**Name: Viral Gajera**

**Big Data and Analytics**

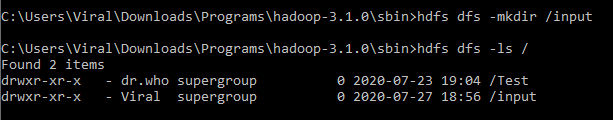
**Roll No.: CE-032**

**LAB4**

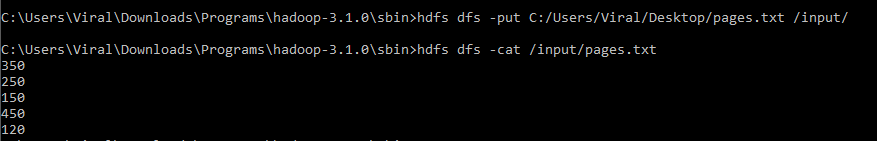
**Aim: Write a map-reduce program to count the books from distributed.**

**Steps:**

**1.** Create a text file pages.txt and store page numbers line by line in that file.

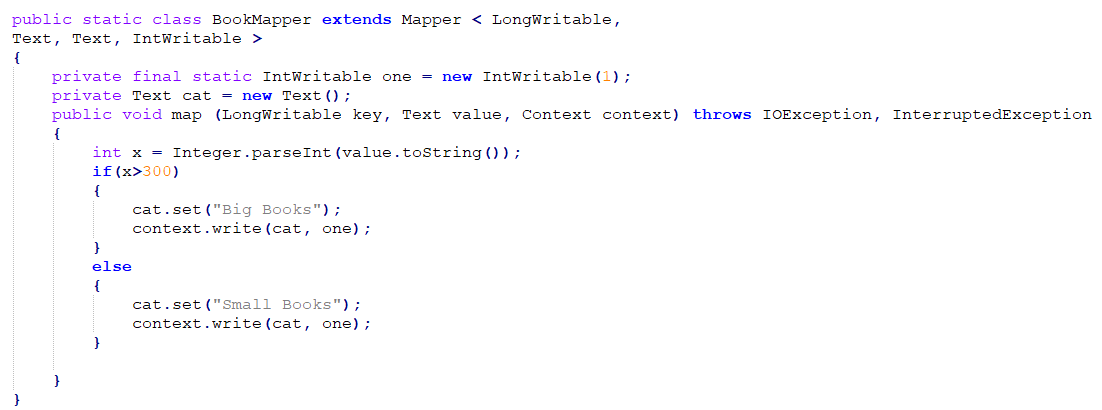


**2.** Copy that file to hdfs from local storage and we can see content of file using cat command.

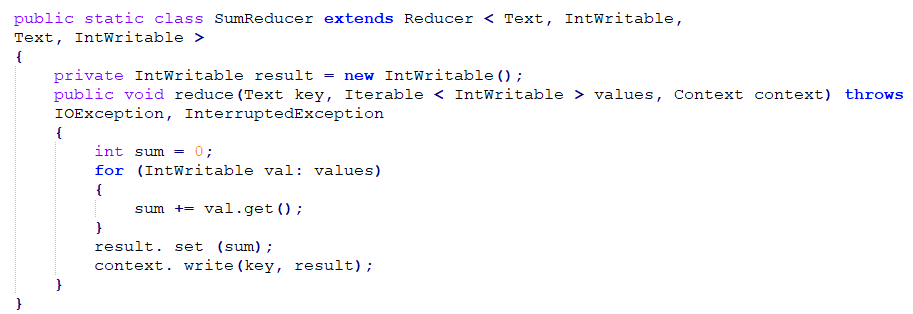


**3.** Create Bookcount.java class file which contains the mapper and reducer class for the particular problem statement.

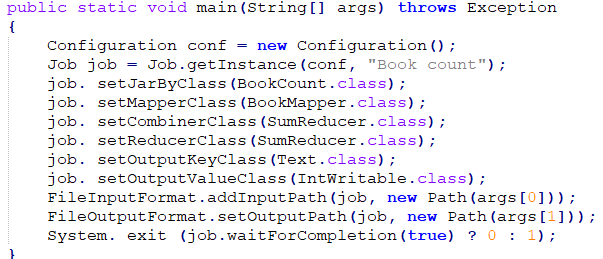
* Mapper



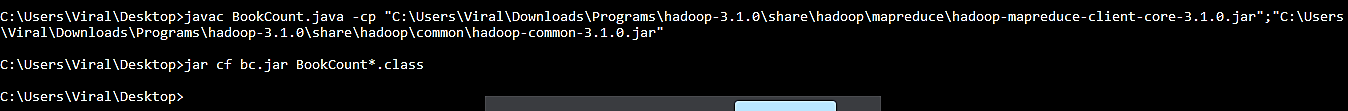
* Reducer



* Main method

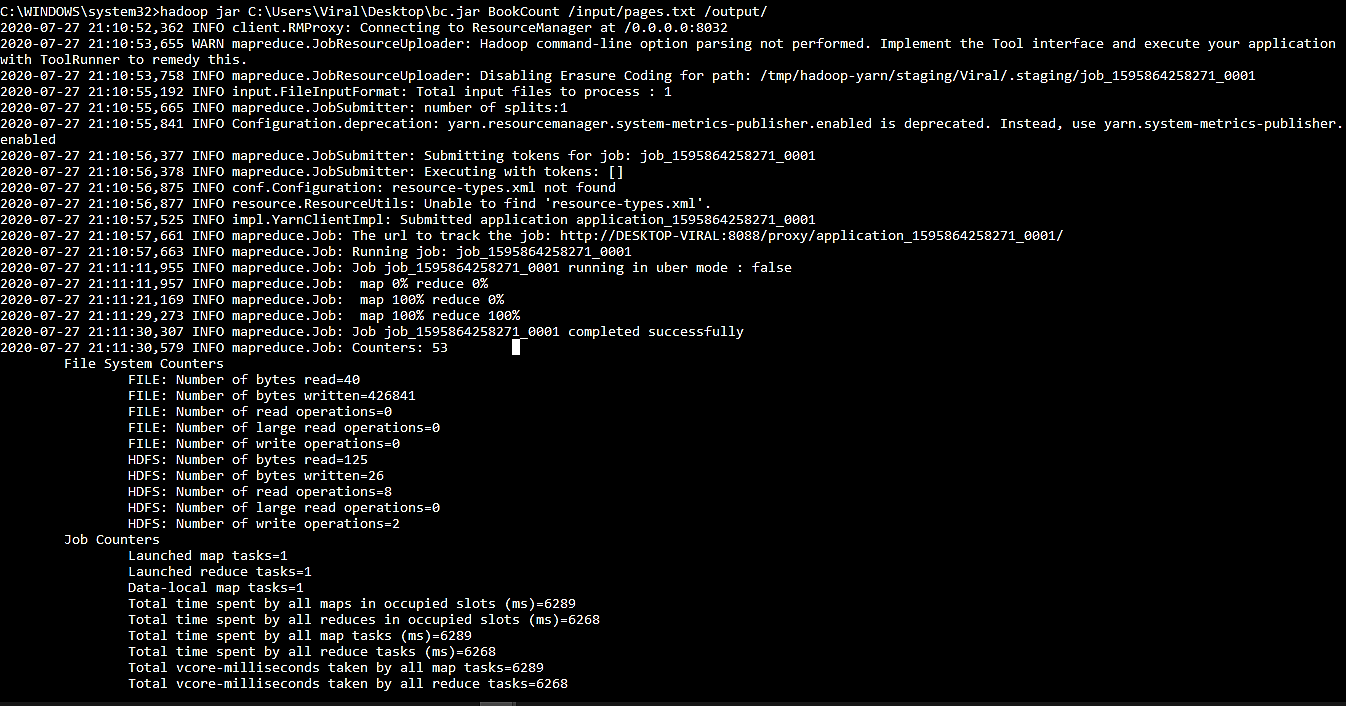


**4.** Compile and create the jar file bc.jar of MapReduce java file.



**5.** Now, by using “hadoop jar” command we can execute our program:

**Command: hadoop jar C:\Users\Viral\Desktop\bc.jar BookCount /input/pages.txt /output**



**6.** Now using cat command we can observe the output of file.

