Week 1 Homework Assignment - Excel Challenge

Contents

[Given the provided data, what are the three conclusions we can draw about Kickstarter campaigns? 1](#_Toc34160411)

[What are some limitations of this dataset? 4](#_Toc34160412)

[What are some other possible tables and/or graphs that we could create? 5](#_Toc34160413)

[**Location based Analysis** 5](#_Toc34160414)

[Pledge based analysis: 5](#_Toc34160415)

[Goal based Analysis: 5](#_Toc34160416)

[Duration based Analysis: 6](#_Toc34160417)

[Bonus Statistical Analysis 7](#_Toc34160418)

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

The Kickstarter data set analyzed included campaigns from 21 countries that fall into one of 9 categories (film & video, food, games, journalism, music, photography, publishing, technology & theater) and 41 sub-categories.

Overall, the dataset contained more successful campaigns (53%) in comparison to failed campaigns (37%). As only 53% of the campaigns reached their funding goal, it is important to know the factors that impact the outcome of the campaigns before launch.

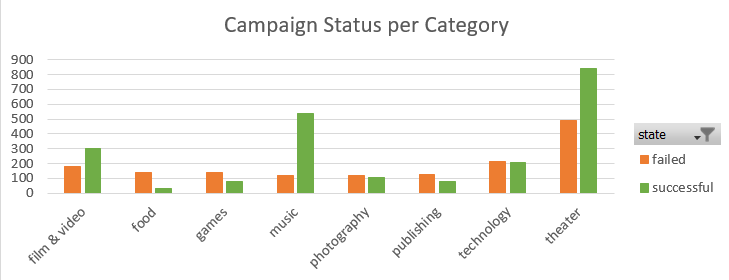
**Summary Statistics:**

The below table shows the summary statistics used to discover trends and indicators that differentiate between successful and failed campaigns.

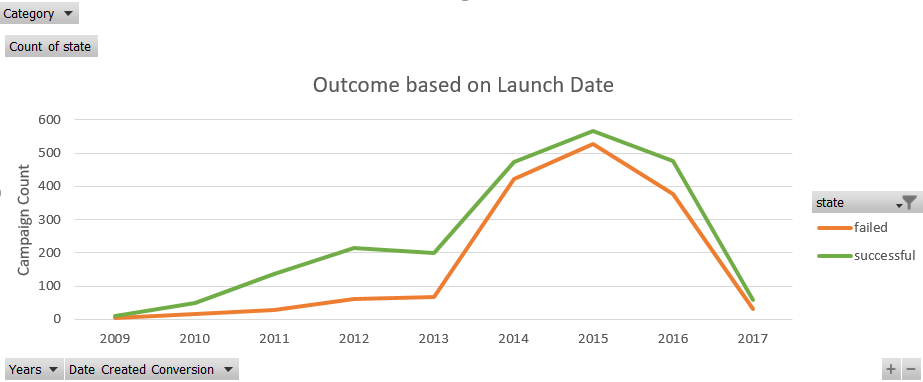
* Successful campaigns have shorter durations (32 days) compared to failed campaigns (35 days).
* Successful campaigns have lower funding goals (averaging $9,867/campaign) than failed campaigns (averaging $60,556).
* Successful campaigns have a higher number of backers (averaging 194 backers/campaign) than failed campaigns (averaging 17.71 backers/campaign).

|  |  |  |
| --- | --- | --- |
| **Campaign Analysis** | **Successful Campaigns** | **Failed Campaigns** |
| Campaigns (Total): | 2185 | 1530 |
| Proportion (%): | 53% | 37% |
| Goal Total ($): | $21,559,364 | $92,651,279 |
| Goal Average ($): | $9,867 | $60,556 |
| Pledged Total ($): | $40,595,733 | $2,838,967 |
| Pledged Average ($): | $18,579 | $1,856 |
| Goal vs. Pledged (%) | 116040% |  |
| Percent Funded (%) |  |  |
| Average Duration (Days): | 32 | 35 |
| Average number of backers: | 194.43 | 17.71 |

**Outcomes per Category:**



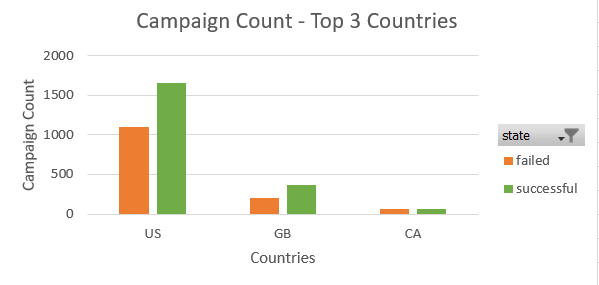
**Outcomes based on Launch dates:**



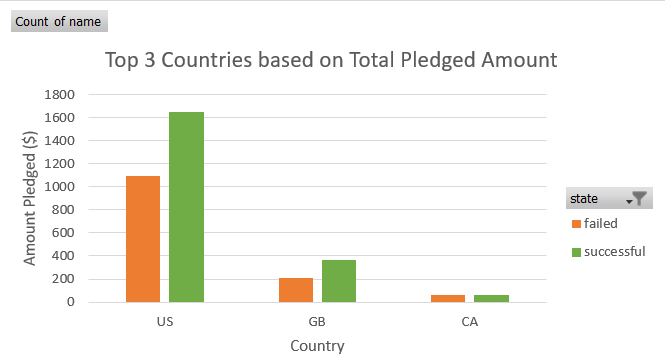
1. **What are some limitations of this dataset?**
   1. Data set only includes campaigns launched between Nov 10, 2009 to Mar 15, 2017.
   2. Data set is limited to 21 countries only.
   3. Data set does not include additional attributes that can help
2. **What are some other possible tables and/or graphs that we could create?**

**Location based Analysis:** Location based analysis can be performed to figure out if certain locations performed better than the rest.

* The data set includes data from campaigns launched in 21 countries.
* Of these 21 countries, maximum number of campaigns were launched in the USA (80%) followed by Great Britain (16%) and Canada (4%).
* USA also had the most successful campaigns (1651).

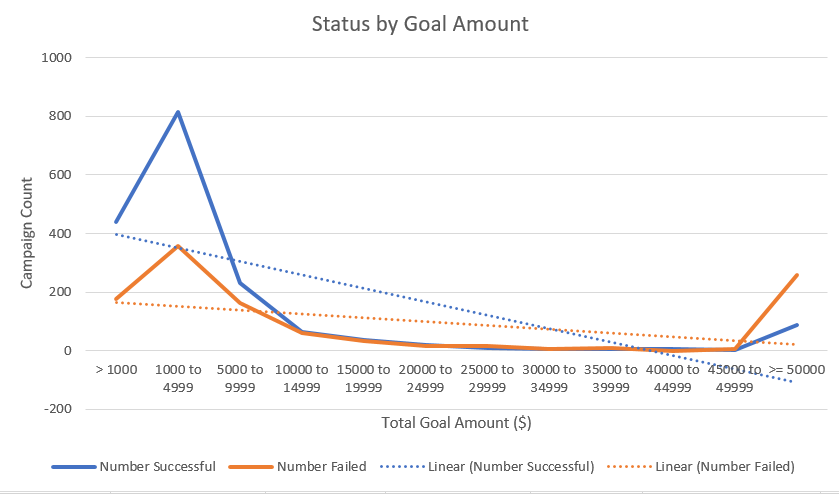


### **Pledge based analysis:**

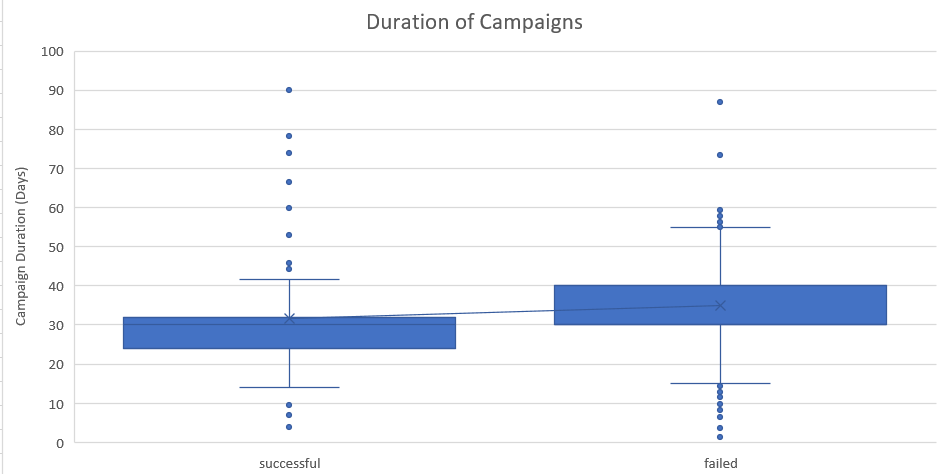


### **Goal based Analysis:**

* Most successful campaigns have a goal of $1000-$4999.
* Campaigns with goals between 10000 to $49999 have a 50% chance of failing.



### **Duration based Analysis:**



**Bonus Statistical Analysis**

1. **Does the mean or median summarize the data more meaningfully?**

Median is more meaningful as there are many outliers (extreme values).

|  |  |  |
| --- | --- | --- |
| **Campaign Analysis** | **Successful Campaigns** | **Failed Campaigns** |
| Average Duration (Days): | 32 | 35 |
| Average number of backers: | 194 | 18 |
| Median number of backers: | 62 | 4 |
| Minimum number of backers: | 1 | 0 |
| Maximum number of backers: | 26457 | 1293 |
| Variance of number of backers: | 712841 | 3773 |
| Standard deviation of number of backers: | 844 | 61 |

1. **Is there more variability with successful campaigns or unsuccessful campaigns? Does this make sense? Why or why not?**

Variance measures how far each number in the set is from the mean and therefore from every other number in the set.

Variance of number of backers is much higher for successful campaigns which