

Queries

```
val sqlContext = new org.apache.spark.sql.SQLContext(sc);
```

```
import sqlContext.implicits._
```

```
val tweetstable = sqlContext.read.json("/project/dataset.json");
```

```
tweetstable.registerTempTable("tweetDatatable");
```

```
val q1 = sqlContext.sql("SELECT place.country,count(*) AS count FROM tweetDatatable  
GROUP BY place.country ORDER BY count DESC limit 10");
```

```
q1.show();
```

```
q1.coalesce(1).write.format("com.databricks.spark.csv").save("/project/csv/q1.csv");
```

```
val q2=sqlContext.sql("SELECT substring(user.created_at,1,3) as day,count(*) as count  
from tweetDatatable group by day");
```

```
val q3=sqlContext.sql("SELECT count(*) as count, user.name from tweetDatatable where  
user.name is not null group by user.name order by count desc limit 10");
```

```
val q4 = sqlContext.sql("SELECT substring(user.created_at,5,3) as month, count(user.id)  
from tweetDatatable group by month");
```

```
val q5 = sqlContext.sql("SELECT user.name, max(user.followers_count) as  
followers_count, user.lang FROM tweetDatatable WHERE text like '%football%' group by  
user.name,user.lang order by followers_count desc limit 15");
```

```
val q6=sqlContext.sql("select count(*) as count,q.text from (select case when text like  
'%fcb%' then 'fc barcelona' when text like '%real madrid%' then 'real madrid' when text  
like '%fifa%' then 'fifa' when text like '%premier league%' then 'premier league' when  
text like '%la liga%' then 'la liga' when text like '%arsenal%' then 'arsenal' when text like  
'%chelsea%' then 'chelsea fc' WHEN text like '%manchester%' THEN 'manchester' WHEN  
text like '%tottenham%' THEN 'tottenham' WHEN text like '%psg%' THEN 'psg' else  
'different sports' end as text from tweetDatatable)q group by q.text");
```

```
val q7 = sqlContext.sql("SELECT  
user.verified,user.screen_name,max(user.followers_count) as followers_count FROM  
tweetDatatable WHERE user.verified = false GROUP BY user.verified, user.screen_name  
ORDER BY followers_count DESC LIMIT 15");
```

```
val q8 = sqlContext.sql("SELECT  
user.screen_name,text,retweeted_status.retweet_count FROM tweetDatatable ORDER  
BY retweeted_status.retweet_count DESC LIMIT 20");
```

```
val q9=sqlContext.sql("SELECT user.location,count(text) as count FROM tweetDatatable  
WHERE place.country='United States' AND user.location is not null GROUP BY  
user.location ORDER BY count DESC LIMIT 15");
```

```
val q10 = sqlContext.sql("SELECT user.name ,retweeted_status.text AS  
Retweet_Text,retweeted_status.retweet_count AS Retweet_Count FROM  
tweetDatatable WHERE retweeted_status.retweet_count IS NOT NULL ORDER BY  
retweeted_status.retweet_count DESC limit 10");
```

```
val q11=sqlContext.sql("SELECT substring(quoted_status.created_at,1,3) as  
day,count(text) as count FROM tweetDatatable GROUP BY day");
```

```
val q12 = sqlContext.sql("SELECT substring(user.created_at,27,4) as year,count(*) as  
Count from tweetDatatable where user.created_at is not null group by  
substring(user.created_at,27,4) order by count(1) desc");
```

```
val q13 = sqlContext.sql("SELECT user.name as UserName,user.location as  
loc,text,created_at," + "CASE WHEN text like '%fcb%' THEN 'fc barcelona'"  
+ "WHEN text like '%real madrid%' THEN 'real madrid'" + "WHEN text like  
'%fifa%' THEN 'fifa'" + "WHEN text like '%premier league%' THEN  
'premier league'" + "WHEN text LIKE '%la liga%' THEN 'la liga'" +  
"WHEN text like '%arsenal%' THEN 'arsenal'" + "WHEN text like '%chelsea%'  
THEN 'chelsea fc'" + "WHEN text like '%manchester%' THEN 'manchester'"  
+"WHEN text like '%tottenham%' THEN 'tottenham'" + "WHEN text like  
'%psg%' THEN 'psg'" + "END AS leagueType from tweetDatatable  
where text is not null");
```

```
q13.createOrReplaceTempView("leaguetable");
```

```
val r1 = sqlContext.sql("SELECT UserName,leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='fc barcelona' " + "group by UserName,
```

```
leagueType order by count desc limit 1");
```

```
val r2 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='real madrid' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r3 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='fifa' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r4 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='premier league' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r5 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='la liga' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r6 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='arsenal' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r7 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='chelsea' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r8 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='manchester' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r9 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='tottenham' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val r10 = sqlContext.sql("SELECT UserName, leagueType,count(*) as count  
FROM leaguetable WHERE leagueType='psg' " + "group by UserName,  
leagueType order by count desc limit 1");
```

```
val query = r1.union(r2).union(r3).union(r4).union(r5).union(r6)  
.union(r7).union(r8).union(r9).union(r10);
```