Assume, that, a file `customers.txt` has the following contents and the following header.

```
cust# name
00001 James
00002 Harry
00003 Peter
00004 Jane
...   ....
```

A file `orders.txt` has the following contents and the following header

```
order#  cust# value
0000001 00001 34.5
0000002 00001 23.0
0000003 00002 123.0
0000004 00003 12.3
...     ...   ...
```

Assume that both files have been loaded to HDFS.

Implementation of Map phase

A file `customer.txt` is converted into `<key,value>` pairs where `key = cust#` and `value = name`.
A file `orders.txt` is converted into into `<key,value>` pairs where `key = cust#` and `value = order#, value`

Implementation of Reduce phase

Reduce phase operates on two files `customers` and `orders` with `<key,value>` pairs where key is exactly the same for pairs. At this step we find all `<key,value>` pairs from `customer` such that does not exist at least one `<key,value>` pairs from `orders` that has the same value of `key`. Then such `<key,value>` pairs are written to `output`.