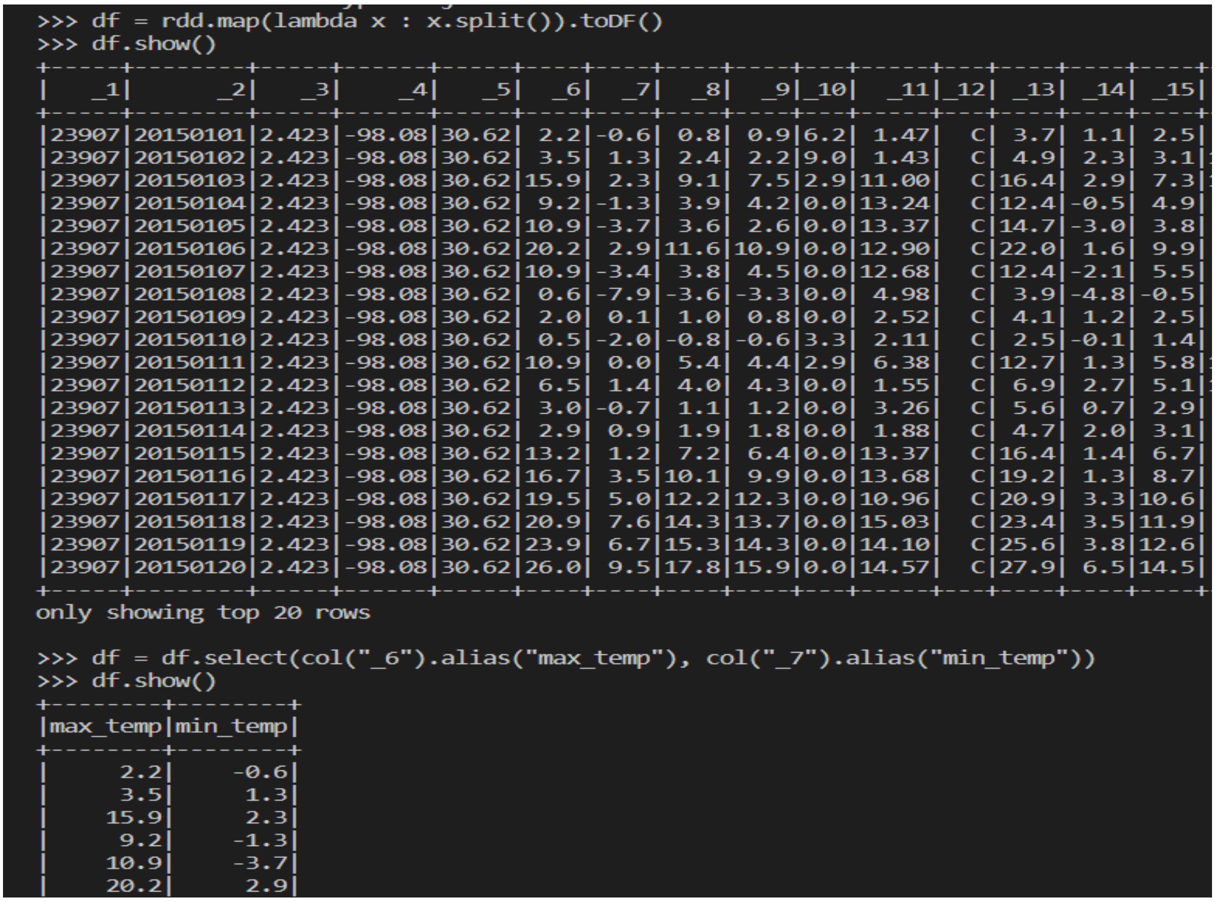
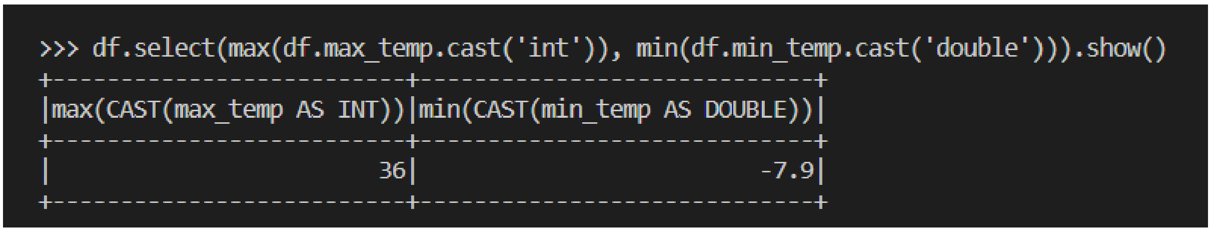
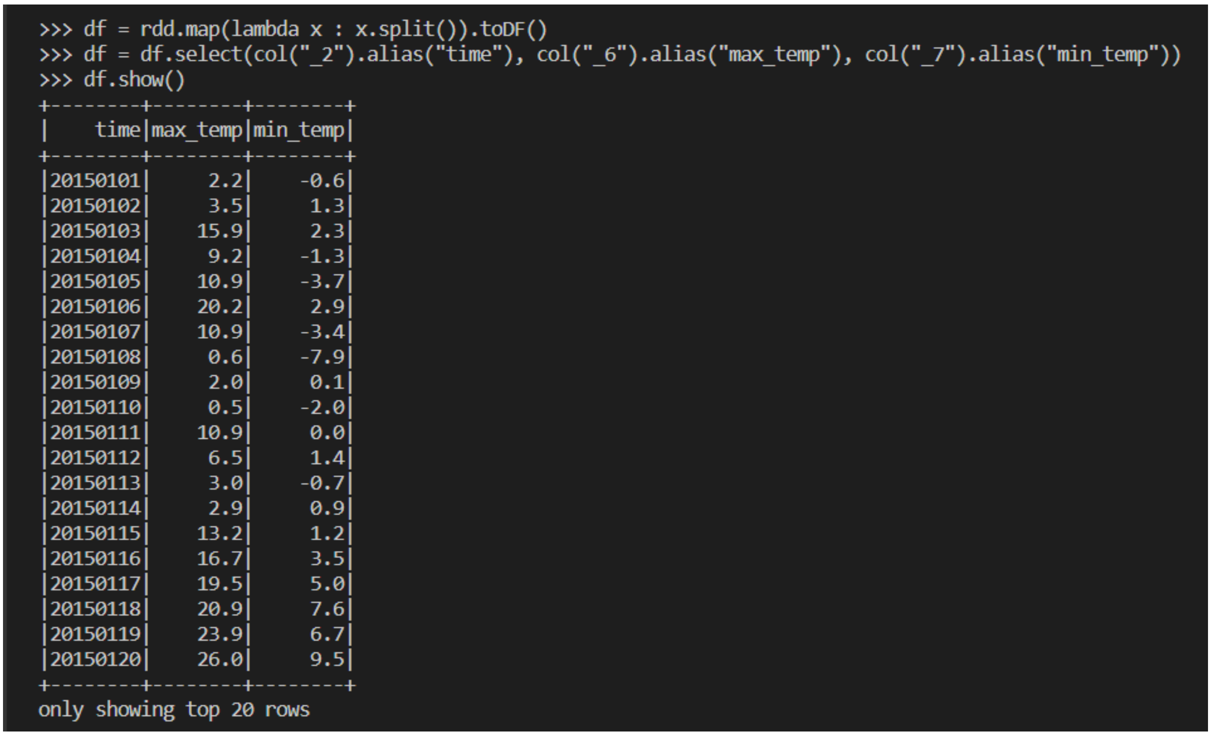


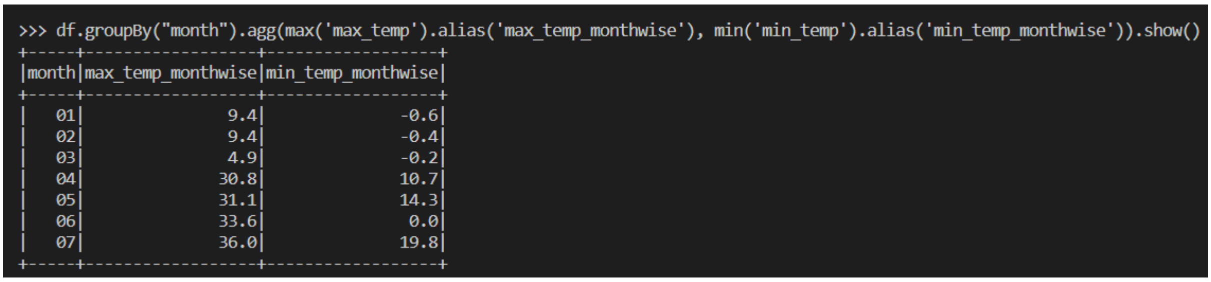
```
>>> df = rdd.map(lambda x : x.split()).toDF()
>>> df.show()
+-----+--------+-----+------+-----+----+----+----+----+---+-----+---+----+----+----+
|   _1|      _2|   _3|    _4|   _5|  _6|  _7|  _8|  _9|_10|  _11|_12| _13| _14| _15|
+-----+--------+-----+------+-----+----+----+----+----+---+-----+---+----+----+----+
|23907|20150101|2.423|-98.08|30.62| 2.2|-0.6| 0.8| 0.9|6.2| 1.47|  C| 3.7| 1.1| 2.5|
|23907|20150102|2.423|-98.08|30.62| 3.5| 1.3| 2.4| 2.2|9.0| 1.43|  C| 4.9| 2.3| 3.1|
|23907|20150103|2.423|-98.08|30.62|15.9| 2.3| 9.1| 7.5|2.9|11.00|  C|16.4| 2.9| 7.3|
|23907|20150104|2.423|-98.08|30.62| 9.2|-1.3| 3.9| 4.2|0.0|13.24|  C|12.4|-0.5| 4.9|
|23907|20150105|2.423|-98.08|30.62|10.9|-3.7| 3.6| 2.6|0.0|13.37|  C|14.7|-3.0| 3.8|
|23907|20150106|2.423|-98.08|30.62|20.2| 2.9|11.6|10.9|0.0|12.90|  C|22.0| 1.6| 9.9|
|23907|20150107|2.423|-98.08|30.62|10.9|-3.4| 3.8| 4.5|0.0|12.68|  C|12.4|-2.1| 5.5|
|23907|20150108|2.423|-98.08|30.62| 0.6|-7.9|-3.6|-3.3|0.0| 4.98|  C| 3.9|-4.8|-0.5|
|23907|20150109|2.423|-98.08|30.62| 2.0| 0.1| 1.0| 0.8|0.0| 2.52|  C| 4.1| 1.2| 2.5|
|23907|20150110|2.423|-98.08|30.62| 0.5|-2.0|-0.8|-0.6|3.3| 2.11|  C| 2.5|-0.1| 1.4|
|23907|20150111|2.423|-98.08|30.62|10.9| 0.0| 5.4| 4.4|2.9| 6.38|  C|12.7| 1.3| 5.8|
|23907|20150112|2.423|-98.08|30.62| 6.5| 1.4| 4.0| 4.3|0.0| 1.55|  C| 6.9| 2.7| 5.1|
|23907|20150113|2.423|-98.08|30.62| 3.0|-0.7| 1.1| 1.2|0.0| 3.26|  C| 5.6| 0.7| 2.9|
|23907|20150114|2.423|-98.08|30.62| 2.9| 0.9| 1.9| 1.8|0.0| 1.88|  C| 4.7| 2.0| 3.1|
|23907|20150115|2.423|-98.08|30.62|13.2| 1.2| 7.2| 6.4|0.0|13.37|  C|16.4| 1.4| 6.7|
|23907|20150116|2.423|-98.08|30.62|16.7| 3.5|10.1| 9.9|0.0|13.68|  C|19.2| 1.3| 8.7|
|23907|20150117|2.423|-98.08|30.62|19.5| 5.0|12.2|12.3|0.0|10.96|  C|20.9| 3.3|10.6|
|23907|20150118|2.423|-98.08|30.62|20.9| 7.6|14.3|13.7|0.0|15.03|  C|23.4| 3.5|11.9|
|23907|20150119|2.423|-98.08|30.62|23.9| 6.7|15.3|14.3|0.0|14.10|  C|25.6| 3.8|12.6|
|23907|20150120|2.423|-98.08|30.62|26.0| 9.5|17.8|15.9|0.0|14.57|  C|27.9| 6.5|14.5|
+-----+--------+-----+------+-----+----+----+----+----+---+-----+---+----+----+----+
only showing top 20 rows

>>> df = df.select(col("_6").alias("max_temp"), col("_7").alias("min_temp"))
>>> df.show()
+--------+--------+
|max_temp|min_temp|
+--------+--------+
|     2.2|    -0.6|
|     3.5|     1.3|
|    15.9|     2.3|
|     9.2|    -1.3|
|    10.9|    -3.7|
|    20.2|     2.9|
```



```
>>> df.select(max(df.max_temp.cast('int')), min(df.min_temp.cast('double'))).show()
+-------------------------+----------------------------+
|max(CAST(max_temp AS INT))|min(CAST(min_temp AS DOUBLE))|
+-------------------------+----------------------------+
|                       36|                        -7.9|
+-------------------------+----------------------------+
```



```
>>> df = rdd.map(lambda x : x.split()).toDF()
>>> df = df.select(col("_2").alias("time"), col("_6").alias("max_temp"), col("_7").alias("min_temp"))
>>> df.show()
+--------+--------+--------+
|    time|max_temp|min_temp|
+--------+--------+--------+
|20150101|     2.2|    -0.6|
|20150102|     3.5|     1.3|
|20150103|    15.9|     2.3|
|20150104|     9.2|    -1.3|
|20150105|    10.9|    -3.7|
|20150106|    20.2|     2.9|
|20150107|    10.9|    -3.4|
|20150108|     0.6|    -7.9|
|20150109|     2.0|     0.1|
|20150110|     0.5|    -2.0|
|20150111|    10.9|     0.0|
|20150112|     6.5|     1.4|
|20150113|     3.0|    -0.7|
|20150114|     2.9|     0.9|
|20150115|    13.2|     1.2|
|20150116|    16.7|     3.5|
|20150117|    19.5|     5.0|
|20150118|    20.9|     7.6|
|20150119|    23.9|     6.7|
|20150120|    26.0|     9.5|
+--------+--------+--------+
only showing top 20 rows
```



```
>>> df.groupBy("month").agg(max('max_temp').alias('max_temp_monthwise'), min('min_temp').alias('min_temp_monthwise')).show()
+-----+------------------+------------------+
|month|max_temp_monthwise|min_temp_monthwise|
+-----+------------------+------------------+
|   01|               9.4|              -0.6|
|   02|               9.4|              -0.4|
|   03|               4.9|              -0.2|
|   04|              30.8|              10.7|
|   05|              31.1|              14.3|
|   06|              33.6|               0.0|
|   07|              36.0|              19.8|
+-----+------------------+------------------+
```