Module 4: Hadoop MapReduce Framework - II

Assignment

edureka!

edureka!

© 2014 Brain4ce Education Solutions Pvt. Ltd.

Module 4: Hadoop MapReduce Framework - II

Assignment - Find the Hot and Cold Days

Table of Contents

Introduction	2
Problem statement	2



Introduction

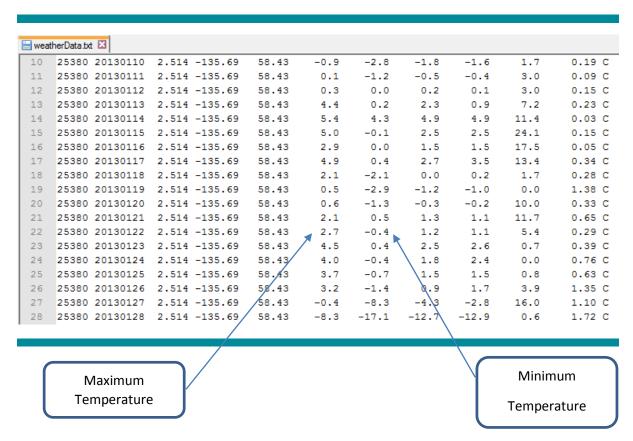
Apply your MapReduce programming knowledge and write a MapReduce program to process a text file. You need to print the count of number of occurrences of each word in the text file.

The dataset for this problem is the **'WeatherData'** records file available in your LMS. This dataset has been taken from <u>National Climatic Data Center (NCDC)</u> public datasets. You can download more datasets from this FTP site and review the <u>README</u> file to understand the available datasets.

Problem statement

Let's understand the problem through a subset of patent records as shown in the following figure:

FIGURE 1-1 WEATHER RECORDS



Your task is to find out the dates with maximum temperature greater than 40 (A Hot Day) and minimum temperature lower than 10 (A Cold Day).

Here is the sample output:

FIGURE 1-2 SAMPLE OUTPUT

04-02-2013	Cold Day
04-03-2013	Cold Day
04-04-2013	Cold Day
04-05-2013	Cold Day
04-06-2013	Hot Day
04-07-2013	Hot Day
04-08-2013	Hot Day
04-09-2013	Hot Day
05-01-2013	Cold Day
05-03-2013	Cold Day
05-04-2013	Cold Day

You can review the solution in your LMS.

