**import** org.apache.spark.sql.SQLContext  
**import** org.apache.spark.{SparkConf, SparkContext}  
**import** org.apache.spark.SparkConf  
**import** org.apache.spark.SparkContext  
**import** org.apache.hadoop.util  
  
**object** newp1 {  
 **def** main(args: Array[String]): Unit = {  
 **val** conf = **new** SparkConf().setAppName(**"query"**).setMaster(**"local[2]"**).set(**"spark.executor.memory"**, **"1g"**);  
 **val** sc = **new** SparkContext(**"local[2]"**, **"PbSpark"**)  
 **val** sqlContext = **new** SQLContext(sc)  
 **val** RDD = sc.textFile(**"C:\\Users\\MAHESH\\Desktop\\tweets2.json/"**)  
  
  
 **val** RDDQuery = sc.textFile(**"C:\\Users\\MAHESH\\Desktop\\tweets2.json"**).map(\_.split(**","**))  
(query 1)

RDDQuery.take(100).foreach(*println*)  
  
 **val** Data = RDD.filter(line => line.contains(**"protest"**))  
 Data.take(20).foreach(*println*)

(query 2)

**val** stext = **new** SQLContext(sc)  
 **val** DataFrames = stext.jsonFile(**"C:\\Users\\MAHESH\\Desktop\\tweets2.json"**)  
  
 DataFrames.groupBy(**"filter\_level"**).count().show()

(query3)

DataFrames.select(**"text"**).show()

(query4)

}  
}

Join query

**object** newp2 {  
 **def** main(args: Array[String]): Unit = {  
 **val** sc = **new** SparkContext(**"local[2]"**, **"PbSpark"**)  
 **val** sqlContext = **new** SQLContext(sc)  
 **val** tweets = sqlContext.jsonFile(**"C:\\Users\\MAHESH\\Desktop\\tweets2.json"**)  
 tweets.registerTempTable(**"table2"**)  
 **val** tweets1 = sqlContext.jsonFile(**"C:\\Users\\MAHESH\\Desktop\\joinQ.json"**)  
 tweets.registerTempTable(**"table3"**)  
 **val** tempo = sqlContext.sql(**"SELECT table2.id, table3.text FROM table2 FULL OUTER JOIN table3 ON table2.id=table3.id"**)  
 tempo.take(100).foreach(*println*)  
 tempo.save(**"dataframe2"**, **"json"**)  
 sc.stop()

API query

**object** newp3 {  
  
 **def** main(args: Array[String]) {  
  
 *// (1) config work to create a twitter object* **val** cb = **new** ConfigurationBuilder()  
 cb.setDebugEnabled(**true**)  
 .setOAuthConsumerKey(**"3qDlfSB9BbJK5IzGrVhtvju8V"**)  
 .setOAuthConsumerSecret(**"EC7Sfoo2G4m5DvnNLnTxd8R6mxt2iRCuBOOUZi3SHbLwcKxZiu"**)  
 .setOAuthAccessToken(**"828931820-IF3FgxtQkiLqmnFeN7IWDhwXQOmPlILSUsjRURkM"**)  
 .setOAuthAccessTokenSecret(**"xUvPB6NozFpljYrc6C7JMwcGA5oy93BJHplE6e5z7a6OE"**)  
  
 **val** tf = **new** TwitterFactory(cb.build())  
 **val** twitter = tf.getInstance()  
  
 *// (2) use the twitter object to get your friend's timeline* **val** statuses = twitter.getFavorites()  
 System.*out*.println(**"Showing friends timeline."**)  
 **val** it = statuses.iterator()  
 **while** (it.hasNext()) {  
 **val** status = it.next()  
 *println*(status.getUser().getName + **":"** +  
 status.getText());  
 }  
  
 }  
}