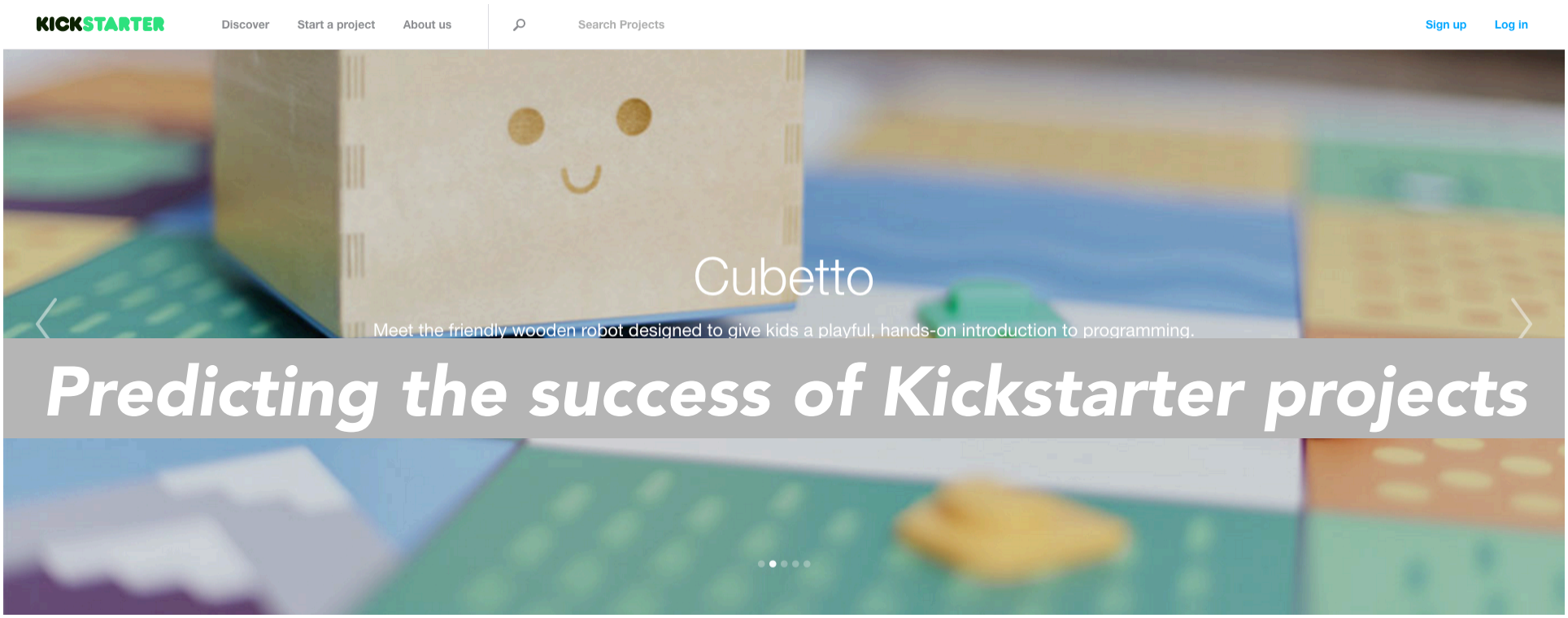
Final Project 1

Don Chung



Problem:

There are thousands of kickstarter projects and many other similar websites that provide crowdfunding platforms for innovators and creators. Problem is… you never know if the project you support is going to end up being fully funded and successful or not.

Data:

I will be using data scrapped from the Kickstarter website provided by Web Robots, which has data from April 2014 through January 2016 of all projects on Kickstarter.

It has data from funding goal to whether it has a video on the page, name of the project, location, number of days open. Basically whatever is available on the project website along with the metadata.

Hypothesis:

Certain categories of projects will increase the chances of success of a project (i.e. technology vs. crafts).

Also, the amount of media (pictures and video) will increase the likelihood of funding as well as the goal amount.

Predicting who will win March Madness

Problem:

Determining who will win March Madness based on the regular season performance as well as the match-up of the teams.

Data:

Kaggle has provided data of all teams who have been in March Madness since 1996 through 2014.

Predicting who will win March Madness

Hypothesis:

For the most part, most of the winners will be predicted based on seeding. However, the upsets and the Cinderella stories can be predicted based on the difficulty of the regular season schedule (measured by number of ranked opponents), regular season record, as well as any streaks coming into March Madness

Predicting the category of crime in SF

Problem:

Determining what type of crime will occur given the location and time within SF's neighborhood

Data:

Kaggle has provided the past 12 years of crime data and I will be using this to train a model to predict