

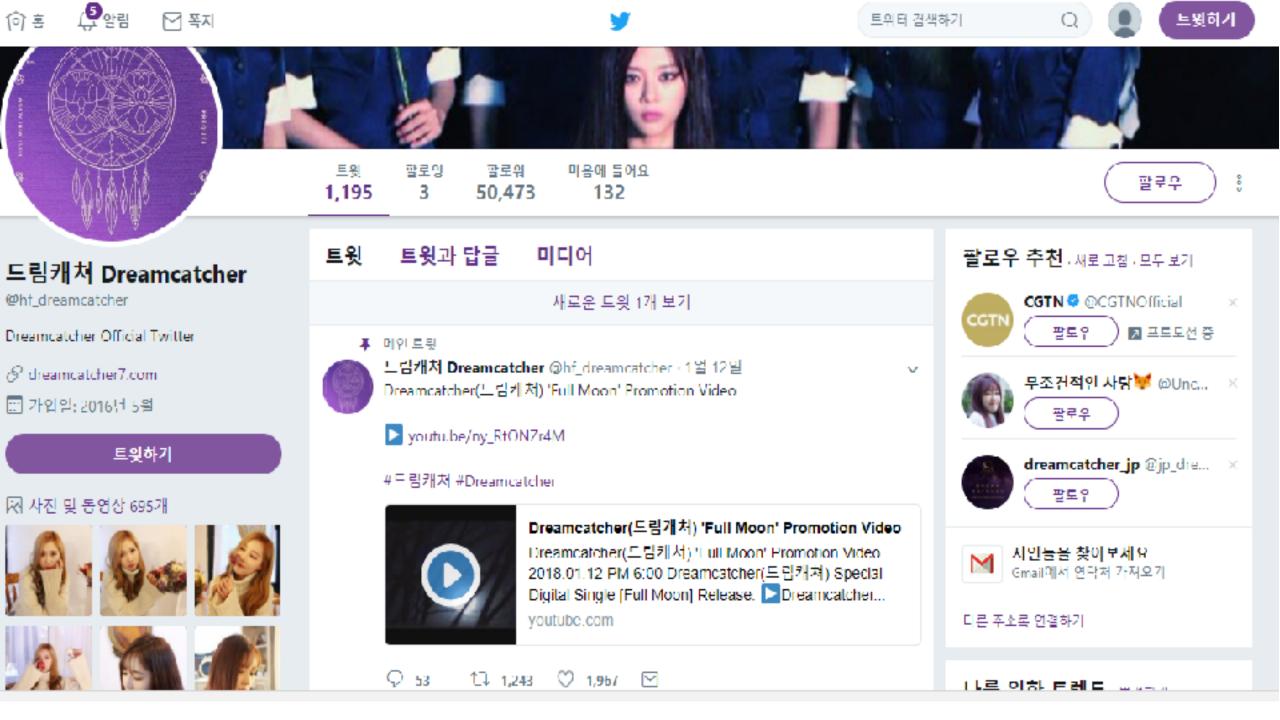
Celebrity's Popularity Prediction Part 1: Profile Celebrities and their Fans

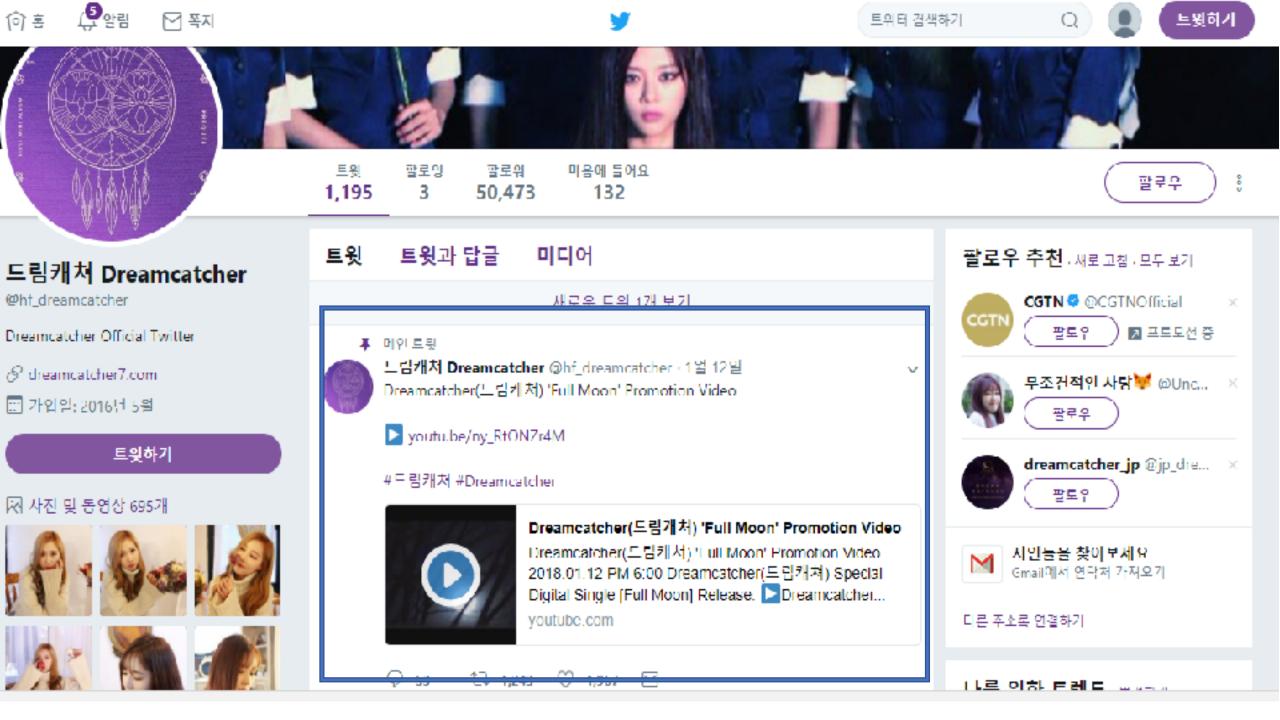
Data

Profiling

Prediction

Data





Json Type Data

```
{ 'contributors': None,
'coordinates': None,
'created at': 'Thu Feb 01 09:01:19 +0000 2018',
'entities':
     {'hashtags': [{'indices': [115, 120], 'text': '드림캐쳐'},
          {'indices': [121, 134], 'text': 'Dreamcatcher'}],
'symbols': [],
'urls': [{'display url': 'youtu.be/wBoWT5DHTdk',
          'expanded url': 'https://youtu.be/wBoWT5DHTdk',
```



1/ Today we're sharing our vision for the future of the Twitter API platform!cards.twitter.com/cards/18ce53wg...

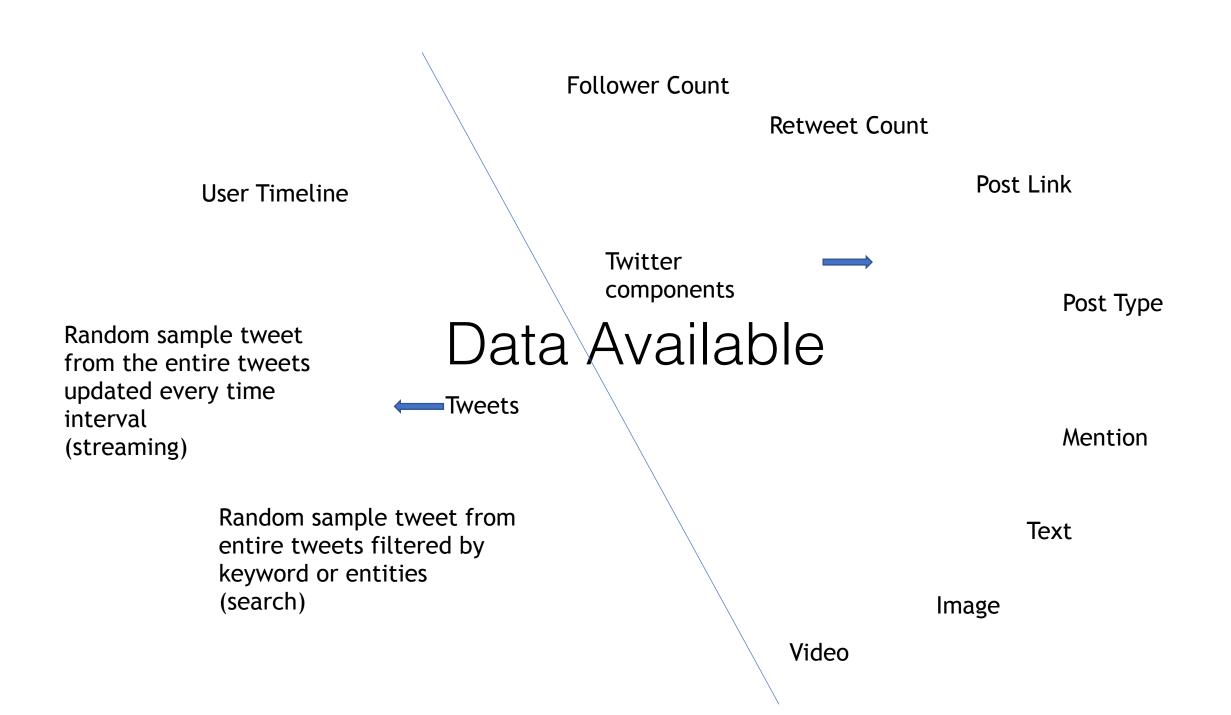
11:24 PM - Apr 6, 2017

Making it easier for you to innovate, build, and scale on Twitter.



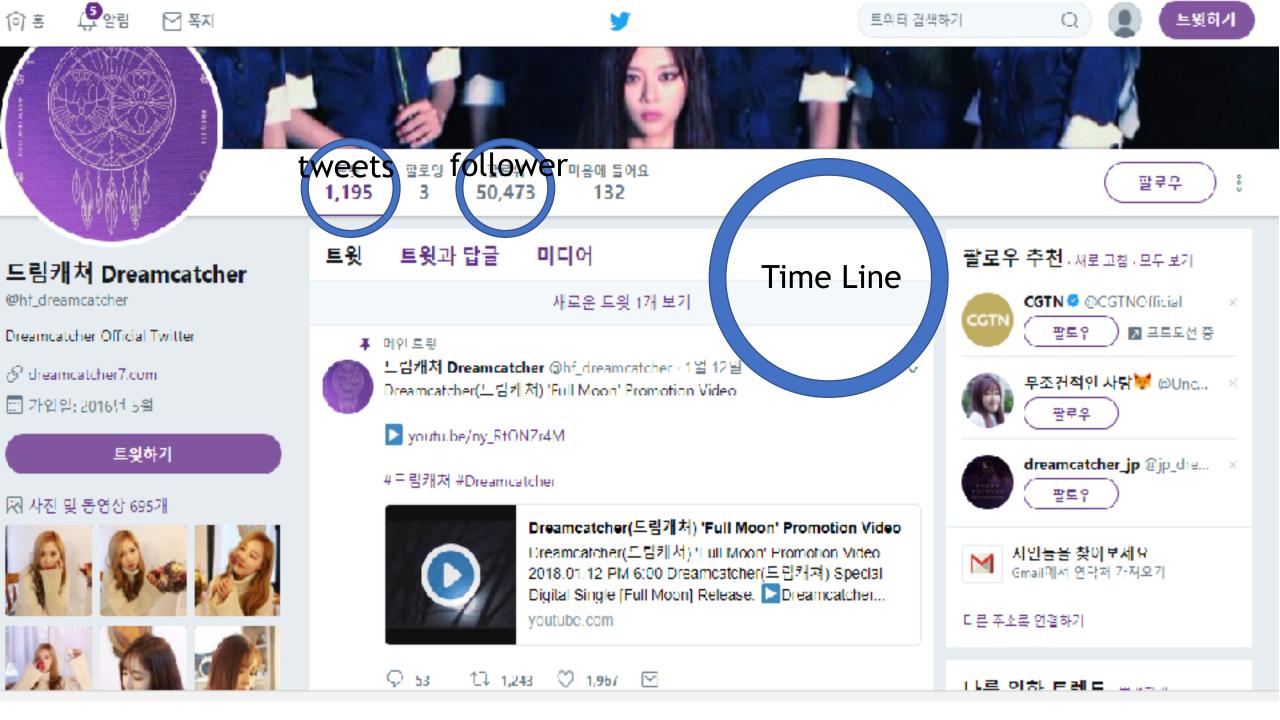
Building the Future of the Twitter API Platform

Tweet object "created at": Thu Apr 06 15:24:15 +0000 2017", # date this tweet is created at "id": 850006245121695744, # tweet specific id "id_str": "850006245121695744", # tweet specific id string "text": "1/ Today we're sharing our vision for the future of the Twitter API platform!nhttps://t.co/XweGngmxIP", # text of this tweet "user": {}, # use who posted this tweet "entities": {} } # entities parsed out from the text: #someone, @someone



Profiling

Create Your Own 'Ratio' for Analysis (Idol)



Ratio1: Participation rate of fans in spreading and regenerating contents

Proportion of followers retweeting tweets

Sum of retweet counts of all the tweets on the time line of DreamCatcher

Number of Tweets * Number of Followers

Find different patterns or ratios between two comparing idol groups divided by 'Ratio' (idols)

Ratio 2: Distinguish core fans with purchase power from light fans

Proportion of tweets ____ with specific textx ____

Number of tweets with the text 'concert' or 'goods'

Number of tweets with mention (#DreamCatcher, @...)

Create Your Own 'Ratio' for Analysis (fans)

Ratio1: Participation rate of fans in spreading and regenerating contents

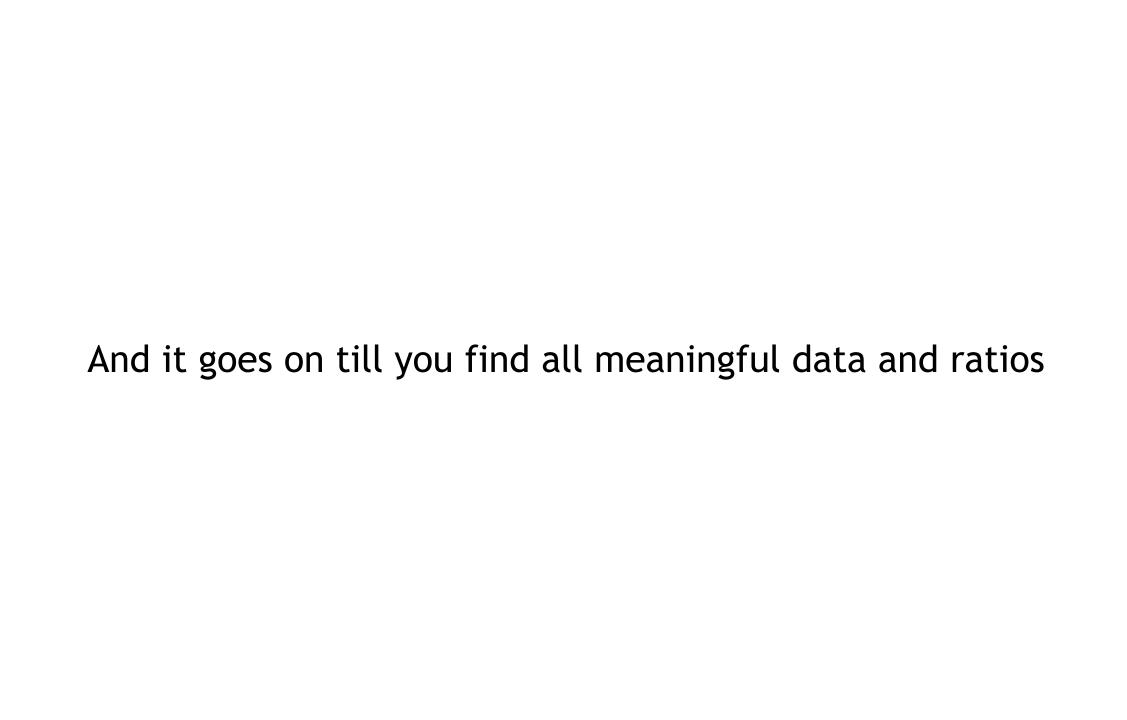
Proportion of followers retweeting tweets

Sum of retweet counts of all the tweets on the time line of DreamCatcher

Number of Tweets * Number of Followers

Find different patterns or ratios between two comparing groups of fans divided by 'Ratio' (Fan)

Model



Feed it to your model for purpose

In my case: Popularity Prediction

Mongodb

Tools Tweepy

Pandas