Carter Newman

GW Data Analytics

Homework 1

May 30, 2018

# What are three conclusions we can make about Kickstarter campaigns given the provided data?

In order to facilitate some analysis, I considered what questions I wanted to answer based on the available data. I decided to determine whether Kickstarter projects are successfully funded on a consistent basis with consideration of the project’s geographic location, monetary amount of goal, and category/sub-category against the “state” of the project.

Contents

[What are three conclusions we can make about Kickstarter campaigns given the provided data? 1](#_Toc515467800)

[1. Does geographic location affect popularity of Kickstarter? 2](#_Toc515467801)

[2. Does the monetary amount of a Kickstarter project goal suggest a likelihood of funding success? 3](#_Toc515467802)

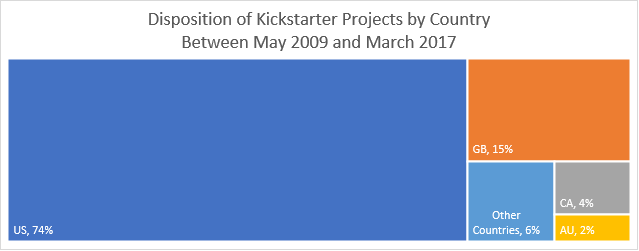
[3. Are certain Kickstarter project categories more successful than others? 4](#_Toc515467803)

[What are some of the limitations of this dataset? 5](#_Toc515467804)

[What are some other possible tables/graphs that we could create? 5](#_Toc515467805)

## 1. Does geographic location affect popularity of Kickstarter?

The United States (US) is clearly the most popular location for Kickstarter projects and is the country in which Kickstarter originated. Great Britain was the second country to adopt Kickstarter (in 2012) followed by Canada (CA), and Australia (AU). Therefore, further investigation needs to be done to determine whether year of adoption drives Kickstarter popularity by country.





Excluding the aforementioned leading countries, the table to the right suggests considerable growth overall for the remaining countries. However, over half of these totals come from the next four leading countries (Germany, Italy, France, and Spain) despite none of them adopting Kickstarter until 2015.

*Note for the sparklines: green indicates high point, red indicates low point.*

**Conclusion 1: Kickstarter popularity based on geographic location is *not* dependent on how early it was adopted, but rather factors outside of this dataset – e.g. country population or native language.**

## 2. Does the monetary amount of a Kickstarter project goal suggest a likelihood of funding success?

*Note: Per the requirements of the assignment (to only use the given data), this does not account for difference in currencies. Thus, this assumes the projects are all based in USD.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Goal (in USD)** | **Cancelled** | **Failed** | **Live** | **Successful** | **TOTAL** |
| Between 1-499 | 5% | 20% | 2% | 74% | 173 |
| Between 500-4,999 | 4% | 29% | 1% | 66% | 1713 |
| Between 5,000-49,999 | 10% | 42% | 1% | 47% | 1779 |
| Between 50,000-499,999 | 21% | 57% | 1% | 21% | 405 |
| Between 500,000-4,999,999 | 32% | 68% | 0% | 0% | 37 |
| 5,000,000+ | 33% | 67% | 0% | 0% | 6 |

The table below suggests project goals under 500 are considerably more successful and that higher amounts become exponentially more likely to fail. Of the 43 projects with goals of at least 500K, none were successful and roughly 2/3rds were cancelled prior to meeting their deadline.

**Conclusion 2: Kickstarter project funding success is dependent on the amount of the goal; the chance of success decreases per ten-fold increase above $500.**

## 3. Are certain Kickstarter project categories more successful than others?

The chart below indicates theater is the most popular type of project to **start**.

|  |  |
| --- | --- |
| **Category** | **Success Rate** |
| theater | 60% |
| music | 77% |
| technology | 35% |
| film & video | 58% |
| publishing | 34% |
| games | 36% |
| photography | 47% |
| food | 17% |
| journalism | 0% |

The table to the right suggests that projects categorized as theater, music, or film & video are more likely than not to succeed.

**Conclusion 3: Theater and music are the most popular project categories both in project initiation and actual funding.**

# What are some of the limitations of this dataset?

Here is an assessment of the dataset:

**Completeness:** Is the dataset representative of all Kickstarter projects?

No. The dataset includes 4,113 individual projects that were open for funding from May 2009 through March 2017. To date, Kickstarter has launched 404,028 projects of all types. Thus, the sample size represents 1% of all Kickstarter projects (if including the 14 months missing from the dataset).

**Uniqueness:** Is each Kickstarter project unique and not repeated?

Yes, each project has an identification number in the “ID” column. However, there is a possibility of a Kickstarter project being cancelled or failing to meet its target and the creator then making another project. Arguably, one would consider these double-counted.

**Timeliness:** Is the dataset of still of timely importance to analysis of Kickstarter in May 2018?

Yes. The dataset is short about 14 months, but still provides nearly eight years of data. Since any resulting analysis would likely include year-over-year trends, the dataset will certainly be sufficient and timely.

**Validity:** Is the dataset consistent in its formats and inputs?

No. The difference in currencies would require further adjustment. Each currency could be converted to US dollars though one would want to account for inflation as well as exchange rates during each year.

**Accuracy:** Do the data items correctly reflect the object or event being described?

Cannot tell for sure. Kickstarter does not readily offer (as far as I can tell) datasets as intensive as this one; they simply offer high-level statistics. However, there are plenty of third-party sources that scrape and then sell data from Kickstarter.

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

* Identifying most commonly used words in the blurbs of successful projects (likely need to break up by category, if not sub-category as well).
* Are successful projects often funded by many small donations or a few large donations?
* What kind of projects receive more than 100% funding?
* How long does it typically take a project to be successfully funded?
* Is there a seasonality effect (comparing months)