

Kickstarter Analysis Proposal

10/9/2020

The Laureates: Farzeen Najam, Olivia Olsher, Vincent Liu, Emile Therrien

Introduction

The introduction section includes an introduction to the subject matter you're investigating, the motivation for your research question (citing any relevant literature), the general research question you wish to explore, and your hypotheses regarding the research question of interest.

Introduction:

Question of Interest:

Our Hypothesis: The more money the project asks for, the less successful it will be in terms of getting funding. Tech projects could most likely be funded.

Data description

In this section, you will describe the data set you wish to explore. This includes a description of the observations in the data set, a general description of the variables in the data set, and a description of how the data was originally collected (not how you found the data but how the original curator of the data collected it).

In this section, you will describe the data set you wish to explore. This includes a description of the observations in the data set, a general description of the variables in the data set, and a description of how the data was originally collected (not how you found the data but how the original curator of the data collected it). This data set contains 15 variables and 378,661 observations, where each observation is one kickstarter project.

The categorical variables include the name of each project, the category of each project (music, narrative film, restaurant, etc.), a broader category of each (food, film, publishing, etc.), the crowdsourcing currency, the state of each project (failed, successful, or cancelled), and the country of origin for each project.

The numerical data are the project's ID number, the monetary goal for each, the money pledged to each project, how many backers of each project; there are three numerical variables that are not self-explanatory, `usd_pledged`: conversion in US dollars of the pledged column (conversion done by kickstarter), `usd_pledge_real`: conversion in US dollars of the pledged column (conversion from Fixer.io API), and `usd_goal_real`: conversion in US dollars of the goal column (conversion from Fixer.io API).

There are also two date columns, one for the project launch and the other for the crowdsourcing deadline.

The data were collected from Kickstarter Platform, to be used by data scientists to possibly model whether or not a project will be successful or not when it is launched.

Glimpse of data

```
library(tidyverse)

kickstarter = read_csv("data/ks-projects-201801.csv")
glimpse(kickstarter)
```

```

## Rows: 378,661
## Columns: 15
## $ ID <dbl> 1000002330, 1000003930, 1000004038, 1000007540, 10...
## $ name <chr> "The Songs of Adelaide & Abullah", "Greeting From ...
## $ category <chr> "Poetry", "Narrative Film", "Narrative Film", "Mus...
## $ main_category <chr> "Publishing", "Film & Video", "Film & Video", "Mus...
## $ currency <chr> "GBP", "USD", "USD", "USD", "USD", "USD", "USD", "...
## $ deadline <date> 2015-10-09, 2017-11-01, 2013-02-26, 2012-04-16, 2...
## $ goal <dbl> 1000, 30000, 45000, 5000, 19500, 50000, 1000, 2500...
## $ launched <dtm> 2015-08-11 12:12:28, 2017-09-02 04:43:57, 2013-01...
## $ pledged <dbl> 0.00, 2421.00, 220.00, 1.00, 1283.00, 52375.00, 12...
## $ state <chr> "failed", "failed", "failed", "failed", "canceled"...
## $ backers <dbl> 0, 15, 3, 1, 14, 224, 16, 40, 58, 43, 0, 100, 0, 0...
## $ country <chr> "GB", "US", "US", "US", "US", "US", "US", "US", "U...
## $ `usd pledged` <dbl> 0.00, 100.00, 220.00, 1.00, 1283.00, 52375.00, 120...
## $ usd_pledged_real <dbl> 0.00, 2421.00, 220.00, 1.00, 1283.00, 52375.00, 12...
## $ usd_goal_real <dbl> 1533.95, 30000.00, 45000.00, 5000.00, 19500.00, 50...

```