In [1]:
```
#!pip install twython
```

In [2]:
```
from twython import Twython
CONSUMER_KEY='5NMrcgtsT3bTAyGsFpioCzDVJ'
CONSUMER_SECRET='MnBe0GqR2joiRiXCbajv2lm0LD5knmPfXxi77bjelSnZROYPrG'

twitter = Twython(CONSUMER_KEY,CONSUMER_SECRET)
# search for tweets containing the phrase "data science"

```

In [3]:
```
twitter
```

Out[3]:
```
<Twython: 5NMrcgtsT3bTAyGsFpioCzDVJ>
```

In [4]:
```
dir(Twython)
```

...

In [5]:
```
help(Twython.search)
```

```
Help on method search in module twython.endpoints:

search(self, **params) unbound twython.api.Twython method
    Returns a collection of relevant Tweets matching a specified query.

    Docs: https://dev.twitter.com/docs/api/1.1/get/search/tweets (http
s://dev.twitter.com/docs/api/1.1/get/search/tweets)
```

In [6]:
```
results = twitter.search(q="data science", count=20)
#language='Eng', geocode='33.6845673,-117.82650490000003,100mi'

type(results)
```

Out[6]:
```
dict
```

In [7]:
```
1 results
```

Out[7]:
```
{u'search_metadata': {u'completed_in': 0.066,
  u'count': 20,
  u'max_id': 983140074874425344,
  u'max_id_str': u'983140074874425344',
  u'next_results': u'?max_id=983139407929618431&q=data%20science&count=20
&include_entities=1',
  u'query': u'data+science',
  u'refresh_url': u'?since_id=983140074874425344&q=data%20science&include
_entities=1',
  u'since_id': 0,
  u'since_id_str': u'0'},
 u'statuses': [{u'contributors': None,
   u'coordinates': None,
   u'created_at': u'Mon Apr 09 00:30:34 +0000 2018',
   u'entities': {u'hashtags': [],
    u'symbols': [],
    u'urls': [{u'display_url': u'twitter.com/i/web/status/9\u2026',
      u'expanded_url': u'https://twitter.com/i/web/status/983140074874425
344',
```

In [8]:
```
1 all = results['statuses']
```

Building a tweets dataframe

In [9]:
```
1 import pandas as pd
2 tweet_df = pd.DataFrame(all)
```

In [10]:
```
1 tweet_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 27 columns):
contributors                  0 non-null object
coordinates                   0 non-null object
created_at                    20 non-null object
entities                      20 non-null object
extended_entities             1 non-null object
favorite_count                20 non-null int64
favorited                     20 non-null bool
geo                           0 non-null object
id                            20 non-null int64
id_str                        20 non-null object
in_reply_to_screen_name       1 non-null object
in_reply_to_status_id         1 non-null float64
in_reply_to_status_id_str     1 non-null object
in_reply_to_user_id           1 non-null float64
in_reply_to_user_id_str       1 non-null object
is_quote_status               20 non-null bool
lang                          20 non-null object
metadata                      20 non-null object
place                         0 non-null object
possibly_sensitive            3 non-null object
retweet_count                 20 non-null int64
retweeted                     20 non-null bool
retweeted_status              18 non-null object
source                        20 non-null object
text                          20 non-null object
truncated                     20 non-null bool
user                          20 non-null object
dtypes: bool(4), float64(2), int64(3), object(18)
memory usage: 3.7+ KB
```

In [11]:
```
1 tweet_df.head()
```

Out[11]:

	contributors	coordinates	created_at	entities	extended_entities	favorite_count	favorite
0	None	None	Mon Apr 09 00:30:34 +0000 2018	{u'symbols': [], u'user_mentions': [{u'id': 29...	NaN	0	Fals
1	None	None	Mon Apr 09 00:30:05 +0000 2018	{u'symbols': [], u'user_mentions': [{u'id': 32...	NaN	0	Fals
2	None	None	Mon Apr 09 00:29:57 +0000 2018	{u'symbols': [], u'user_mentions': [{u'id': 32...	NaN	0	Fals
3	None	None	Mon Apr 09 00:29:56 +0000 2018	{u'symbols': [], u'user_mentions': [{u'id': 74...	NaN	0	Fals
4	None	None	Mon Apr 09 00:29:38 +0000 2018	{u'symbols': [], u'user_mentions': [{u'id': 19...	NaN	0	Fals

5 rows × 27 columns

In [12]:
```
1 #[d['user'] for d in results['statuses']]
2
3 tweet_user = pd.DataFrame([d['user'] for d in results['statuses']])
```

In [13]:
```
1 tweet_user.head()
```

Out[13]:

	contributors_enabled	created_at	default_profile	default_profile_image	description	e
0	False	Mon Apr 20 23:47:46 +0000 2009	True	False	Pragmatic Fiscal Conservative	{u'descr {u'u
1	False	Thu Mar 31 21:15:37 +0000 2011	False	True	Mary's husband, father, grandfather, great gra...	{u'descr {u'u
2	False	Sun May 19 13:42:37 +0000 2013	True	False	Film Junkie. Fantasy Football/Baseball Fanatic...	{u'descr {u'u
3	False	Thu Jan 29 04:43:24 +0000 2009	False	False	Professional Internet breaker. Writer. Woman. ...	{u'url': u'https://t.
4	False	Mon Jun 21 04:15:47 +0000 2010	True	False		{u'descr {u'u

5 rows × 42 columns

In [14]:
```
1 df = tweet_df.merge(tweet_user, left_index=True, right_index=True)
```

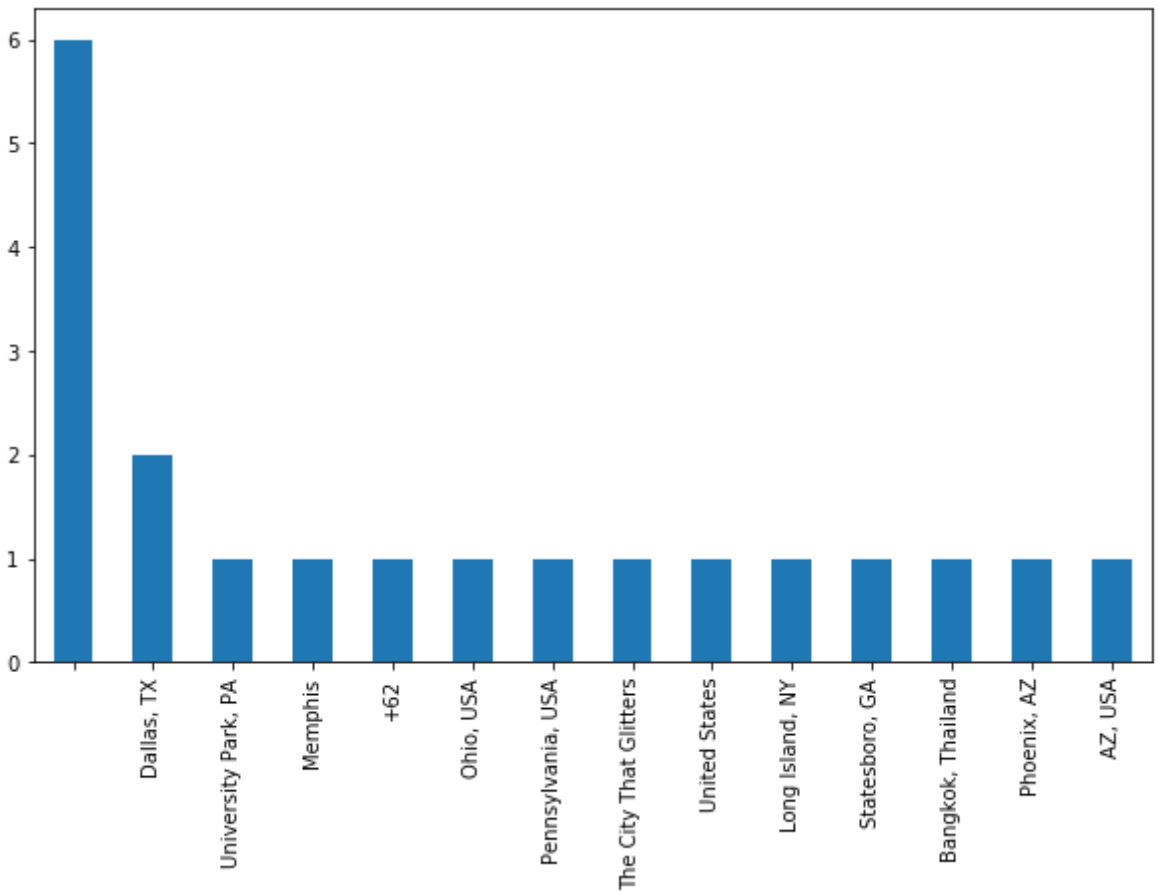
In [15]:
```
1 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 69 columns):
contributors                            0 non-null object
coordinates                             0 non-null object
created_at_x                            20 non-null object
entities_x                              20 non-null object
extended_entities                       1 non-null object
favorite_count                          20 non-null int64
favorited                               20 non-null bool
geo                                     0 non-null object
id_x                                    20 non-null int64
id_str_x                                20 non-null object
in_reply_to_screen_name                 1 non-null object
in_reply_to_status_id                   1 non-null float64
in_reply_to_status_id_str               1 non-null object
in_reply_to_user_id                     1 non-null float64
in_reply_to_user_id_str                 1 non-null object
is_quote_status                         20 non-null bool
lang_x                                  20 non-null object
metadata                                20 non-null object
place                                   0 non-null object
possibly_sensitive                      3 non-null object
retweet_count                           20 non-null int64
retweeted                               20 non-null bool
retweeted_status                        18 non-null object
source                                  20 non-null object
text                                    20 non-null object
truncated                               20 non-null bool
user                                    20 non-null object
contributors_enabled                    20 non-null bool
created_at_y                            20 non-null object
default_profile                         20 non-null bool
default_profile_image                   20 non-null bool
description                             20 non-null object
entities_y                              20 non-null object
favourites_count                        20 non-null int64
follow_request_sent                     0 non-null object
followers_count                         20 non-null int64
following                               0 non-null object
friends_count                           20 non-null int64
geo_enabled                             20 non-null bool
has_extended_profile                    20 non-null bool
id_y                                    20 non-null int64
id_str_y                                20 non-null object
is_translation_enabled                  20 non-null bool
is_translator                           20 non-null bool
lang_y                                  20 non-null object
listed_count                            20 non-null int64
location                                20 non-null object
name                                    20 non-null object
notifications                           0 non-null object
profile_background_color                20 non-null object
profile_background_image_url            15 non-null object
profile_background_image_url_https      15 non-null object
```

```
profile_background_tile               20 non-null bool
profile_banner_url                    13 non-null object
profile_image_url                     20 non-null object
profile_image_url_https               20 non-null object
profile_link_color                    20 non-null object
profile_sidebar_border_color          20 non-null object
profile_sidebar_fill_color            20 non-null object
profile_text_color                    20 non-null object
profile_use_background_image          20 non-null bool
protected                             20 non-null bool
screen_name                           20 non-null object
statuses_count                        20 non-null int64
time_zone                             12 non-null object
translator_type                       20 non-null object
url                                   5 non-null object
utc_offset                            12 non-null float64
verified                              20 non-null bool
dtypes: bool(15), float64(3), int64(9), object(42)
memory usage: 8.8+ KB
```

In [16]:
```
1 import matplotlib.pyplot as plt
2 %matplotlib inline
3 df.location.value_counts().plot(kind='bar', figsize=(10,6))
```

Out[16]: <matplotlib.axes._subplots.AxesSubplot at 0x114c86650>



In [ ]:
```
1
```

In [ ]:
```
1
```

In [17]:
```
1 user_screenname = []
2 user_location = []
3 user_name = []
4 user_tz = []
5 tweet_text = []
6 tweet_tm = []
```

In [18]:

```
for status in twitter.search(q='"data science"', result_type='popular',
    #print status
    #print
    user = status["user"]["screen_name"].encode('utf-8')
    text = status["text"].encode('utf-8')
    tm = status["created_at"]
    loc = status["user"]["location"]
    nm = status["user"]["name"].encode('utf-8')
    tz = status["user"]["time_zone"]
    #tweet_df = tweet_df
    print
    print user, "--", loc, "--", nm,"--", tm, ":", text
    print
    user_screenname.append(user)
    user_location.append(loc)
    user_name.append(nm)
    user_tz.append(tz)
    tweet_text.append(text)
    tweet_tm.append(tm)
```

```
slashdot -- San Francisco -- Slashdot -- Sun Apr 08 18:18:20 +0000 2018 :
Berkeley Offers Its Data Science Course Online For Free https://t.co/NUpO
ek4AbR (https://t.co/NUpOek4AbR)


KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sun Apr 08 05:03:25 +0
000 2018 : 🎉🏆80 Best #DataScience Books That Are Worth Reading: http
s://t.co/LutHjBx2Zz (https://t.co/LutHjBx2Zz) #abdsc #BigData #MachineLea
rning #AI… https://t.co/msivBkNEgM (https://t.co/msivBkNEgM)


Stanford -- Stanford, CA -- Stanford University -- Mon Apr 09 00:00:00 +0
000 2018 : Word embeddings can be used to study historical changes in ste
reotypes, said James Zou, an assistant professor of b… https://t.co/Tdqub
Gf2du (https://t.co/TdqubGf2du)


KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sat Apr 07 18:40:54 +0
000 2018 : The first deadline for the $100,000 #DataSciBowl https://t.co/
gO0vzOWZzS (https://t.co/gO0vzOWZzS) is April 9th!

The next deadline is April… https://t.co/cWVP7yCQZu (https://t.co/cWVP7yC
QZu)


KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sun Apr 08 04:54:54 +0
000 2018 : Comprehensive Repository of #DataScience and #MachineLearning
 Resources, including “15 Deep Learning Tutorials”… https://t.co/59InyQUr
NC (https://t.co/59InyQUrNC)


KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sat Apr 07 18:16:10 +0
000 2018 : #DataScience Interview Guide: https://t.co/jZyoBzaGUd (http
s://t.co/jZyoBzaGUd) #BigData #Analytics #MachineLearning #Algorithms… ht
tps://t.co/GHVb0mCjDq (https://t.co/GHVb0mCjDq)
```

IBMDataScience -- -- IBM Data Science -- Sun Apr 08 21:00:01 +0000 2018 : 💪 💥 Discover this brand-new comic book, "Putting the Statistics in Data Science", starring @jennjshin, analytics su… https://t.co/QlhAN2Myjq (https://t.co/QlhAN2Myjq)

KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sun Apr 08 19:01:21 +0000 2018 : Recommendations to #DataScientists for reducing bias and ensuring fairness in #DataScience: https://t.co/bEKeAB6fX9… (https://t.co/bEKeAB6fX9…) https://t.co/o0WbdeumaU (https://t.co/o0WbdeumaU)

SpirosMargaris -- US | UK | DE | CH -- Spiros Margaris -- Mon Apr 09 00:02:02 +0000 2018 : Gainers and Losers in @Gartner_inc 2018 Magic Quadrant

for #DataScience and #MachineLearning Platforms… https://t.co/ptEp8SQ8SX (https://t.co/ptEp8SQ8SX)

drewconway -- New York, NY -- Drew Conway -- Sat Apr 07 00:04:44 +0000 2018 : Here we go engineering and data science managers, @DHSgov just gave you your ethics interview question... https://t.co/3b9qfVI25H (https://t.co/3b9qfVI25H)

KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sat Apr 07 02:27:48 +0000 2018 : Find all #Jupyter Notebook content for the #Python #DataScience Handbook by @jakevdp here: https://t.co/hG5nhy9brn… (https://t.co/hG5nhy9brn…) https://t.co/Fp5CVQCbG6 (https://t.co/Fp5CVQCbG6)

KirkDBorne -- Booz Allen Hamilton -- Kirk Borne -- Sat Apr 07 02:45:58 +0000 2018 : 5 Free #DataScience Books: https://t.co/4MO9gDGp1e (https://t.co/4MO9gDGp1e) #abdsc #BigData #Statistics #MachineLearning #DataScientists… https://t.co/keH5zbDkNJ (https://t.co/keH5zbDkNJ)

SASsoftware -- Cary, NC -- SAS Software -- Sun Apr 08 18:10:01 +0000 2018 : Attending #SASGF? Now's your chance to win a FREE e-Learning seat for the SAS Academy for Data Science! Visit the… https://t.co/k8SmveRtyY (https://t.co/k8SmveRtyY)

WorldBankPSD -- Washington, DC -- World Bank PSD -- Sun Apr 08 19:15:10 +0000 2018 : Need easy-to-use #datascience tools on #trade and #competitiveness data? #TCdata360 Data Science corner has user-fr… https://t.co/N0Lb6KV6qJ (https://t.co/N0Lb6KV6qJ)

IBMDataScience -- -- IBM Data Science -- Sun Apr 08 12:00:07 +0000 2018 : Learn how machine learning works in Hybrid Data Management, Unified Governance and Data Science. Watch the replay:… https://t.co/JYGBximhlo (https://t.co/JYGBximhlo)

In [19]:

```
import pandas as pd
tweets = pd.DataFrame([user_screenname, user_location, user_name, user_tz
tweets.columns = ['user_screenname', 'user_location', 'user_name', 'user_
```

In [20]:
```
1 tweets
```

Out[20]:

	user_screenname	user_location	user_name	user_tz	tweet_text	tweet_tm
0	slashdot	San Francisco	Slashdot	Pacific Time (US & Canada)	Berkeley Offers Its Data Science Course Online...	Sun Apr 08 18:18:20 +0000 2018
1	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	🎉🏆80 Best #DataScience Books That Are Worth ...	Sun Apr 08 05:03:25 +0000 2018
2	Stanford	Stanford, CA	Stanford University	Pacific Time (US & Canada)	Word embeddings can be used to study historica...	Mon Apr 09 00:00:00 +0000 2018
3	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	The first deadline for the $100,000 #DataSciBo...	Sat Apr 07 18:40:54 +0000 2018
4	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	Comprehensive Repository of #DataScience and #...	Sun Apr 08 04:54:54 +0000 2018
5	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	#DataScience Interview Guide: https://t.co/jZy...	Sat Apr 07 18:16:10 +0000 2018
6	IBMDataScience		IBM Data Science	Pacific Time (US & Canada)	💪💥 Discover this brand-new comic book, "Put...	Sun Apr 08 21:00:01 +0000 2018
7	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	Recommendations to #DataScientists for reducin...	Sun Apr 08 19:01:21 +0000 2018
8	SpirosMargaris	US | UK | DE | CH	Spiros Margaris	Bern	Gainers and Losers in @Gartner_inc 2018 Magic ...	Mon Apr 09 00:02:02 +0000 2018
9	drewconway	New York, NY	Drew Conway	Eastern Time (US & Canada)	Here we go engineering and data science manage...	Sat Apr 07 00:04:44 +0000 2018
10	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	Find all #Jupyter Notebook content for the #Py...	Sat Apr 07 02:27:48 +0000 2018
11	KirkDBorne	Booz Allen Hamilton	Kirk Borne	Eastern Time (US & Canada)	5 Free #DataScience Books: https://t.co/4MO9gD...	Sat Apr 07 02:45:58 +0000 2018
12	SASsoftware	Cary, NC	SAS Software	Eastern Time (US & Canada)	Attending #SASGF? Now's your chance to win a F...	Sun Apr 08 18:10:01 +0000 2018
13	WorldBankPSD	Washington, DC	World Bank PSD	Eastern Time (US & Canada)	Need easy-to-use #datascience tools on #trade ...	Sun Apr 08 19:15:10 +0000 2018
14	IBMDataScience		IBM Data Science	Pacific Time (US & Canada)	Learn how machine learning works in Hybrid Dat...	Sun Apr 08 12:00:07 +0000 2018

In [ ]:
```
1
```