

```
import org.apache.spark.sql._
import org.apache.spark.sql.SQLContext._

import org.apache.spark.sql._
import org.apache.spark.sql.SQLContext._
```

```
val flightData = sqlContext.read.json("/FileStore/tables/55rjz3sz1500302954550/flightData2015.json")

flightData: org.apache.spark.sql.DataFrame = [DEST_COUNTRY_NAME: string, ORIGIN_COUNTRY_NAME: string ... 1 more field]
```

flightData.printSchema

```
root
 |-- DEST_COUNTRY_NAME: string (nullable = true)
 |-- ORIGIN_COUNTRY_NAME: string (nullable = true)
 |-- count: long (nullable = true)
```

flightData.createOrReplaceTempView("FlightData2015")

%sql select \* from FlightData2015

DEST_COUNTRY_NAME	ORIGIN_COUNTRY_NAME	count
United States	Romania	15
United States	Croatia	1
United States	Ireland	344
Egypt	United States	15
United States	India	62
United States	Singapore	1
United States	Grenada	62
Costa Rica	United States	588
Senegal	United States	40

```
val sqlWay = spark.sql("""
SELECT DEST_COUNTRY_NAME, count(1)
FROM FlightData2015
GROUP BY DEST_COUNTRY_NAME
""")
```

sqlWay: org.apache.spark.sql.DataFrame = [DEST\_COUNTRY\_NAME: string, count(1): bigint]

sqlWay.show

DEST_COUNTRY_NAME	count(1)
Anguilla	1
Russia	1
Paraguay	1
Senegal	1
Sweden	1
Kiribati	1
Guyana	1
Philippines	1
Djibouti	1
Malaysia	1
Singapore	1
Fiji	1
Turkey	1
Iraq	1
Germany	1
Jordan	1
Palau	1
Turks and Caicos ...	1

sqlWay.explain

```
== Physical Plan ==
*HashAggregate(keys=[DEST_COUNTRY_NAME#189], functions=[count(1)])
+- Exchange hashpartitioning(DEST_COUNTRY_NAME#189, 200)
   +- *HashAggregate(keys=[DEST_COUNTRY_NAME#189], functions=[partial_count(1)])
      +- *FileScan json [DEST_COUNTRY_NAME#189] Batched: false, Format: JSON, Location: InMemoryFileIndex[dbfs:/FileStore/tables/55rjz3sz1500302954550/flightData2015.json], PartitionFilters: [], PushedFilters: [], ReadSchema: struct<DEST_COUNTRY_NAME:string>
```

%sql select \* from FlightData2015

DEST_COUNTRY_NAME	ORIGIN_COUNTRY_NAME	count
United States	Romania	15

United States	Romania	15
United States	Croatia	1
United States	Ireland	344
Egypt	United States	15
United States	India	62
United States	Singapore	1
United States	Grenada	62
Costa Rica	United States	588
Senegal	United States	40



```
val FlightDataDest = sql("select DEST_COUNTRY_NAME from FlightData2015 ")

FlightDataDest: org.apache.spark.sql.DataFrame = [DEST_COUNTRY_NAME: string]
```

FlightDataDest.show

DEST_COUNTRY_NAME
United States
United States
United States
Egypt
United States
United States
United States
Costa Rica
Senegal
Moldova
United States
United States
Guyana
Malta
Anguilla
Bolivia
United States
Algeria