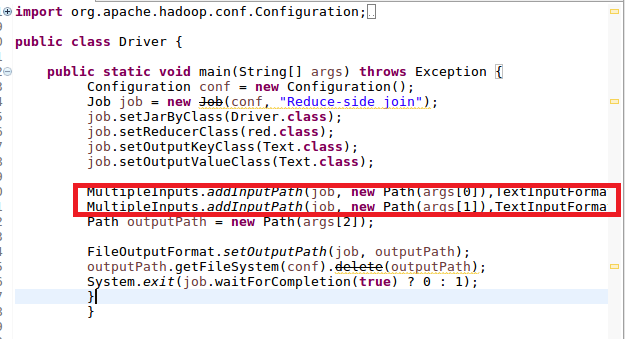
**Execute Session 16 - Assignment 1 (Machine and Sensor data analysis) using Reduce side join and compare which join is faster in normal conditions**.

**DRIVER**

Use mulitiple Input paths to get multiple in puts for mapper



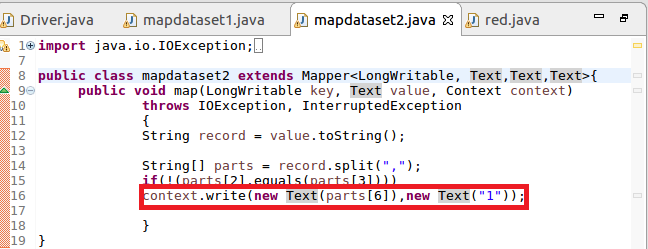
**MAPPER1**

In mapper 1 sending the BUILDING ID and Country as key and value respectively



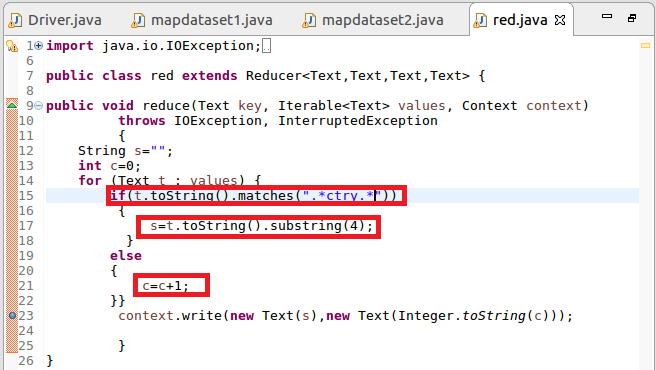
**MAPPER2**

In this checked for variation(actual temperature not equal to targeted temperature) and then sent BUILDING ID AS KEY AND A COUNT OF 1 AS VALUE

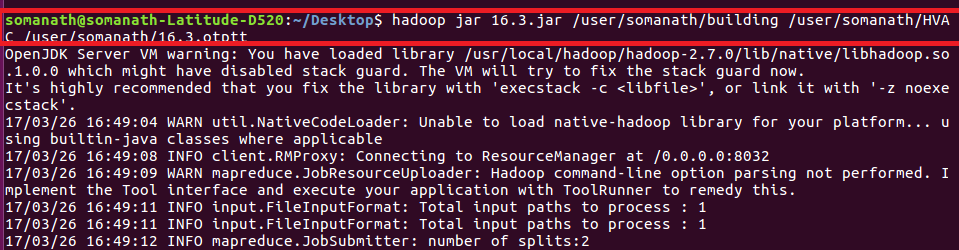


**REDUCER**

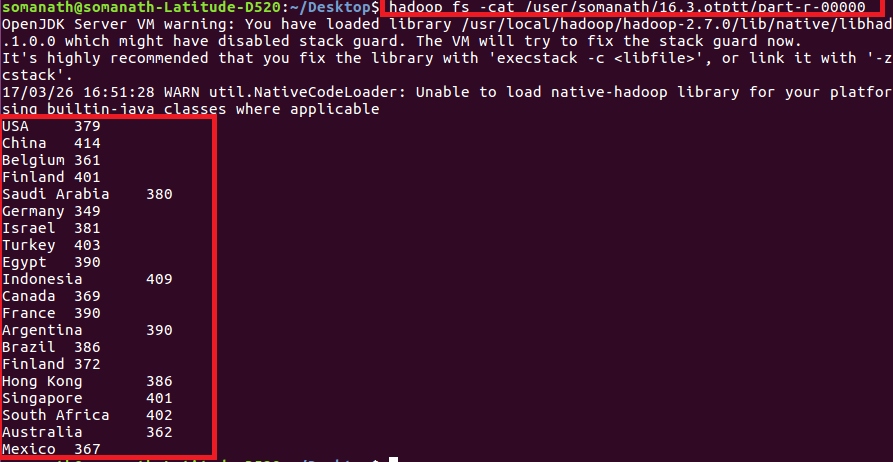
IF VALUE MATCHED “CTRY ” I USED SUBSTRING TO TAKE THE COUNTRY FOR A GIVEN ID ELSE MADE TOTAL COUNT AND GAVE COUNTRY AS KEY AND TOTAL COUNT AS VALUE



**RUNNING JAR**

****

**OUTPUT**

****

**Comparison**

joining the Reduce Side will happen after the completion of Map Phase. So data will come from different map/Machine.

-Reduce-Side joins are more simple than Map-Side joins since the input datasets need not to be structured.  
- It is slower and less efficient then the Map Side join.