**Session 16**

**Assignment 1**

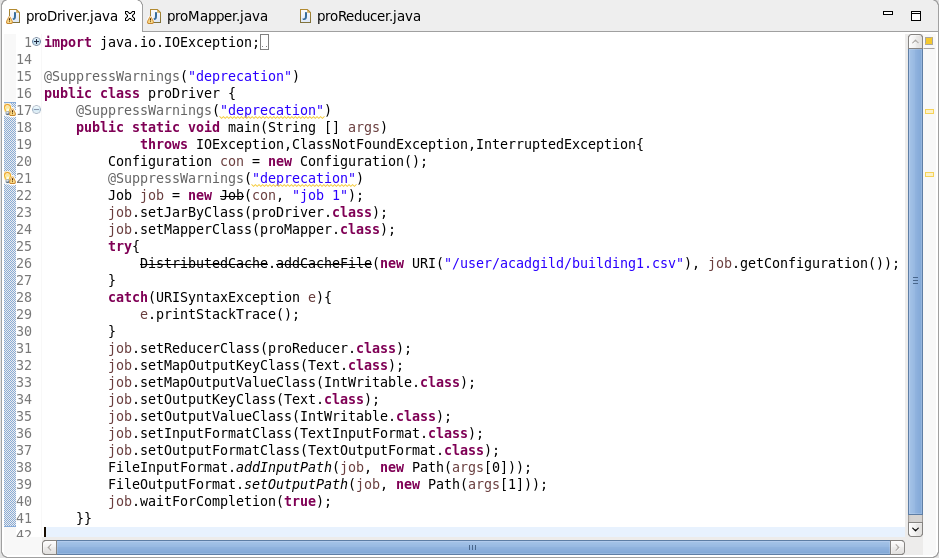
Using Distributed cache join the two datasets and find out the number of temperature variations in each country using MapReduce.

Here are the columns that are present in the datasets:

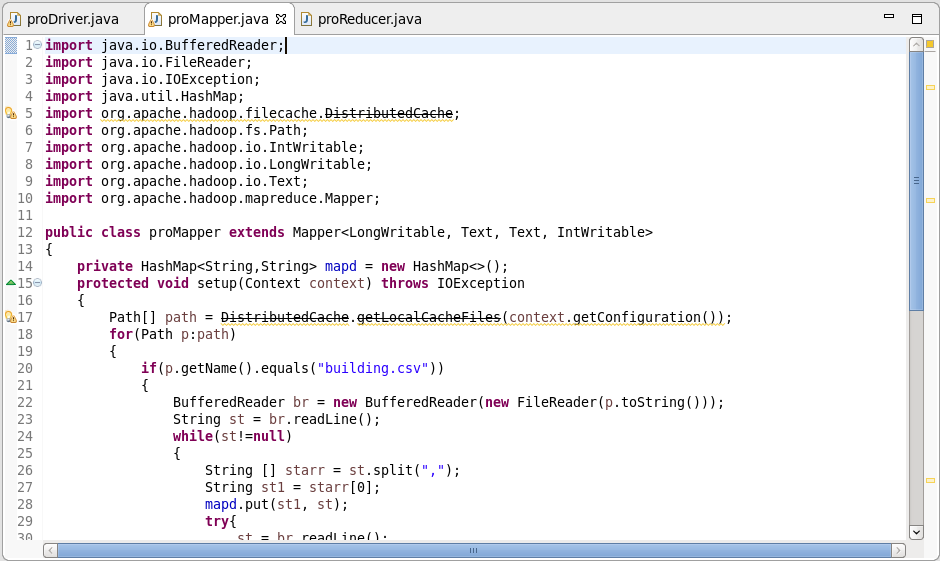
**Building.csv** – BuildingID, BuildingMgr, BuildingAge, HVACproduct, Country

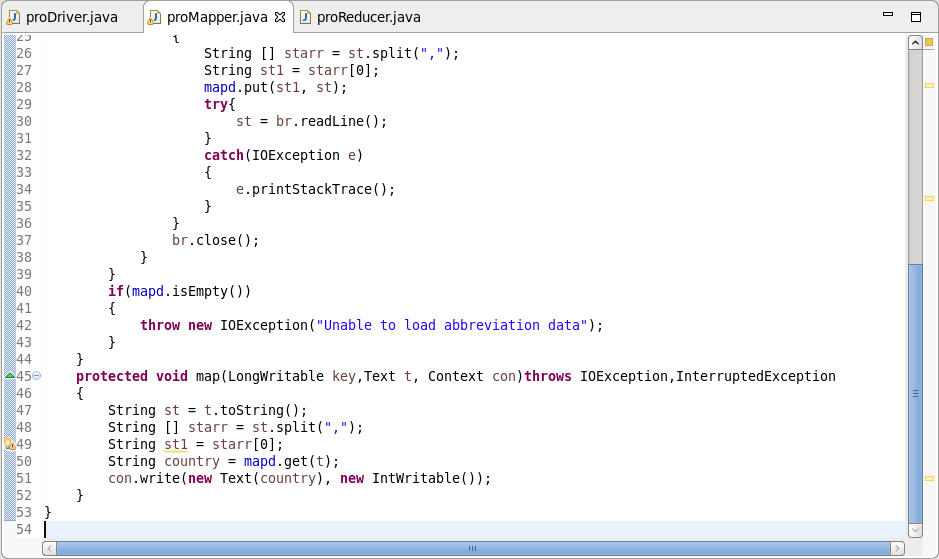
**HVAC.csv** – Date, Time, TargetTemp, ActualTemp, System, SystemAge, BuildingID

**Driver Class:**

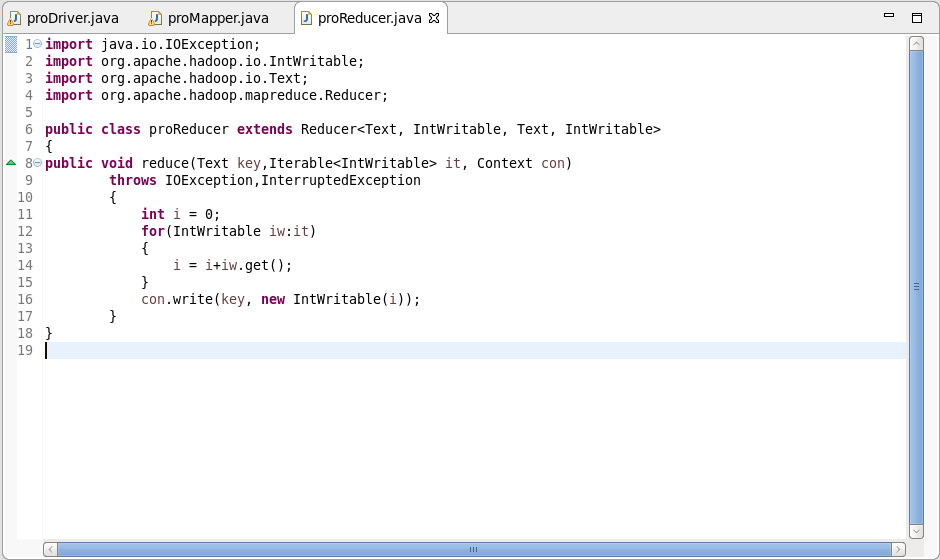


**Mapper Class:**





**Reducer Class:**



**Output:**

