##SALES DATA ANALYSIS

###Setup for the pySpark

```
!apt-get install openidk-8-jdk-headless -gg > /dev/null
!wget -g http://archive.apache.org/dist/spark/spark-3.1.1/spark-3.1.1-
bin-hadoop3.2.tgz
!tar xf spark-3.1.1-bin-hadoop3.2.tgz
!pip install -q findspark
import os
os.environ["JAVA HOME"] = "/usr/lib/jvm/java-8-openjdk-amd64"
os.environ["SPARK HOME"] = "/content/spark-3.1.1-bin-hadoop3.2"
import findspark
findspark.init()
from pyspark.sql import SparkSession
spark = SparkSession.builder.master("local[*]").getOrCreate()
#for MaxRows
#spark.conf.set('spark.sql.repl.eagerEval.maxNumRows', True)
#To Enable the Spark Rows Show
spark.conf.set('spark.sql.repl.eagerEval.enabled', True)
#spark.conf.set("spark.sql.repl.eagerEval.enabled", True) # Property
used to format output tables better
```

###Importing Libraries

###Importing Sales Dataest

```
sales_data = spark.read.csv("sales_data.csv", header=True,
inferSchema=True)
```

The given dataset contains total 2824 rows and 25 columns.

###Analysing Dataset

```
sales data.printSchema()
#Before Data Cleaning
sales data.show()
root
 |-- ORDERNUMBER: integer (nullable = true)
 -- QUANTITYORDERED: integer (nullable = true)
 -- PRICEEACH: double (nullable = true)
 -- ORDERLINENUMBER: integer (nullable = true)
 -- SALES: double (nullable = true)
 -- ORDERDATE: string (nullable = true)
 -- STATUS: string (nullable = true)
 -- QTR ID: integer (nullable = true)
 -- MONTH ID: integer (nullable = true)
 -- YEAR ID: integer (nullable = true)
 -- PRODUCTLINE: string (nullable = true)
 -- MSRP: integer (nullable = true)
 -- PRODUCTCODE: string (nullable = true)
 -- CUSTOMERNAME: string (nullable = true)
 |-- PHONE: string (nullable = true)
 -- ADDRESSLINE1: string (nullable = true)
 -- ADDRESSLINE2: string (nullable = true)
 -- CITY: string (nullable = true)
 -- STATE: string (nullable = true)
 -- POSTALCODE: string (nullable = true)
 -- COUNTRY: string (nullable = true)
 -- TERRITORY: string (nullable = true)
 -- CONTACTLASTNAME: string (nullable = true)
 |-- CONTACTFIRSTNAME: string (nullable = true)
 |-- DEALSIZE: string (nullable = true)
+-----
+-----
+-----
+----+
|ORDERNUMBER|QUANTITYORDERED|PRICEEACH|ORDERLINENUMBER| SALES|
ORDERDATE | STATUS | QTR ID | MONTH ID | YEAR ID | PRODUCTLINE | MSRP |
PRODUCTCODE |
               CUSTOMERNAME |
                                  PHONE |
ADDRESSLINE1|ADDRESSLINE2|
                                 STATE | POSTAL CODE | COUNTRY |
                           CITY
TERRITORY | CONTACTLASTNAME | CONTACTFIRSTNAME | DEALSIZE |
+-----
+-----
+-----
+----+
| 10107| 30| 95.7| 2| 2871.0| 2/24/2003 0:00|Shipped| 1| 2| 2003|Motorcycles| 95| S10_1678| Land of Toys Inc.| 2125557818|897 Long Airport ...|
```

```
null|
             NYCI
                       NY I
                              10022| USA|
                                                   NA I
                     Small
Yu|
             Kwail
      10121|
                        341
                              81.35|
                                                5 | 2765.9 |
5/7/2003 0:00|Shipped|
                                     2003|Motorcycles| 95|
                         2|
                                 5|
S10_1678| Reims Collectables|
                                26.47.1555| 59 rue de l'Abbaye|
null|
                              51100| France| EMEA|
           Reims|
                     null|
                  Paul| Small|
Henriot|
      10134|
                        411
                              94.741
                                                 213884.341
7/1/2003 0:00|Shipped|
                                     2003|Motorcycles| 95|
                        3|
                               7|
           Lyon Souveniers | +33 1 46 62 7555 | 27 rue du Colonel... |
S10 1678|
null|
           Paris| null|
                              75508| France| EMEA| Da
Cunha|
              Daniel| Medium|
      10145|
                        45|
                              83.261
                                                 6| 3746.7|
8/25/2003 0:00|Shipped|
                                  8|
                                      2003|Motorcycles| 95|
                         31
S10_1678 | Toys4GrownUps.com
                                 6265557265| 78934 Hillside Dr.|
null|
         Pasadena|
                              90003| USA| NA|
                       CA|
Young |
               Julie|
                       Mediuml
                                                14|5205.27|
      10159|
                        49|
                              100.0|
                                 10| 2003|Motorcycles| 95|
10/10/2003 0:00|Shipped|
S10 1678|Corporate Gift Id...|
                                 6505551386| 7734 Strong St.|
null|San Francisco|
                       CA|
                               null| USA| NA|
               Julie|
Brown|
                       Medium|
      10168|
                              96.66
                                                1|3479.76|
                        36|
                                  10| 2003|Motorcycles| 95|
10/28/2003 0:00|Shipped|
S10 1678|Technics Stores Inc.|
                                 6505556809| 9408 Furth Circle|
null| Burlingame|
                              94217| USA| NA|
                       CA|
Hirano|
                 Juril
                        Medium
      10180|
                              86.13|
                                                 9|2497.77|
                        29|
11/11/2003 0:00|Shipped|
                                 11|
                                       2003|Motorcycles| 95|
S10 1678|Daedalus Designs ...|
                                 20.16.1555|184, chausse de T...|
                              59000| France| EMEA|
null| Lille| null|
Rance|
             Martine|
                        Small
      10188|
                        48|
                              100.0
                                                1|5512.32|
11/18/2003 0:00|Shipped|
                               11| 2003|Motorcycles| 95|
                              +47 2267 3215|Drammen 121, PR 7...|
               Herkku Gifts|
S10 1678|
null|
                             N 5804|
                                      Norway| EMEA|
           Bergen| null|
0eztan|
               Veysel|
                       Medium|
      10201
                              98.571
                                                212168.541
                        22|
12/1/2003 0:00|Shipped|
                               12|
                                      2003|Motorcycles| 95|
                       41
S10 1678 | Mini Wheels Co.|
                                 6505555787|5557 North Pendal...|
null|San Francisco|
                               null| USA| NA|
                       CA|
Murphy|
                Julie|
                         Small
                                         14|4708.44|
      10211|
                              100.0|
                        41|
1/15/2004 0:00|Shipped|
                       1|
                               1|
                                      2004|Motorcycles| 95|
S10 1678 | Auto Canal Petit
                             (1) 47.55.6555| 25, rue Lauriston|
null|
                             75016|
                                     France| EMEA|
            Paris| null|
Perrier
             Dominique | Medium |
      10223|
                       37|
                            100.0
                                                1|3965.66|
2/20/2004 0:00|Shipped|
                      1|
                             2| 2004|Motorcycles| 95|
```

```
S10_1678|Australian Collec...| 03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria| 3004|Australia| APAC|
           Peter| Medium|
37| 23| 100
Ferguson|
       10237|
                                 100.0|
                                                     7 | 2333, 12 |
4/5/2004 0:00|Shipped| 2|
4/5/2004 0:00|Shipped| 2| 4| 2004|Motorcycles| 95|
S10_1678| Vitachrome Inc.| 2125551500| 2678 Kingston Rd.|
             NYC| NY| 10022| USA| NA|
Suite 101|
               Michael| Small|
Frick|
       10251|
                                                     2|3188.64|
                          28|
                                 100.0
                                 5| 2004|Motorcycles| 95|
5/18/2004 0:00|Shipped|
                         2|
S10_1678|Tekni Collectable...|
                                    2015559350| 7476 Moss Rd.|
null| Newark| NJ| '
Brown| William| Medium
                                 94019| USA| NA|
                       Medium
      10263|
                                 100.01
                         341
                                                     2|3676.76|
                                 6 | 2004|Motorcycles | 95|
6/28/2004 0:00|Shipped|
                         2|
                                    2035552570 | 25593 South Bay Ln. |
S10 1678 | Gift Depot Inc.
                                 97562| USA| NA|
null| Bridgewater|
                        CT|
King|
                Julie | Medium |
                                92.83| 1|4177.35|
7| 2004|Motorcycles| 95|
      10275|
7/23/2004 0:00|Shipped|
                          3|
                                 40.67.8555|67, rue des Cinqu...|
S10 1678| La Rochelle Gifts|
null | Nantes | null | 44
Labrune | Janine | Medium |
                                 44000| France| EMEA|
      10285|
                          36 | 100.0 |
                                               6|4099.68|
                                8| 2004|Motorcycles| 95|
6175558555| 39323 Spinnaker Dr.|
8/27/2004 0:00|Shipped| 3|
S10 1678|Marta's Replicas Co.|
null| Cambridge| MA|
                                 51247| USA| NA|
                     Marta| Medium|
Hernandez|
                                         9|2597.39|
      10299|
                      23| 100.0|
                                9| 2004|Motorcycles| 95|

90-224 8555| Keskuskatu 45|

21240| Finland| EMEA|
9/30/2004 0:00|Shipped| 3|
S10_1678|Toys of Finland, Co.|
null| Helsinki| null|
                     Matti| Small|
Karttunen|
| 10309| 41| 100.0|
                                         5|4394.38|
| 10/15/2004 0:00|Shipped| 4| 10| 2004|Motorcycles| 95| | S10_1678| Baane Mini Imports| 07-98 9555|Erling Skakkes ga...| | null| Stavern| null| 4110| Norway| EMEAL
       Stavern| null|
                      Jonas| Medium|
Bergulfsen|
                        46 94.74
                                         1|4358.04|
| 10318|
11/2/2004 0:00|Shipped| 4| 11| 2004|Motorcycles| 95|
S10 1678|Diecast Classics ...| 2155551555| 7586 Pompton St.|
null| Allentown|
                         PA| 70267| USA| NA|
Yul
             Kyung| Medium|
                                 100.0|
                         42
| 10329|
                                                    1|4396.14|
11/15/2004 0:00|Shipped| 4| 11| 2004|Motorcycles| 95| S10_1678| Land of Toys Inc.| 21255578181897 Long Airport
S10_1678| Land of Toys Inc.|
                                    2125557818|897 Long Airport ...|
null|
           NYC|
                                 10022| USA| NA|
                         NY |
Yul
             Kwai| Medium|
+-----
```

###Dropping NULL Values

```
cleaned sales data=sales data.na.drop()
#After Data cleaning
cleaned sales data.show()
+-----
+-----
+-----
+-----
+-----
|ORDERNUMBER|QUANTITYORDERED|PRICEEACH|ORDERLINENUMBER|
ORDERDATE | STATUS | QTR ID | MONTH ID | YEAR ID | PRODUCTLINE | MSRP |
              CUSTOMERNAME |
PRODUCTCODE |
                               PHONE |
                                        ADDRESSLINE1|
                    STATE | POSTAL CODE | COUNTRY | TERRITORY |
ADDRESSLINE21
               CITY
CONTACTLASTNAME | CONTACTFIRSTNAME | DEALSIZE |
+-----
+-----
+-----
+-----
+----+
     102231
                  37|
                       100.01
                                      1|3965.66|
2/20/2004 0:00|Shipped|
                          21
                             2004| Motorcycles|
                   11
S10_1678|Australian Collec...|
                       03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                         3004|Australia| APAC|
              Peter | Medium |
Ferguson|
                  23|
                                      7 | 2333.12 |
     10237 I
                       100.01
4/5/2004 0:00|Shipped|
                   2|
                         41
                             2004| Motorcycles| 95|
S10 1678|
          Vitachrome Inc.|
                         2125551500|
                                   2678 Kingston Rd.
Suite 101|
                    NY |
                          10022|
                                   USA| NA|
             NYC
Frick
          Michael|
                  SmallI
     10361
                                     13 | 1451.0 |
                       72.55
                  20|
12/17/2004 0:00|Shipped|
                    41
                          12|
                              2004| Motorcycles| 95|
S10_1678|Souveniers And Th...|+61 2 9495 8555|Monitor Money Bui...|
Level 6|
       Chatswood|
                  NSW|
                         2067|Australia| APAC|
Huxley|
                   Small|
            Adrianl
     10163
                       100.0
                                      1|4860.24|
                  21|
10/20/2003 0:00|Shipped|
                    4|
                          10|
                              2003 | Classic Cars | 214 |
S10 1949 Classic Legends Inc.
                         2125558493| 5905 Pompton St.|
                    NY|
Suite 750
             NYC|
                          10022|
                                   USA| NA|
Hernandez|
               Maria| Medium|
```

```
100.0|
                         21|
                                                    9|4905.39|
      10270
                                         2004|Classic Cars| 214|
7/19/2004 0:00|Shipped|
                           3|
                                   7|
S10 1949|Souveniers And Th...|+61 2 9495 8555|Monitor Money Bui...|
                                   2067|Australia|
Level 6| Chatswood|
                         NSW|
                                                       APACI
Huxley|
                Adrianl
                         Medium|
                                100.0
      10347|
                                                    1 | 3944.7 |
                         30|
                                    11| 2004|Classic Cars| 214|
11/29/2004 0:00|Shipped|
                            4|
S10 1949|Australian Collec...|
                                03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                                   3004|Australia| APAC|
Ferguson
                   Peter | Medium |
                         24|
      10391
                                100.0
                                                    4|2416.56|
3/9/2005 0:00|Shipped|
                          1|
                                   31
                                        2005|Classic Cars| 214|
S10 1949|Anna's Decoration...|
                                02 9936 8555| 201 Miller Street|
                          NSW|
Level 15|North Sydney|
                                    2060|Australia| APAC|
0'Hara|
                  Annal
                          Small|
                                96.34
                                                    3|2793.86|
       10120|
                         29|
                                    4|
4/29/2003 0:00|Shipped|
                           2|
                                        2003| Motorcycles| 118|
S10 2016|Australian Collec...|
                                03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                                   3004|Australia| APAC|
Ferguson|
                   Peter|
                           Small
      10223|
                          47|
                                100.0
                                                    4|5422.39|
2/20/2004 0:00|Shipped|
                           1|
                                    2|
                                         2004| Motorcycles| 118|
S10 2016|Australian Collec...|
                                03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                                   3004|Australia| APAC|
Ferguson|
                   Peter | Medium |
      10361
                                51.15
                                                    8 | 1329.9 |
                         26|
                           4|
12/17/2004 0:00|Shipped|
                                    12| 2004| Motorcycles| 118|
S10 2016|Souveniers And Th...|+61 2 9495 8555|Monitor Money Bui...|
                         NSW|
Level 6| Chatswood|
                                   2067|Australia| APAC|
Huxley|
                Adrianl
                          Small|
                                                    2|9264.86|
       10120|
                                100.0
                          46|
4/29/2003 0:00|Shipped|
                                         2003| Motorcycles| 193|
                           2|
                                    41
S10 4698|Australian Collec...|
                                03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                                   3004|Australia| APAC|
Ferguson|
                   Peter|
                            Large
                                                    3|9774.03|
       10223|
                          49|
                                100.0
2/20/2004 0:00|Shipped|
                                    2|
                                         2004| Motorcycles| 193|
                           1|
S10 4698|Australian Collec...|
                                03 9520 4555| 636 St Kilda Road|
Level 3| Melbourne|Victoria|
                                   3004|Australia| APAC|
                   Peter|
Ferguson|
                           Large
                         39|
       10237
                                100.0
                                                    9 | 7023.9 |
                          2|
4/5/2004 0:00|Shipped|
                                        2004 | Motorcycles | 193 |
                                  4|
                                  2125551500| 2678 Kingston Rd.|
S10 4698|
            Vitachrome Inc.|
Suite 101|
                  NYC
                            NY |
                                    10022| USA| NA|
              Michael|
                         Large|
Frick|
       10337|
                          25|
                                48.05
                                                    8 | 1201.25 |
                                    11| 2004|Classic Cars| 136|
11/21/2004 0:00|Shipped|
                            4|
                                  2125558493|
S10 4757|Classic Legends Inc.|
                                                5905 Pompton St.
Suite 750|
                  NYCI
                            NY|
                                    10022|
                                                USA|
                                                           NA I
```

```
Hernandez I
                  Maria|
                         Small
                            100.01
                                             2|4302.08|
      10270|
                      32|
7/19/2004 0:00|Shipped|
                        31
                               7|
                                    2004|Classic Cars| 147|
S10 4962|Souveniers And Th...|+61 2 9495 8555|Monitor Money Bui...|
Level 6|
         Chatswood|
                      NSW
                               2067|Australia|
                                                APACI
Huxley|
              Adrianl
                      Medium |
                            100.0|
                                             2 | 4428.0 |
      10347
                      27|
11/29/2004 0:00|Shipped|
                                     2004 | Classic Cars | 147 |
                         4|
                                111
S10 4962|Australian Collec...|
                            03 9520 4555| 636 St Kilda Road|
Level 31
         Melbourne|Victoria|
                               3004|Australia|
                                                APACI
                 Peter | Medium
Ferguson|
      10370
                      35|
                            65.631
                                             4|2297.05|
1/20/2005 0:00|Shipped|
                                    2005|Classic Cars| 147|
                        11
                               11
S10 4962|Anna's Decoration...|
                            02 9936 85551
                                         201 Miller Street
Level 15|North Sydney|
                       NSWI
                               2060|Australia| APAC|
0'Hara|
                       Small|
                Anna|
      10391
                      37 I
                             46.91
                                             7 | 1735.3 |
3/9/2005 0:00|Shipped|
                       1|
                               3|
                                   2005|Classic Cars| 147|
                            02 9936 8555| 201 Miller Street|
S10 4962|Anna's Decoration...|
Level 15|North Sydney|
                       NSW|
                               2060|Australia| APAC|
0'Hara|
                Anna|
                       Small|
      10169|
                      30|
                            100.0|
                                             2 | 5019.9 |
11/4/2003 0:00|Shipped|
                                    2003 | Classic Cars | 194 |
                        4|
                               111
S12 1099|Anna's Decoration...|
                            02 9936 8555| 201 Miller Street|
Level 15|North Sydney|
                       NSW|
                               2060|Australia| APAC|
0'Hara|
                Annal
                      Medium
                            100.0
                                             2|11279.2|
      10127
                      46 l
                       2|
                                   2003|Classic Cars| 207|
6/3/2003 0:00|Shipped|
                               6|
                              2125557413|
S12 1108 | Muscle Machine Inc
                                          4092 Furth Circle
                NYCI
                         NYI
Suite 400|
                               10022|
                                          USAI
               Jeffl
Young |
                      Largel
+-----
+-----
+-----
+-----
+----+
only showing top 20 rows
```

###Dropping Duplicate Values

```
|ORDERNUMBER|QUANTITYORDERED|PRICEEACH|ORDERLINENUMBER|
                                                 SALESI
         STATUS|QTR ID|MONTH ID|YEAR ID| PRODUCTLINE|MSRP|
ORDERDATE |
PRODUCTCODE |
                 CUSTOMERNAME |
                                     PHONE |
                                                 ADDRESSLINE1|
                  CITY| STATE|POSTALCODE| COUNTRY|TERRITORY|
ADDRESSLINE21
CONTACTLASTNAME | CONTACTFIRSTNAME | DEALSIZE |
+-----
+---+---
+-----
+----+
     10163|
                      31 | 100.0|
                                             2|3329.09|
      2003 0:00| Shipped| 4| 10| 2003| Vintage Cars|
| S18_2949|Classic Legends Inc.| 2125558493| 5905 Pompton
10/20/2003 0:00|
                                              5905 Pompton
      Suite 750|
                      NYC| NY| 10022| USA| NA|
Hernandez|
                  Maria| Medium|
                                             6|3871.92|
      10361
                      34|
                          100.0
                          4 12 2004
12/17/2004 0:00| Shipped|
                                              Vintage Carsl
65| S24 3420|Souveniers And Th...|+61 2 9495 8555|Monitor Money
Bui...| Level 6| Chatswood| NSW|
                                         2067|Australia|
          Huxlevl
APACI
                          Adrian| Medium|
                      29|
                           100.0|
      10337|
                                             2|4498.19|
     /2004 0:00| Shipped| 4| 11| 2004| Vintage Cars|
S18_4522|Classic Legends Inc.| 2125558493| 5905 Pompton
11/21/2004 0:00|
                                              5905 Pompton
                     NYC| NY| 10022| USA| NA|
      Suite 750|
Hernandez|
                  Maria| Medium|
                          100.0
                      31|
                                             7|3184.94|
      10139|
                          3| 7| 2003| Vintage Cars|
7/16/2003 0:00| Shipped|
102| S18_1342|Souveniers And Th...|+61 2 9495 8555|Monitor Money
Bui...| Level 6| Chatswood| NSW|
                                         2067|Australia|
APAC|
            Huxley|
                         Adrian| Medium|
                            62.46
      10361
                                             1 | 1561.5 |
                      25|
12/17/2004 0:00| Shipped|
                                12 | 2004 |
                          4|
                                                    Planesl
72| S24 4278|Souveniers And Th...|+61 2 9495 8555|Monitor Money
Bui...| Level 6| Chatswood| NSW|
                                         2067|Australia|
APAC |
           Huxley|
                          Adrian| Small|
                            96.74
                                             4|4256.56|
      10267
                                 7 | 2004 |
7/7/2004 0:00| Shipped|
                         3|
                                             Classic Cars
                                 2125557413| 4092 Furth
85| S24 3191| Muscle Machine Inc|
Circle| Suite 400| NYC|
                                 NYI
                                        10022| USA|
                          Jeffl
                                Medium|
           Young |
                           100.0
      10169|
                                             2| 5019.9|
                      30|
                              11| 2003| Classic Cars
11/4/2003 0:00| Shipped|
                          4|
194| S12_1099|Anna's Decoration...| 02 9936 8555| 201 Miller
Street| Level 15|North Sydney|
                                NSW|
                                         2060|Australia|
            0'Hara|
                           Annal Mediuml
APACI
                           54.33|
                                             1|1466.91|
      10324
                      27|
11/5/2004 0:00| Shipped| 4| 11| 2004| Classic Cars|
151| S12_3148| Vitachrome Inc.| 2125551500| 2678 Kingston
Rd.| Suite 101| NYC| NY| 10022| USA| NA|
```

```
Frick|
                Michaell
                            Small
                            29| 85.49| 9|2479.21|
     10120|
4/29/2003 0:00| Shipped| 2| 4| 2003| Planes| 72| S24_4278|Australian Collec...| 03 9520 4555| 636 St Kilda
      __Level 3| Melbourne|Victoria| 3004|Australia|
Ferguson| Peter| Small|
Road l
APAC |
                        37|
                                  100.0
                                                         5| 5283.6|
       10420|
5/29/2005 0:00|In Process|
                                 2| 5| 2005| Vintage Cars|
170| S18 1749|Souveniers And Th...|+61 2 9495 8555|Monitor Money
Bui...| Level 6| Chatswood| NSW|
                                                    2067|Australia|
APAC|
             Huxley|
                                Adrian| Medium|
                                  87.24
       10267
                            38|
                                                         3|3315.12|
                              3| 7| 2004| Classic Cars|
ine Inc| 2125557413| 4092 Furth
NYC| NY| 10022| USA|
7/7/2004 0:00|
                 Shipped|
90| S24 2766| Muscle Machine Incl
Circle| Suite 400|
NA| Young|
| 10127| 20|
                                Jeff|
                                         Medium|
                            20| 60.69|
                                                         8 | 1213.8 |
                                2| 6| 2003| Classic Cars|
6/3/2003 0:00| Shipped|
                                        2125557413| 4092 Furth
57| S24 1444| Muscle Machine Incl
                                NYC| NY| 10022| USA|
Jeff| Small|
Circle| Suite 400|
                               NYC|
             Young
| 10337| 42| 97.16| 5|4080.72| 11/21/2004 0:00| Shipped| 4| 11| 2004| Ships| 54| S72_3212|Classic Legends Inc.| 2125558493| 5905 Pompton
St.| Suite 750| NYC| NY| 10022| USA| NA|
                      Maria| Medium|
Hernandezl
                           42| 100.0|
       10391
                                                         3| 4998.0|
3/9/2005 0:00| Shipped| 1| 3| 2005| Vintage Cars|
53| S18_1367|Anna's Decoration...| 02 9936 8555| 201 Miller
Street| Level 15|North Sydney| NSW| 2060|Australia|
               O'Hara| Anna| Medium|
APAC|
                                  91.02
       102041
                            481
                                                        11|4368.96|
                              4| 12| 2003| Classic Cars|
12/2/2003 0:00| Shipped|
107| S24_3432| Muscle Machine Inc|
                                         2125557413| 4092 Furth
Circle | Suite 400 | NYC | NAI Young | Jeff |
                                         NY| 10022| USA|
             Young|
|
                                Jeff| Medium|
                           40 | 100.0 |
       10163|
                                                         3 | 4900.8 |
| 10163| 40| 100.0| 3| 4900.8|

10/20/2003 0:00| Shipped| 4| 10| 2003| Vintage Cars|

104| S18_3136|Classic Legends Inc.| 2125558493| 5905 Pompton

St.| Suite 750| NYC| NY| 10022| USA| NA|
       Marial Mediuml
Hernandez|
                            35| 100.0|
                                                      11|4277.35|
12/17/2004 0:00| Shipped| 4| 12| 2004| Plane 74| S700_4002|Souveniers And Th...|+61 2 9495 8555|Monitor Money
                                                                  Planes|
Bui...| Level 6| Chatswood|
                                        NSW| 2067|Australia|
APAC|
            Huxley|
                                 Adrian| Medium|
                            48| 69.96|
      10163|
                                                         4|3358.08|
10/20/2003 0:00| Shipped| 4| 10| 2003| Vintage Cars| 62| S18_2957|Classic Legends Inc.| 2125558493| 5905 Pompton
```

```
10022|
                                   USAI
                                           NAI
St.l
    Suite 750|
                NYCI
                       NY I
Hernandez|
             Maria| Medium|
    10120|
                29|
                    96.34
                                  3|2793.86|
                         4|
                   2|
4/29/2003 0:00|
           Shipped
                            2003|
                                   Motorcycles|
    S10 2016|Australian Collec...|
                         03 9520 4555|
                                   636 St Kilda
      Level 3|
             Melbourne|Victoria|
                             3004|Australia|
Roadl
                          Small|
APAC |
       Ferguson|
                    Peter|
                                 11| 2559.6|
    101271
                24|
                     100.0
                   2|
                         6|
                            2003|Trucks and Buses|
6/3/2003 0:00|
          Shipped|
    S12 4473| Muscle Machine Inc|
                           2125557413|
                                   4092 Furth
Circle
      Suite 400|
                         NY|
                              10022|
                  NYC|
NA I
        Young
                   Jeffl
                         Small
+---+----
+----+
only showing top 20 rows
```

###Calculating Totals Sales By Each Product

```
calculate total sales data =
dropped duplicates data.groupBy("PRODUCTLINE").sum("SALES").withColumn
Renamed("sum(SALES)", "TotalSales")
calculate_total_sales_data.show()
      PRODUCTLINE| TotalSales|
     Motorcycles|
                         116738.31
     Vintage Cars|
                         199436.24
            Ships|
                          15146.83
Trucks and Buses
                          120698.5
     Classic Cars|379958.2299999999
           Trains
                            7530.651
           Planes | 63772.0900000001|
```

###Creating New CSV File

```
calculate_total_sales_data.coalesce(1).write.csv("product_sales_summar
y.csv", header=True)
spark.stop()
```