BDA2 – Spark SQL

alebo256 & danibi675

1.

```
+----+----+          +----+-----+
|year| Max|          |year|  Min|
+----+----+          +----+-----+
|1975|36.1|          |1990|-35.0|
|1992|35.4|          |1952|-35.5|
|1994|34.7|          |1974|-35.6|
|2010|34.4|          |1954|-36.0|
|2014|34.4|          |1992|-36.1|
|1989|33.9|          |1975|-37.0|
|1982|33.8|          |1972|-37.5|
|1968|33.7|          |1995|-37.6|
|1966|33.5|          |2000|-37.6|
|2002|33.3|          |1957|-37.8|
|1983|33.3|          |1983|-38.2|
|1986|33.2|          |1989|-38.2|
|1970|33.2|          |1953|-38.4|
|2000|33.0|          |2009|-38.5|
|1956|33.0|          |1993|-39.0|
|1959|32.8|          |1984|-39.2|
|2006|32.7|          |2008|-39.3|
|1991|32.7|          |1973|-39.3|
|1988|32.6|          |1991|-39.3|
|2011|32.5|          |2005|-39.4|
+----+----+          +----+-----+
only showing top 20 rows   only showing top 20 rows
```

2. a) The cunt for each month of all the times temperature above 10 degrees were recorded.

```
+----+-----+------+
|year|month| count|
+----+-----+------+
|2014|   07|147681|
|2011|   07|146656|
|2010|   07|143419|
|2012|   07|137477|
|2013|   07|133657|
|2009|   07|133008|
|2011|   08|132734|
|2009|   08|128349|
|2013|   08|128235|
|2003|   07|128133|
|2002|   07|127956|
|2006|   08|127622|
|2008|   07|126973|
|2002|   08|126073|
|2005|   07|125294|
|2011|   06|125193|
|2012|   08|125037|
|2006|   07|124794|
|2010|   08|124417|
|2014|   08|124045|
+----+-----+------+
only showing top 20 rows
```

2. b) The count for distinct stations.

```
+----+-----+-----+
|year|month|count|
+----+-----+-----+
|1972|   10|  378|
|1973|   05|  377|
|1973|   06|  377|
|1973|   09|  376|
|1972|   08|  376|
|1971|   08|  375|
|1972|   05|  375|
|1972|   06|  375|
|1972|   09|  375|
|1971|   09|  374|
|1972|   07|  374|
|1971|   06|  374|
|1971|   05|  373|
|1973|   08|  373|
|1974|   08|  372|
|1974|   06|  372|
|1974|   05|  370|
|1971|   07|  370|
|1970|   08|  370|
|1974|   09|  370|
+----+-----+-----+
only showing top 20 rows
```

3.

```
+-------+----+-----+------------------+
|station|year|month|        avg(value)|
+-------+----+-----+------------------+
|  96000|2014|   07|              26.3|
|  65450|1994|   07| 23.65483870967742|
|  95160|1994|   07|23.505376344086027|
|  75120|1994|   07| 23.26881720430107|
| 105260|1994|   07|23.143820224719107|
|  85280|1994|   07|23.108602150537635|
|  54550|1983|   08|              23.0|
|  54550|1975|   08|           22.9625|
|  96550|1994|   07|22.957894736842114|
|  96000|1994|   07|22.931182795698923|
| 106070|1994|   07|22.822580645161295|
| 173960|1972|   07|22.776666666666667|
|  54300|1994|   07| 22.76021505376344|
|  85210|1994|   07|22.755913978494615|
|  65450|2006|   07| 22.74086021505376|
|  75120|2006|   07| 22.73010752688173|
| 103080|1994|   07|22.708602150537626|
|  92100|1994|   07|22.698924731182792|
|  94180|1994|   07| 22.68172043010753|
|  83230|1994|   07|22.577419354838707|
+-------+----+-----+------------------+
only showing top 20 rows
```

4. The result was empty.

```
+-------+--------+--------+
|station|Max_Perc|Max_Temp|
+-------+--------+--------+
+-------+--------+--------+
```

5.

```
+----+-----+------------------+
|year|month|   avg(Sum_Month) |
+----+-----+------------------+
|2016|   07|               0.0|
|2016|   06|           47.6625|
|2016|   05|29.250000000000007|
|2016|   04| 26.90000000000001|
|2016|   03|19.962500000000002|
|2016|   02|           21.5625|
|2016|   01|22.325000000000003|
|2015|   12|            28.925|
|2015|   11| 63.88750000000002|
|2015|   10|            2.2625|
|2015|   09|101.29999999999997|
|2015|   08|26.987500000000004|
|2015|   07|119.09999999999998|
|2015|   06| 78.66250000000001|
|2015|   05| 93.22499999999998|
|2015|   04|           15.3375|
|2015|   03|42.612500000000004|
|2015|   02|            24.825|
|2015|   01| 59.11250000000003|
|2014|   12|35.462500000000006|
+----+-----+------------------+
only showing top 20 rows
```