LUCAS

Decoding Kickstarter Success

- Thandanvittil Krishnankutty Jaju Abhinav [U2223031L]
- Lucas Ng [U2220046K]

Team 3

What is Kickstarter?

LUCAS

KICKSTARTER

Kickstarter is a crowdfunding platform

It allows individuals or groups to launch creative projects and raise funds from a community of backers

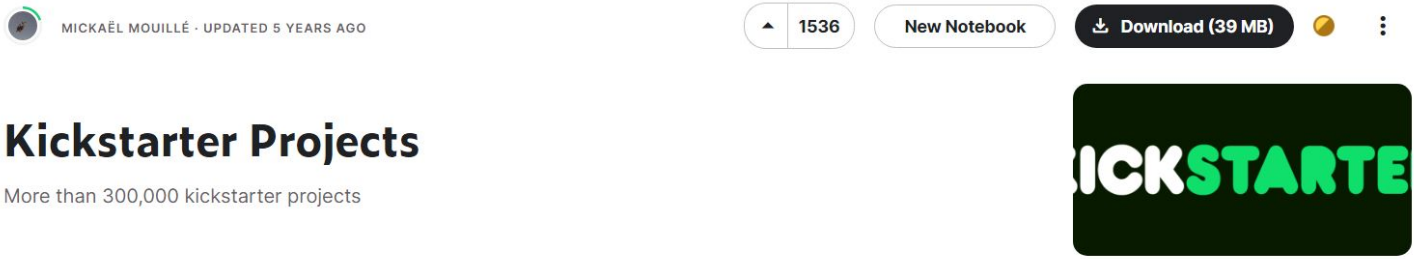
Creators set a funding goal and a deadline and Backers pledge money to help the project reach its goal

The Dataset

We will be making using of the Kickstarter Projects Dataset by MICKAËL MOUILLÉ from Kaggle.

It consists of data from 2016 and 2018. For this Project we will be making use of the 2018 dataset. Specifically the “ks-projects-201801.csv” file.

Link to the Dataset on kaggle



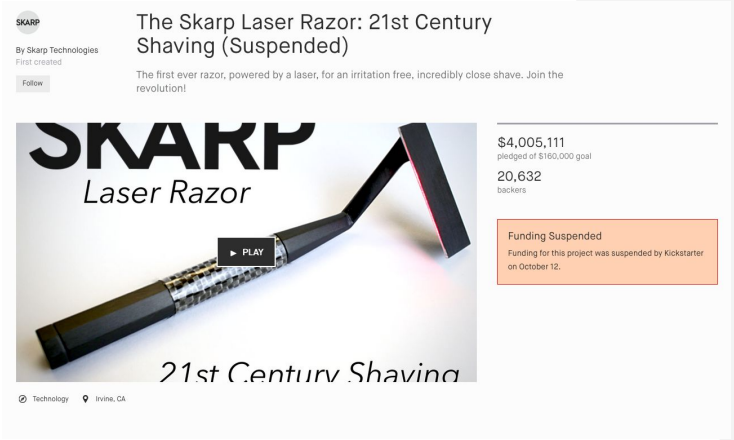
So, Why Kickstarter?

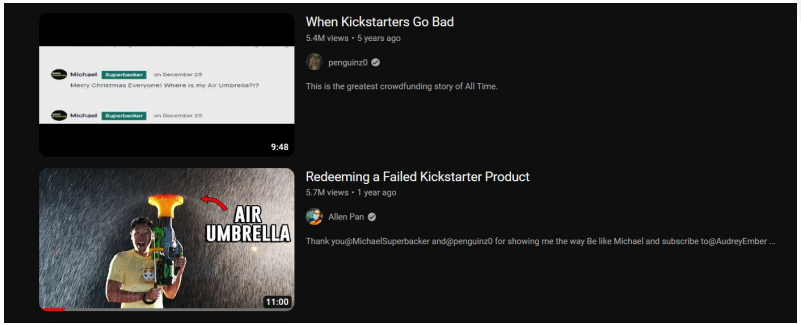
LUCAS

Motivation

Kickstarter campaigns are investments and pose a risk to the investor, similar to stocks in the stock market.

There have been many instances of campaigns being either scams or being cancelled, with the investors not getting anything in return



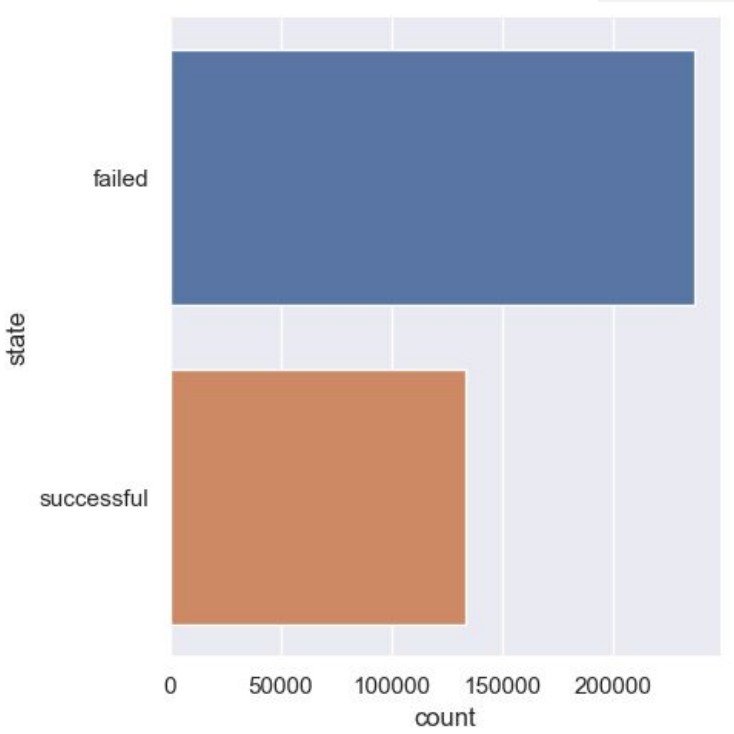


Motivation

We decided to investigate and we can simply see from an initial analysis on the data set.

Failure rate > Success rate for completed Kickstarter campaigns

Can we find out what campaigns are doom to fail and are unable to deliver their product or promise to the investor?



LUCAS

Preparing the Data

- Removing Rows with NULL values
- Standardizing the states
- Removing unnecessary Columns
- Removing Rows with undefined values
- Removing Rows with Duplicate values

Data Cleaning

LUCAS

Preparing Data

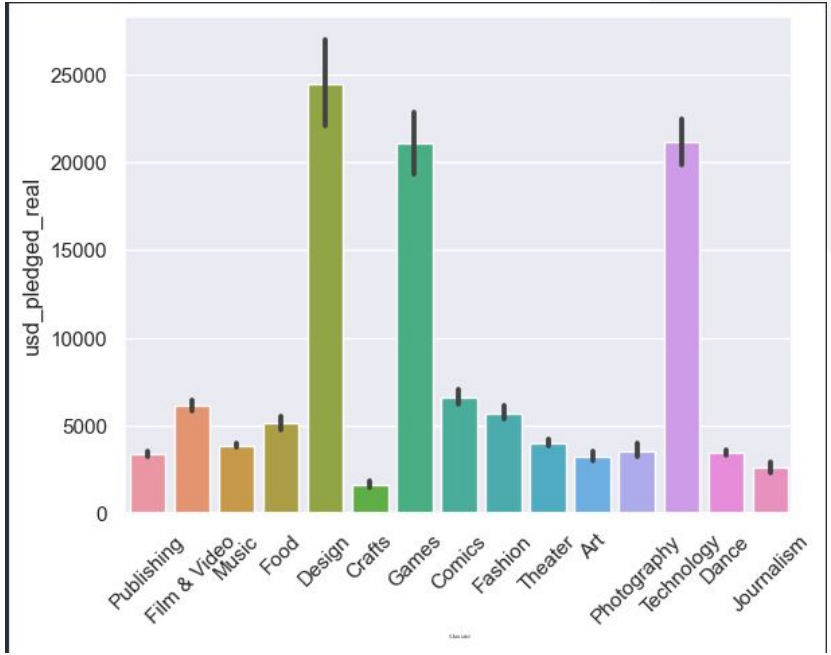
	name	category	main_category	deadline	goal	launched	state	backers	country	usd_pledged_real	usd_goal_real
0	The Songs of Adelaide & Abullah	Poetry	Publishing	9 10 2015	1000.0	2015-08-11 12:12:28	failed	0	GB	0.0	1533.95
1	Greeting From Earth: ZGAC Arts Capsule For ET	Narrative Film	Film & Video	1 11 2017	30000.0	2017-09-02 04:43:57	failed	15	US	2421.0	30000.00
2	Where is Hank?	Narrative Film	Film & Video	26 02 2013	45000.0	2013-01-12 00:20:50	failed	3	US	220.0	45000.00
3	ToshiCapital Rekordz Needs Help to Complete Album	Music	Music	16 04 2012	5000.0	2012-03-17 03:24:11	failed	1	US	1.0	5000.00
4	Community Film Project: The Art of Neighborhoo...	Film & Video	Film & Video	29 08 2015	19500.0	2015-07-04 08:35:03	failed	14	US	1283.0	19500.00

- Reorganise ‘state’ of the project, to be either Successful or Failed
- Further down, we will extrapolate duration of project

LUCAS

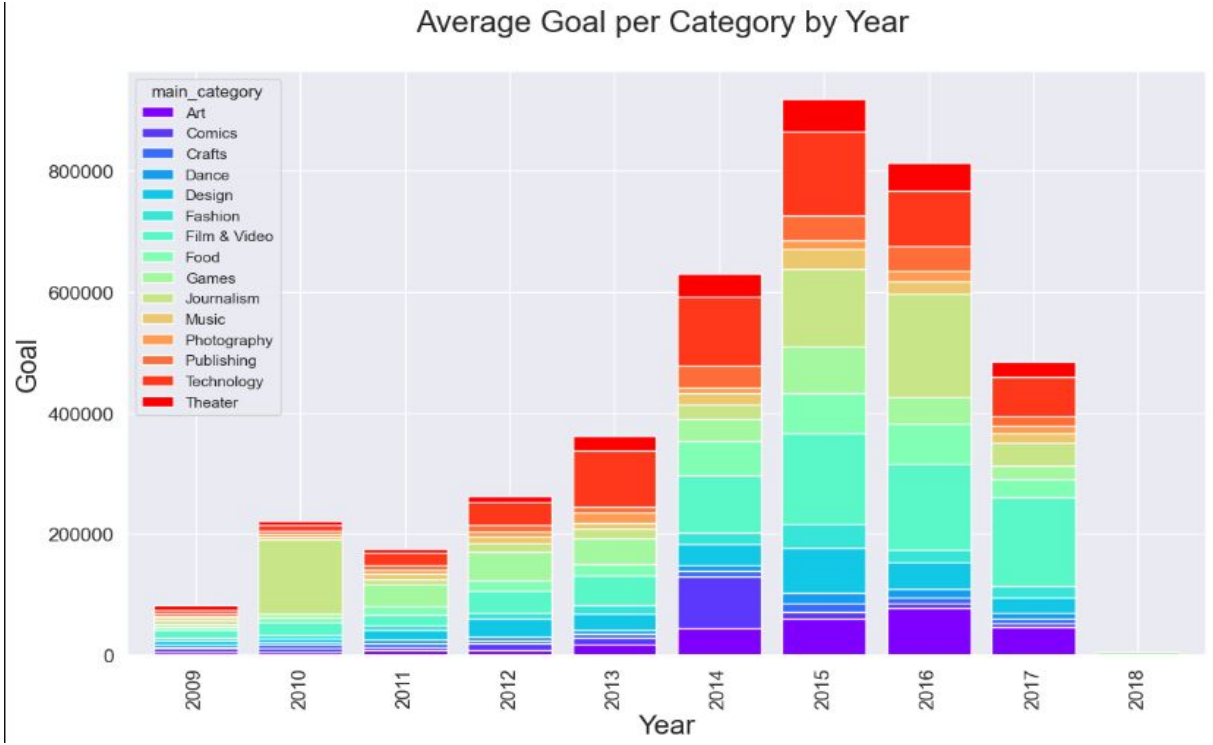
Initial Insights

- Which main categories have higher pledges?
- Design
- Technology



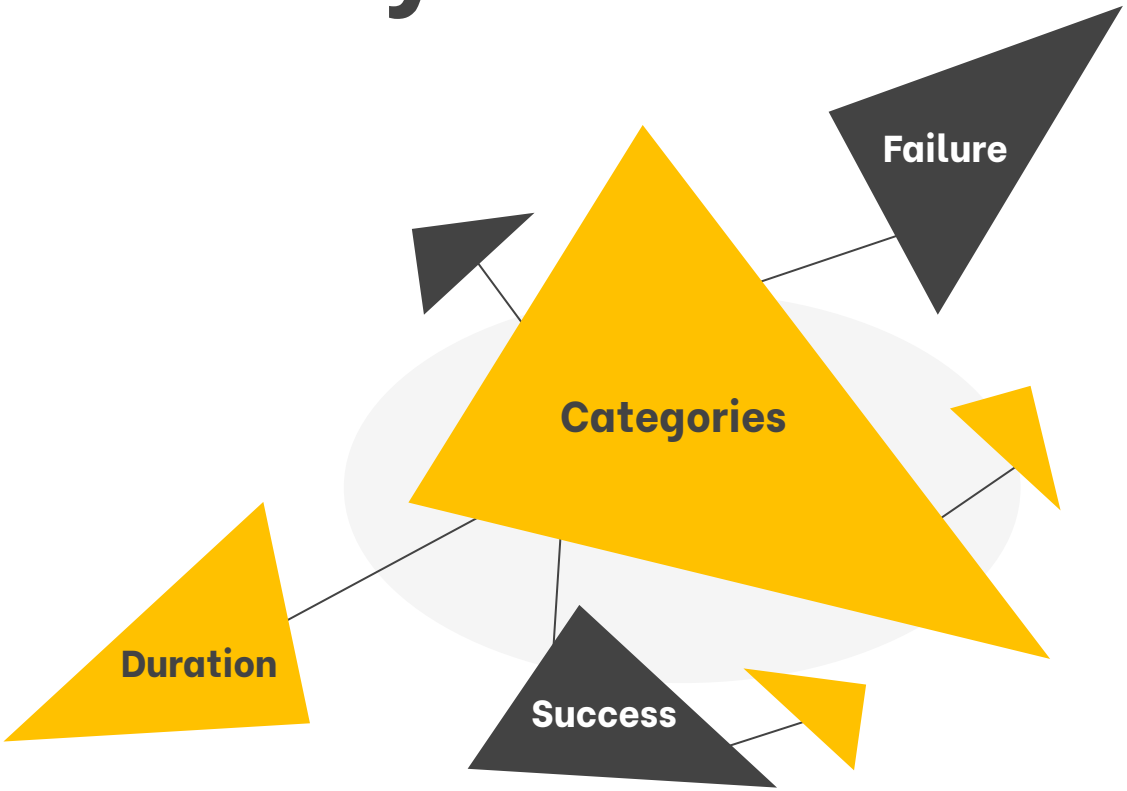
LUCAS

Initial Insights



Core analysis

- Logistic Regression
- Linear Regression



Logistic Regression

Using our root problem statement, we want to have a good way of measuring the chances of a campaign failing or succeeding with the attributes from our dataset.

Predictors

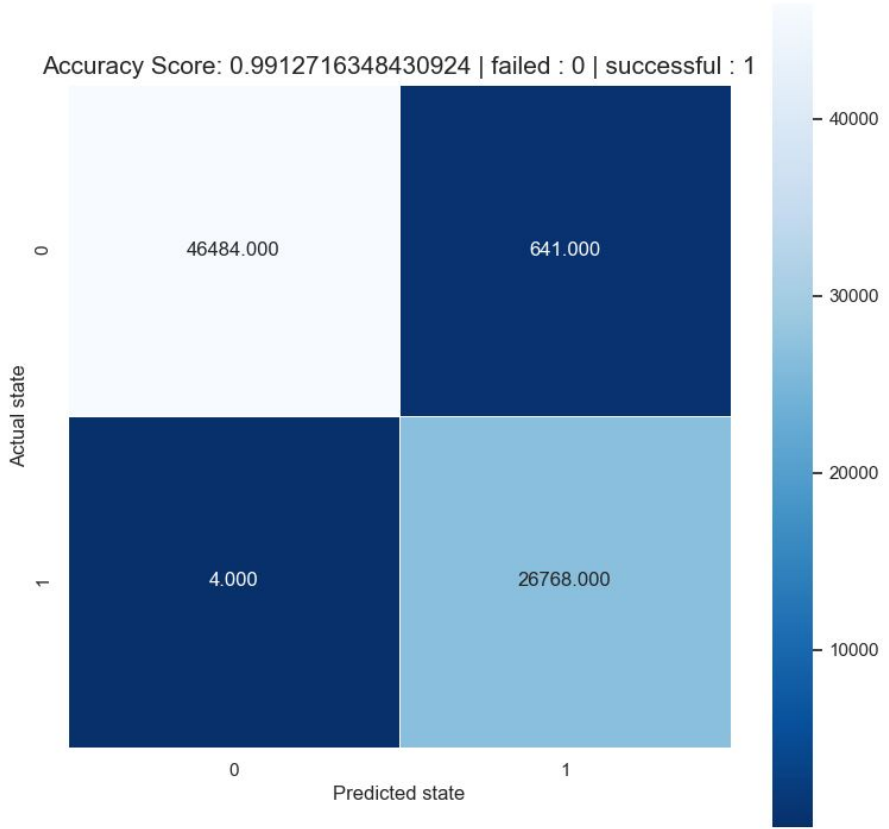
- usd_goal_real
- backers
- duration
- usd_pledged_real

Logistic Regression

Predictors

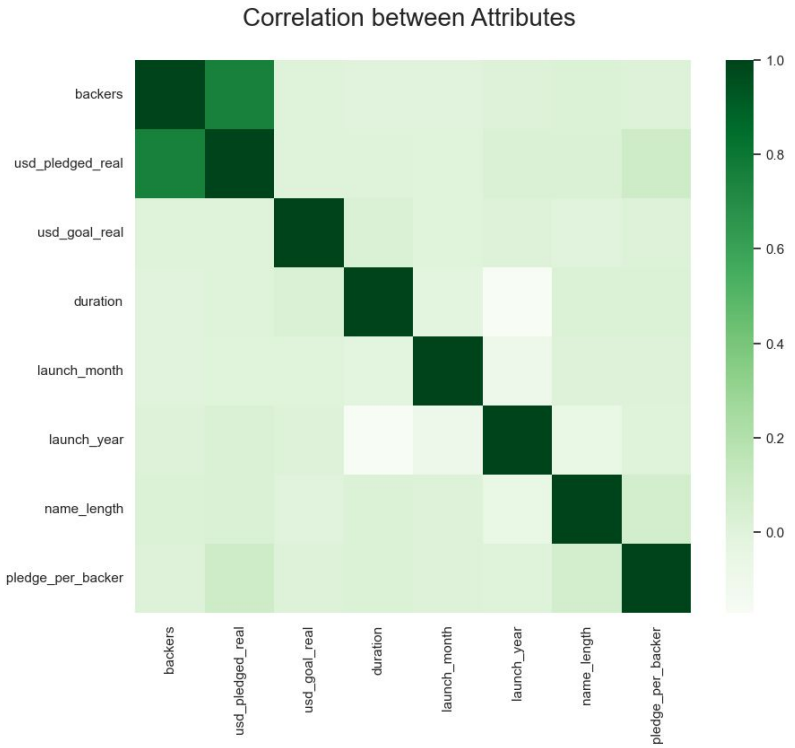
- usd_goal_real
- backers
- duration
- usd_pledged_real

Accuracy: 99.13%



How does amount pledged per backer relate to success rate?

- Linear correlation between usd_pledged and number of backers
- Observe the trend of number of backers vs usd_pledged



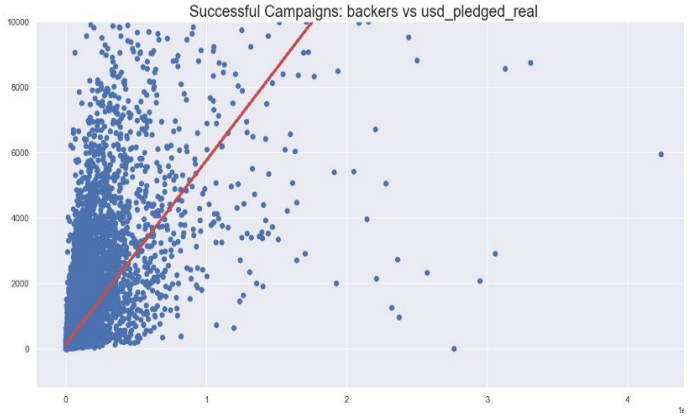
Success vs Failed

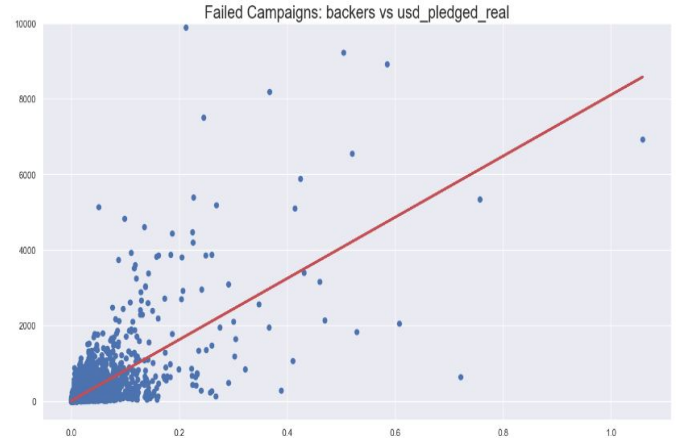
Number of backers per pledged dollar:

Successful > Failed

Total pledged dollar amount:

Successful ~ 3x Failed

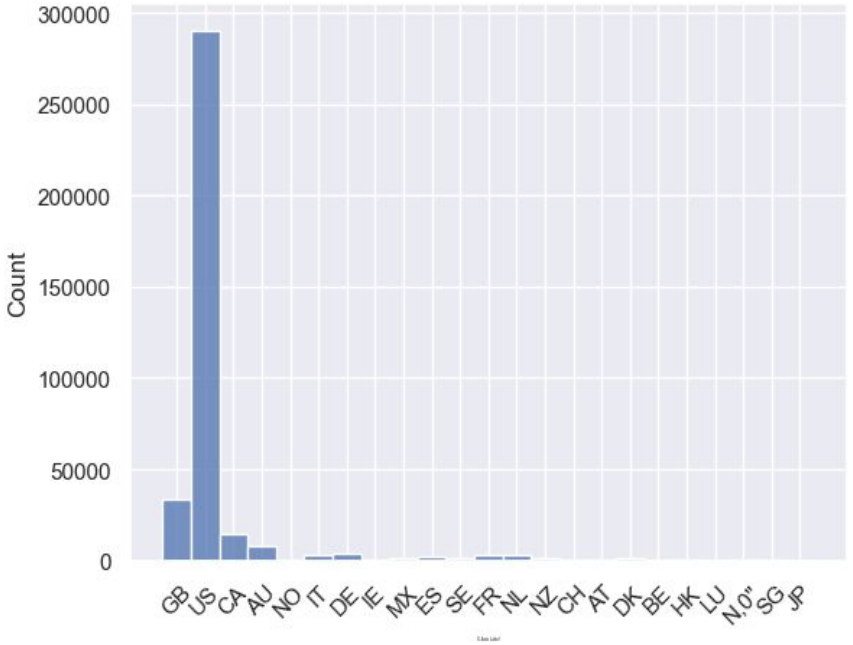




Insight and Conclusion

Potential Bias

- Overwhelming majority of kickstarter projects in this dataset started in US
- Everything taken in USD



Conclusion

The problem is a classification task. This involves using the selected predictors to predict which state a campaign is likely to head towards.

Logistic Regression, along with the following predictors, can be used to predict the campaign's state:

- usd_goal_real
- duration
- backers
- usd_pledged_real


THANK
YOU


END