

# Twitter Sentiment Analysis



default ▾

```
import org.apache.spark.streaming._
import org.apache.spark.streaming.twitter._
import org.apache.spark.storage.StorageLevel
import scala.io.Source
import scala.collection.mutable.HashMap
import java.io.File
import org.apache.log4j.Logger
import org.apache.log4j.Level
import sys.process.stringSeqToProcess
```

FINISHED ▶ ⌵ 📖 ⚙️

```
import org.apache.spark.streaming._
import org.apache.spark.streaming.twitter._
import org.apache.spark.storage.StorageLevel
import scala.io.Source
import scala.collection.mutable.HashMap
import java.io.File
import org.apache.log4j.Logger
import org.apache.log4j.Level
import sys.process.stringSeqToProcess
```

```
/** Configures the OAuth Credentials for accessing Twitter */
def configureTwitterCredentials(apiKey: String, apiSecret: String, accessToken: String, accessTokenSecret: String) {
  val configs = new HashMap[String, String] += Seq(
    "apiKey" -> apiKey, "apiSecret" -> apiSecret, "accessToken" -> accessToken, "accessTokenSecret" -> accessTokenSecret
  )
  println("Configuring Twitter OAuth")
  configs.foreach{ case(key, value) =>
    if (value.trim.isEmpty) {
      throw new Exception("Error setting authentication - value for " + key + " not set")
    }
    val fullKey = "twitter4j.oauth." + key.replace("api", "consumer")
  }
```

FINISHED ▶ ⌵ 📖 ⚙️

```

    System.setProperty(fullKey, value.trim)
    println("\tProperty " + fullKey + " set as [" + value.trim + "]")
  }
  println()
}

```

configureTwitterCredentials: (apiKey: String, apiSecret: String, accessToken: String, accessTokenSecret: String)Unit

```

// Configure Twitter credentials
val apiKey = "yCOVn8xTNa08J1iotY8McHWce"
val apiSecret = "LOtYS5XP5IvCbHJZvMA8TTdUhMb2FYTDDBa8IYrwYQJWJaZ8cj"
val accessToken = "273063041-gSEiwLVtMuJKVW5NVExgek7ZUIpHY0enPkohJPK6"
val accessTokenSecret = "5YIMg7b10Hpsv9LVVn0oP06lvu5ACnu4WH5mVRZz1ZPWk"
configureTwitterCredentials(apiKey, apiSecret, accessToken, accessTokenSecret)

```

FINISHED ▶ 🔍 📖 ⚙️

```

apiKey: String = yCOVn8xTNa08J1iotY8McHWce
apiSecret: String = LOtYS5XP5IvCbHJZvMA8TTdUhMb2FYTDDBa8IYrwYQJWJaZ8cj
accessToken: String = 273063041-gSEiwLVtMuJKVW5NVExgek7ZUIpHY0enPkohJPK6
accessTokenSecret: String = 5YIMg7b10Hpsv9LVVn0oP06lvu5ACnu4WH5mVRZz1ZPWk
Configuring Twitter OAuth

```

```

    Property twitter4j.oauth.consumerKey set as [yCOVn8xTNa08J1iotY8McHWce]
    Property twitter4j.oauth.accessToken set as [273063041-gSEiwLVtMuJKVW5NVExgek7ZUIpHY0enPkohJPK6]
    Property twitter4j.oauth.consumerSecret set as [LOtYS5XP5IvCbHJZvMA8TTdUhMb2FYTDDBa8IYrwYQJWJaZ8cj]
    Property twitter4j.oauth.accessTokenSecret set as [5YIMg7b10Hpsv9LVVn0oP06lvu5ACnu4WH5mVRZz1ZPWk]

```

```

import org.apache.spark.streaming.twitter._
val ssc = new StreamingContext(sc, Seconds(2))
val tweets = TwitterUtils.createStream(ssc, None)
val twt = tweets.window(Seconds(60))

```

FINISHED ▶ 🔍 📖 ⚙️

```
import org.apache.spark.streaming.twitter._
ssc: org.apache.spark.streaming.StreamingContext = org.apache.spark.streaming.StreamingContext@86d658a
tweets: org.apache.spark.streaming.dstream.ReceiverInputDStream[twitter4j.Status] = org.apache.spark.streaming.twitter.TwitterInputDStream@214b09c9
tw: org.apache.spark.streaming.dstream.DStream[twitter4j.Status] = org.apache.spark.streaming.dstream.WindowedDStream@5ed0d9f5
```

```
case class Tweet(createdAt:Long, text:String)
tw.map(status=>
  Tweet(status.getCreatedAt().getTime()/1000, status.getText()))
).foreachRDD(rdd=>
  // Below line works only in spark 1.3.0.
  // For spark 1.1.x and spark 1.2.x,
  // use rdd.registerTempTable("tweets") instead.
  rdd.toDF().registerTempTable("tweets")
)
```

FINISHED ▶ ⌘ 📖 ⚙️

defined class Tweet  
warning: there was one deprecation warning; re-run with -deprecation for details

```
tw.print
```

FINISHED ▶ ⌘ 📖 ⚙️

```
ssc.start()
```

FINISHED ▶ ⌘ 📖 ⚙️



Time: 1485476670000 ms

```
StatusJSONImpl{createdAt=Thu Jan 26 19:23:29 EST 2017, id=824774796256886784, text='@JHONPE9 😊', source='<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>', isTruncated=false, inReplyToStatusId=824774288792227844, inReplyToUserId=61558366, isFavorited=false, isRetweeted=false, favoriteCount=0, inReplyToScreenName='JHONPE9', geoLocation=null, place=null, retweetCount=0, isPossiblySensitive=false, lang='und', contributorsIDs=[], retweetedStatus=null, userMentionEntities=[UserMentionEntityJSONImpl{name='Jhon', screenName='JHONPE9', id=61558366}], urlEntities=[], hashtagEntities=[], mediaEntities=[], symbolEntities=[], currentUserRetweetId=-1, user=UserJSONImpl{id=2790861295, name='LA CARECULO', screenName='Yennlachiqui', location='Villavo Colombia', description='De Tetas TÍMIDAS', isContributorsEnabled=false, profileImageUrl='http://pbs.twimg.com/profile_images/809070528749273088/NeG2Sx0a_normal.jpg', profileImageUrlHttps='https://pbs.twimg.com/profile_images/809070528749273088/NeG2Sx0a_normal.jpg', isDefaultProfileImage=false, url='null', isProtected=false, followersCount=2913, status=null, profileBackgroundColor='000000', profileTextColor='000000', profileLinkColor='FF691F', profileSidebarFillColor='000000', profileSidebarBorderColor='000000', profileUseBackgroundImage=false, isDefaultProfile=false, showAllInlineMedia=false, friendsCount=400, createdAt=Thu Sep 04 19:57:17 EDT 2014, favouritesCount=21569, utcOffset=-18000, timeZone='Bogota', profileBackgroundImageUrl='http://abs.twimg.com/images/themes/theme1/bg.png', profileBackgroundImageUrlHttps='https://abs.twimg.com/images/themes/theme1/bg.png', profileBackgroundTiled=false, lang='es', statusesCount=36622, isGeoEnabled=true, isVerified=false, translator=false, listedCount=12, isFollowRequestSent=false, withheldInCountries=null}
```

%sql select \* from tweets where text like '%trump%' limit 10

FINISHED ▷ ✖ 📖 ⚙



| createdAt     | text  |
|---------------|---|
| 1,485,475,781 | RT @MaximusTheRuler: What? #crazy #trump <a href="https://t.co/uYEr57avfr">https://t.co/uYEr57avfr</a>                |
| 1,485,475,794 | RT @VicenteFoxQue: @ajplus @DeniseDresserG trump, I want to see your tax returnsjj Maybe not paying taxes will be e   |
| 1,485,475,798 | RT @JordanUhl: rt if trump needs to stop talking about his inauguration crowd size                                    |
| 1,485,475,803 | @MAGAtump5 Was Hardiman on Trump's list?  |
| 1,485,475,816 | RT @princessjarii: trump wants to build a mexican border but wants Mexico to pay for it..what kinda shit is that      |
| 1,485,475,821 | @realDonaldTrump @MiamiHerald go trump!   |
| 1,485,475,825 | RT @nia4_trump: #climatefacts George Soros paid Al Gore MILLIONS to push junk science. #climatechange is a SCAM!      |
| 1,485,475,828 | RT @NaYaKnoMi: all you celebrities that claim you are against trump re-tweet, rephrase, copy and paste, or even sugge |

```
%sql select createdAt, count(1) from tweets group by createdAt order by createdAt
```

FINISHED ▶ 🔍 📖 ⚙️



| createdAt     | count |
|---------------|-------|
| 1,485,475,668 | 1     |
| 1,485,475,669 | 36    |
| 1,485,475,670 | 37    |
| 1,485,475,671 | 42    |
| 1,485,475,672 | 42    |
| 1,485,475,673 | 36    |
| 1,485,475,674 | 34    |
| 1,485,475,675 | 42    |
| 1,485,475,676 | 28    |

```
def sentiment(s:String) : String = {  
  val positive = Array("like", "love", "good", "great", "happy", "cool", "the", "one", "that")  
  val negative = Array("hate", "bad", "stupid", "is")  
  
  var st = 0;  
  
  val words = s.split(" ")  
  positive.foreach(p =>  
    words.foreach(w =>  
      if(p==w) st = st+1  
    )  
  )  
  
  negative.foreach(p=>  
    words.foreach(w=>  
      if(p==w) st = st-1  
    )  
  )  
  if(st>0)  
    "positive"  
  else if(st<0)
```

FINISHED ▶ ⌕ 📖 ⚙️

```
        "negative"  
    else  
        "neutral"  
    }  
  
sentiment: (s: String)String
```

```
sqlc.udf.register("sentiment", sentiment _)
```

FINISHED ▶ ⌵ 📖 ⚙️

```
res14: org.apache.spark.sql.expressions.UserDefinedFunction = UserDefinedFunction(<function1>,StringType,Some(List(StringT  
ype)))
```

```
%sql select sentiment(text), count(1) from tweets where text like '%trump%' group by sentiment(text)  FINISHED ▶ ⌵ 📖 ⚙️
```



| UDF(text) | count(1) |
|-----------|----------|
| neutral   | 2        |
| negative  | 3        |
| positivie | 2        |

