**Session 21**

**Assignment 3**

**Problem Statement:**

The given dataset consists of two columns i.e., first\_name, last\_name & city.

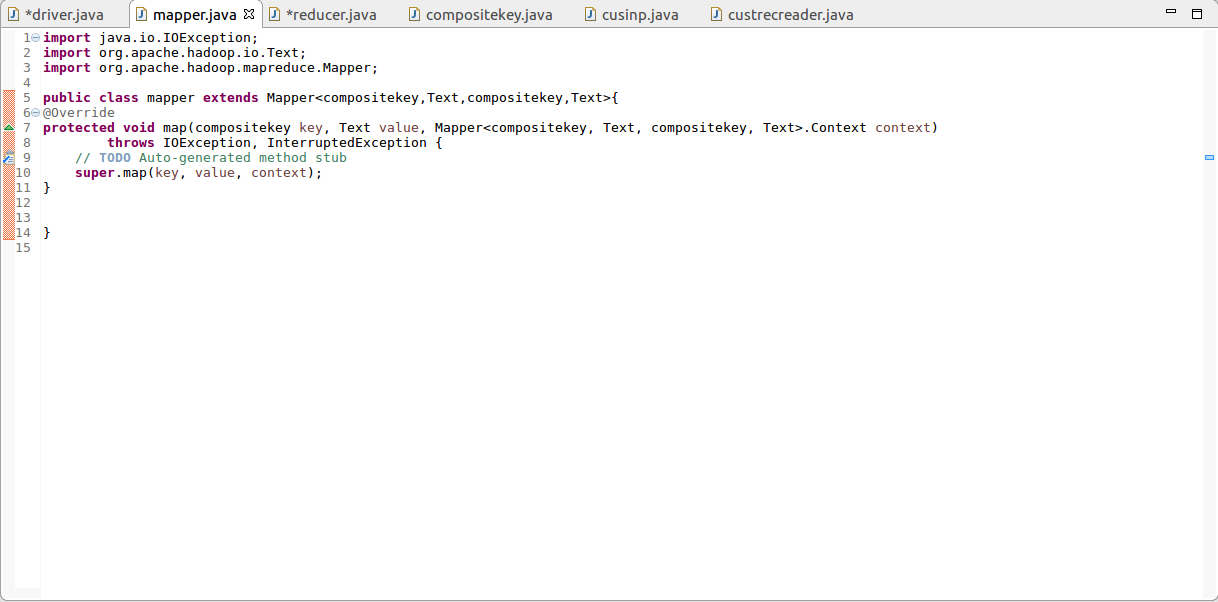
Use custom input format to read the key as the combination of first\_name and last\_name and arrange the output of the program should be the key sorted by last\_name.

* Since the problem wants the record reader to read first\_name and last\_name as key, a composite key can be defined which contains first\_name and last\_name and then sort last\_name by ascending order.
* For reading the file by composite key, a custom Input format class is created which accepts the parameters key and value.
* Now, these values are passed to the Record reader, which does the actual formatting of the inputs.

**Driver Class:**



**Mapper Class:**



**Reducer Class:**



**Composite Key Class:**





**Custom Input Class:**



**Custom Record Reader:**

