

```

import org.apache.spark.SparkContext
import org.apache.spark.SparkConf
import org.apache.spark.sql.functions._
import org.apache.spark.sql.SQLContext
import org.apache.spark.sql.hive.HiveContext
import org.apache.spark.sql.hive.orc

val stocks = sc.textFile("file:///root/yahoo_stocks.csv")

val header = stocks.first()

val data = stocks.mapPartitionsWithIndex{ (idx, iter) => if (idx == 0)
iter.drop(1) else iter }

data.count()

stocks.count()

import sqlContext.implicits._

case class YahooStockPrice(date: String, open: Float, high: Float,
low: Float, close: Float, volume: Integer, adjClose: Float)

val stocksDF = data.map( rec => rec.split(",")).map(rec =>
YahooStockPrice(rec(0).toString,rec(1).toFloat,rec(2).toFloat,rec(3).t
oFloat,rec(4).toFloat,rec(5).toInt,rec(6).toFloat)).toDF()

stocksDF.registerTempTable("YAH00_STOCKS")

val YAH00_STOCKS = sqlContext.sql("select * from YAH00_STOCKS")

YAH00_STOCKS.write.format("orc").save("file:///root/yahoo")

val hivectx = new HiveContext(sc)

hivectx.sql("CREATE TABLE  yahoo_orc_table (date STRING, open FLOAT,
high FLOAT, low FLOAT, close FLOAT, volume INT, adjClose FLOAT) STORED
AS ORC ")

hivectx.sql("LOAD DATA LOCAL INPATH '/root/yahoo' INTO TABLE
yahoo_orc_table")

```

```
hivectx.sql("select * from  
yahoo_orc_table").collect().foreach(println)
```

```
hivectx.sql("show create table  
yahoo_orc_table").collect().foreach(println)
```