

Spark Query Results

1. Year wise carrier delay from 2003-2010

```
scala> spark.sql("SELECT Year, SUM(CarrierDelay) as total_delay FROM DelayedFlights WHERE Year >= 2003 AND Year <= 2010 AND CarrierDelay IS NOT NULL GROUP BY Year ORDER BY Year").show()
-----+-----+
|Year|total_delay|
-----+-----+
|2003|    1577.0|
|2004|    2064.0|
|2005|    1436.0|
|2006|    1540.0|
|2007|    1835.0|
|2008|    1298.0|
|2009|    1191.0|
|2010|    1992.0|
-----+-----+
```

2. Year wise NAS delay from 2003-2010

```
scala> spark.sql("SELECT Year, SUM(NASDelay) as total_delay FROM DelayedFlights WHERE Year >= 2003 AND Year <= 2010 AND CarrierDelay IS NOT NULL GROUP BY Year ORDER BY Year").show()
-----+-----+
|Year|total_delay|
-----+-----+
|2003|    1856.0|
|2004|     493.0|
|2005|     680.0|
|2006|     869.0|
|2007|    4269.0|
|2008|    1410.0|
|2009|    1963.0|
|2010|    3568.0|
-----+-----+
```

3. Year wise Weather delay from 2003-2010

```
scala> spark.sql("SELECT Year, SUM(LateAircraftDelay) as total_delay FROM DelayedFlights WHERE Year >= 2003 AND Year <= 2010 AND CarrierDelay IS NOT NULL GROUP BY Year ORDER BY Year").show()
-----+-----+
|Year|total_delay|
-----+-----+
|2003|    1986.0|
|2004|    1451.0|
|2005|    2332.0|
|2006|    2295.0|
|2007|    5726.0|
|2008|    2069.0|
|2009|    1819.0|
|2010|    4233.0|
-----+-----+
```

4. Year wise late aircraft delay from 2003-2010

```
scala> spark.sql("SELECT Year, SUM(LateAircraftDelay) as total_delay FROM DelayedFlights WHERE Year >= 2003 AND Year <= 2010 AND CarrierDelay IS NOT NULL GROUP BY Year ORDER BY Year").show()
-----+-----+
|Year|total_delay|
-----+-----+
|2003|    1986.0|
|2004|    1451.0|
|2005|    2332.0|
|2006|    2295.0|
|2007|    5726.0|
|2008|    2069.0|
|2009|    1819.0|
|2010|    4233.0|
-----+-----+
```

5. Year wise security delay from 2003-2010

```
scala> spark.sql("SELECT Year, SUM(SecurityDelay) as total_delay FROM DelayedFlights WHERE Year >= 2003 AND Year <= 2010 AND CarrierDelay IS NOT NULL GROUP BY Year ORDER BY Year").show()
-----+-----+
|Year|total_delay|
-----+-----+
|2003|         0.0|
|2004|         0.0|
|2005|         0.0|
|2006|         0.0|
|2007|         12.0|
|2008|         0.0|
|2009|         0.0|
|2010|         0.0|
-----+-----+
```