

Module 4: Hadoop MapReduce Framework - II

Assignment

edureka!

edureka!

© 2014 Brain4ce Education Solutions Pvt. Ltd.

Module 4: Hadoop MapReduce Framework - II

Assignment – Calculate Maximum Temperature

Table of Contents

Introduction	2
Problem statement	2

edureka!

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 text file '**Temperature**' available in your LMS

Problem statement

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

FIGURE 1-1 PATENT RECORDS

1900	36
1900	29
1901	32
1901	40
1901	29
1901	48
1901	16
1901	11
1901	21
1901	6
1901	22
1902	49
1902	49

In this data set, the first field represents the year and the second field represents the temperature in that year. As the temperature will not be constant throughout the year, each year has multiple temperatures listed in the dataset. You need to process the dataset and find the maximum temperature during a year. Here is the sample output:

Sample Output

Year	Maximum Temperature
1900	36
1901	48
1902	49

Your task in this assignment is to process the **'temperature'** records using MapReduce program and find out the maximum temperature during each year in this dataset. You can review the solution in your LMS.

edureka!