campaigns in these seven countries is significantly low.

Figure 2. Category and State in USA

Figure 3. Film and Video state in different countries

*Subcategories*

There are five subcategories in Music that had 100% success rate: Classical Music, Electronic Music, Metal, Rock, and Pop (Figure 4). Rock is the largest subcategory with 260 campaigns, which is 37% of the total campaigns in Music. Of all subcategories, the most successful is Plays in the Theater category with a 66% of successful campaigns. The subcategory Plays represents 77% of all categories under Theater. Film and Video has three subcategories that had 100% success rate: Documentary, Shorts and Television. The largest subcategory is Documentary with 180 campaigns. These represent 34% of the grand total of campaigns in the category Film and Video. There are several subcategories represented on the graph (Figure 4) with 100% canceled or failed campaigns.

Figure 4. Sub-category and state

*Time of the Year*

Based on the graph (Figure 5), the highest number of successful campaigns occurred within the months of February (202), May (234) and June (211). A grand total of 1225 campaigns were reported in the year 2015, which represents the highest number of campaigns within the nine years period (2009-2017). The month of July has the highest reported number of failed and canceled campaigns compared to the other months. Live campaigns were only reported in January, February and March. The reported number of failed campaigns in December is 3% higher than the successful campaigns. This month has the least grand total of campaigns.

Figure 5. Dates and state

*Outcome Based on Goal*

Based on the graph (Figure 6), the percentage of successful campaigns decrease as the goals increase. The percentage of failed campaigns increase as the goals increase. There is a slight decrease in the percentage of failed campaigns when the goal is 40000 to 44999. Then, it increases by 15%.

Figure 6. Outcome based on goal

*Statistical Analysis*

The statistical analysis is based on successful and failed campaigns (Figure 7). The average number of backers for successful campaigns is 194, and 18 for failed campaigns. The median for both campaigns is significantly lower than the average.

In this analysis, the median is of more value than the average for both successful and failed campaigns because the average is affected by the outliers.

The standard deviation indicates more variability in successful campaigns. This variability is related to the number of backers for each campaign. The maximum number of backers for successful campaigns is 26457, which is extreme compared to the failed campaigns (1293).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Successful Campaigns** | |  | **Failed Campaigns** | |
| Mean | 194.43 |  | Mean | 17.71 |
| Median | 62 |  | Median | 4 |
| Minimum | 1 |  | Minimum | 0 |
| Maximum | 26457 |  | Maximum | 1293 |
| Sample Variance | 713167.3791 |  | Sample Variance | 3775.689439 |
| Standard Deviation | 844.4923796 |  | Standard Deviation | 61.44663896 |

Figure 7. statistical analysis

**Limitations**

* The data provided for this analysis does not include metrics to understand how the state (canceled, failed, live, successful) was determined for each campaign.
* There is no metric to understand the difference between a live and successful campaign.
* The data does not include how much time an organization has to complete each Kickstarter campaign.
* The data reported for 2009 excludes the months Jan-April, and for the year 2017 it only includes data from January-March.
* The variable country is too broad. This variable does not allow one to understand the trends based on local marketing.

**Alternative Tables and/or Graphs**

* Display graphs representing each category with its specific subcategories to analyze what makes one category successful.
* Use a pie chart to represent the distribution of the count of each category.
* Display a table with statistical analysis of the most successful categories/subcategories versus the most unsuccessful categories/subcategories.
* Display a graph to compare the duration of each campaign and its outcome.