

1) ***Launch Spark shell*******

```
spark-shell --packages com.databricks:spark-csv_2.10:1.5.0
```

2) ***Create and Import sqlContext.*******

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

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

3) ***Creating data frame & Importing csv file.*******

```
val data = sqlContext.read.format("com.databricks.spark.csv").option("header",
"true").option("delimiter",",").load("/home/acadgild/Desktop/Hadoop_SME_Coding_Data_Set.do
cx.csv")
```

4) ***Create table from the data*******

```
data.registerTempTable("olympics")
```

Question 1) Find the total number of bronze medals won by each country in Football.

```
val result1= sqlContext.sql("select Country,count(Medal) as Medal from olympics where
Sport=='Football' and Medal=='Bronze' group by Country").show()
```

```
scala> val result1= sqlContext.sql("select Country,count(Medal) as Medal,medal from olympics where Sport=='Football' and Medal=='Bronze' group by Cou
ntry,medal").show(50)
+-----+-----+
|Country|Medal| medal|
+-----+-----+
| ITA | 39 | Bronze |
| FRG | 20 | Bronze |
| CAN | 20 | Bronze |
| GER | 53 | Bronze |
| NED | 36 | Bronze |
| BUL | 14 | Bronze |
| SWE | 34 | Bronze |
| JPN | 16 | Bronze |
| URS | 53 | Bronze |
| CHI | 18 | Bronze |
| YUG | 17 | Bronze |
| BRA | 36 | Bronze |
| GDR | 17 | Bronze |
| DEN | 14 | Bronze |
| EUA | 19 | Bronze |
| GHA | 13 | Bronze |
| HUN | 17 | Bronze |
| BEL | 11 | Bronze |
| NOR | 32 | Bronze |
| USA | 12 | Bronze |
| KOR | 18 | Bronze |
+-----+-----+
```

Question 2) Find the number of Medals won by USA grouped by sport

```
val result2= sqlContext.sql("select Count(Medal) as Medal,country,sport from olympics where Country=='USA' group by Sport,country").show()
```

```
scala> val result2= sqlContext.sql("select Count(Medal) as Medal,country,sport from olympics where Country=='USA' group by Sport,country").show(50)
```

Medal	country	sport
21	USA	Canoe / Kayak
52	USA	Fencing
9	USA	Polo
132	USA	Football
66	USA	Cycling
1184	USA	Aquatics
129	USA	Wrestling
121	USA	Equestrian
1	USA	Skating
17	USA	Modern Pentathlon
178	USA	Gymnastics
190	USA	Shooting
60	USA	Softball
111	USA	Boxing
54	USA	Tennis
364	USA	Rowing
12	USA	Judo
139	USA	Sailing
11	USA	Ice Hockey
28	USA	Hockey
46	USA	Rugby
13	USA	Lacrosse
3	USA	Roque
43	USA	Weightlifting
318	USA	Basketball
994	USA	Athletics
1	USA	Jeu de paume
15	USA	Tug of War
55	USA	Archery
68	USA	Baseball
114	USA	Volleyball
27	USA	Golf

Question 3) Find the total number of medals won by each country displayed by type of medal.

```
val result3= sqlContext.sql("select Country,Medal,Count(Medal) as Count from olympics group by Medal,country").show()
```

```
scala> val result3= sqlContext.sql("select Country,Medal,Count(Medal) as Count from olympics group by Medal,country").show(50)
```

Country	Medal	Count
FIN	Bronze	212
RSA	Silver	39
CRO	Bronze	34
ZIM	Gold	18
ISR	Gold	1
HUN	Bronze	351
ZZX	Bronze	10
THA	Gold	7
LAT	Silver	11
SYR	Gold	1
USA	Gold	2235
JAM	Gold	31
ARG	Gold	69
CHI	Gold	3
HUN	Silver	316
POL	Bronze	231
YUG	Bronze	118
SUR	Gold	1
LTU	Gold	6
TRI	Gold	1
NED	Gold	233
UKR	Bronze	90
CUB	Silver	129
RSA	Bronze	37
HAI	Silver	1
UZB	Gold	5
SYR	Bronze	1
JPN	Bronze	303
IOP	Silver	1
HKG	Bronze	1
TCH	Gold	80
NOR	Gold	209

Question 4) Find how many Silver medals have been given to Mexico and the year of each

```
val result4= sqlContext.sql("select Country,Medal,Count(Medal) as Count,year from olympics where Medal='Silver' and Country='MEX' group by Medal,country,year").show(false)
```

```
scala> val result4= sqlContext.sql("select Country,Medal,Count(Medal) as Count,year from olympics where Medal='Silver' and Country='MEX' group by Medal,country,year").show(50)
```

Country	Medal	Count	year
MEX	Silver	1	1952
MEX	Silver	5	2012
MEX	Silver	1	1992
MEX	Silver	1	1980
MEX	Silver	1	1972
MEX	Silver	3	1984
MEX	Silver	2	1932
MEX	Silver	3	1968
MEX	Silver	2	2000
MEX	Silver	3	2004
MEX	Silver	1	1948

```
result4: Unit = ()
scala> █
```