

```
[cloudera@quickstart ~]$ hadoop fs -put /home/cloudera/onlineretail.csv /user/cloudera/
```

```
$ hive
```

```
hive> CREATE TABLE OnlineRetail (
```

```
    InvoiceNo STRING,  
    StockCode STRING,  
    Description STRING,  
    Quantity INT,  
    InvoiceDate STRING,  
    UnitPrice DOUBLE,  
    CustomerID INT,  
    Country STRING)
```

```
ROW FORMAT DELIMITED
```

```
FIELDS TERMINATED BY ','
```

```
STORED AS TEXTFILE;
```

```
hive> LOAD DATA INPATH 'onlineretail.csv' INTO TABLE OnlineRetail;
```

```
//Sum of total sales and average
```

```
hive> SELECT SUM(Quantity * UnitPrice) AS total_sales, AVG(Quantity * UnitPrice) AS avg_sales  
FROM OnlineRetail;
```

```
//Find Order details with maximum cost:
```

```
SELECT InvoiceNo, SUM(Quantity * UnitPrice) AS order_total  
FROM OnlineRetail  
WHERE InvoiceNo IS NOT NULL  
GROUP BY InvoiceNo  
ORDER BY order_total DESC  
LIMIT 1;
```

```
//customer details with maximum order total
```

```
SELECT CustomerID, SUM(Quantity * UnitPrice) AS order_total  
FROM OnlineRetail  
WHERE CustomerID IS NOT NULL  
GROUP BY CustomerID  
ORDER BY order_total DESC  
LIMIT 1;
```

```
//create external table
```

```
//first create a table in hbase and then run this
```

```
hive> CREATE EXTERNAL TABLE OnlineRetailHBase (
```

```
    InvoiceNo STRING,  
    StockCode STRING,  
    Description STRING,
```

```
Quantity INT,  
InvoiceDate STRING,  
UnitPrice DOUBLE,  
CustomerID INT,  
Country STRING)  
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'  
WITH SERDEPROPERTIES (  
    "hbase.columns.mapping" =  
":key,OnlineRetail:stock_code,OnlineRetail:description,OnlineRetail:quantity,OnlineRetail:invoice_date,OnlineRetail:unit_price,OnlineRetail:customer_id,OnlineRetail:country")  
TBLPROPERTIES ("hbase.table.name" = "OnlineRetail");
```