

What is an API?

An Application Programming Interface, or API, is a standard set of commands that allow programs to interact with a specific application or website. For instance, there are APIs for Facebook, Google Maps, and the New York Times. You can find a list of commercial APIs here (<http://www.programmableweb.com/apis/directory>), or a list of US government APIs here (http://catalog.data.gov/dataset?q=-aapi+api+OR++res_format%3aapi#topic%3ddevelopers_navigation).

As a data scientist, you might use an API to collect data (for example, by fetching crime statistics for a neighborhood from a government website), or to publish data (for example, by automatically sharing photographs to Facebook).

Twitter

Twitter (<https://twitter.com/>) is a social network which allows members to post photos and brief status updates or "tweets". Twitter has only two privacy settings: users must choose between sharing all of their tweets with everyone, or only showing tweets to a single list of approved followers. Most Twitter users choose to make their tweets public. Thus, tracking Twitter allows us to see millions of people's comments on current events and their daily lives.

Here are a couple of famous science-related tweets.

The Most Retweeted Tweets



CERN 
@CERN

#CMS: "we have observed a new boson with a mass of 125.3 ± 0.6 GeV at 4.9 sigma significance." Thunderous applause. #Higgs #ICHEP2012

Reply Retweet Favorite Storify More Assign To
HootSuite

4,750 RETWEETS 811 FAVORITES

9:53 AM - 4 Jul 12

Burson-Marsteller Burson-Marsteller Being More



Philae Lander 
@Philae2014

Follow

Back to work! I'm now drilling into the surface of #67P...
I'll give you updates as soon as I can! #CometLanding

9:03 AM - 14 Nov 2014

1,633 RETWEETS 1,311 FAVORITES

Twitter allows users to categorize their tweets using hashtags, brief phrases introduced by the pound sign (#). Twitter users may post under a hashtag to make a point, as when women posted photos with the tag #ILookLikeAnEngineer (<https://twitter.com/hashtag/looklikeanengineer>) to encourage greater diversity in engineering. Users also use hashtags for emphasis, as in the hashtag #ILoveMyCat (<https://twitter.com/hashtag/ilovemycat>).

Twitter's API allows programmers to gather data from Twitter or to post to the site. Many people have created bots, which post to Twitter automatically when certain conditions are satisfied. For example, a company might create a bot that retweets praise of its product. Some entertaining bots are Big Ben (https://twitter.com/big_ben_clock), which tweets on the hour every hour (Greenwich Mean Time), Accidental Haiku (<https://twitter.com/accidental575>), which shares links to tweets in the 5-7-5 syllable pattern of a haiku, and Museum Bot (<https://twitter.com/MuseumBot>), which shares pictures from the Metropolitan Museum of Art.

We'll gather data from public Twitter posts, and analyze it using the techniques we've learned in this course.