



DEEP
LEARNING
INSTITUTE



DLI Accelerated Data Science Teaching Kit

Lecture 10.4 - Example MapReduce Program



The Accelerated Data Science Teaching Kit is licensed by NVIDIA, Georgia Institute of Technology, and Prairie View A&M University under the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Example: Compute Word Counts

Input

“Apple Orange Mango Orange Grapes Plum”

“Apple Plum Mango Apple Apple Plum”

Output

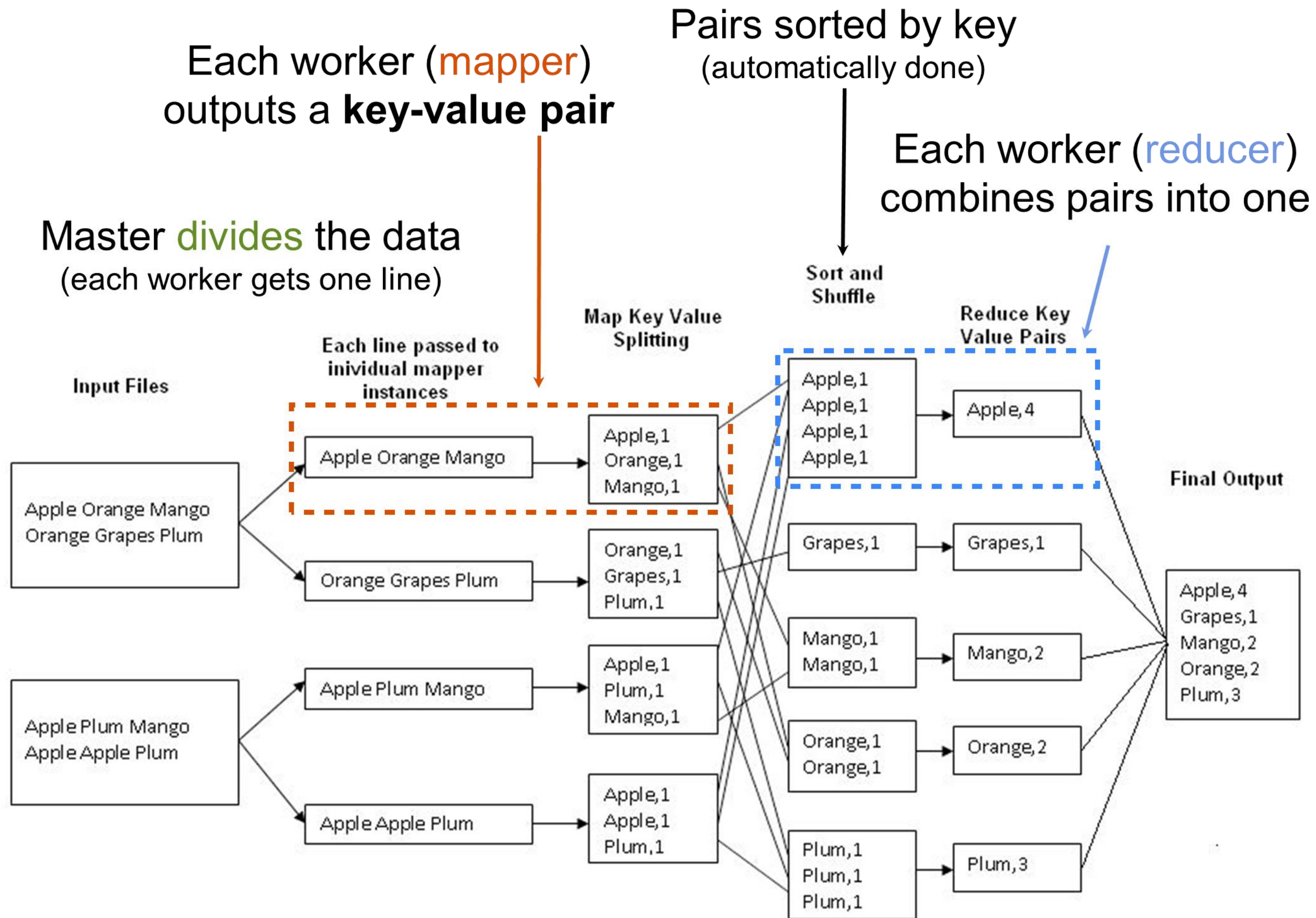
Apple, 4

Grapes, 1

Mango, 2

Orange, 2

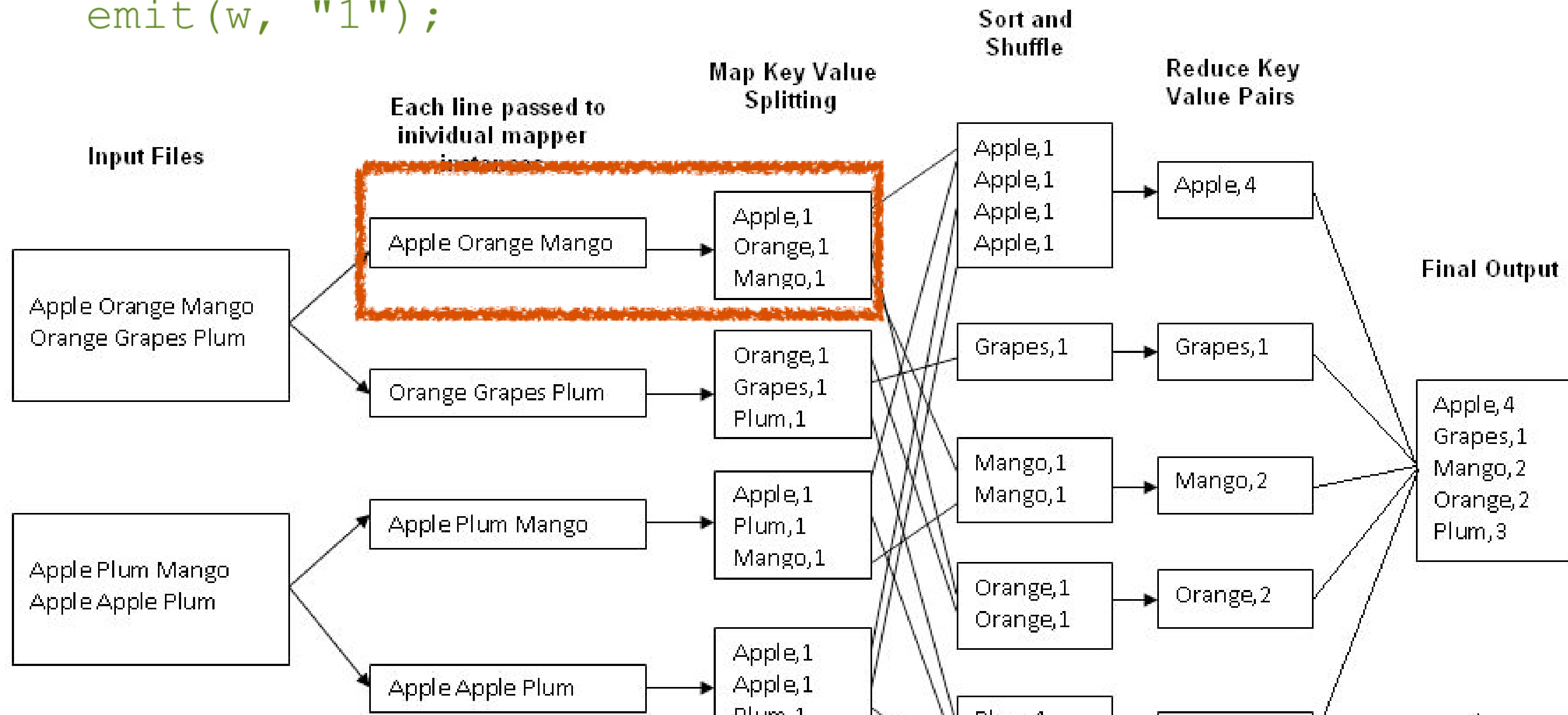
Plum, 3



A machine can be **both** a mapper and a reducer

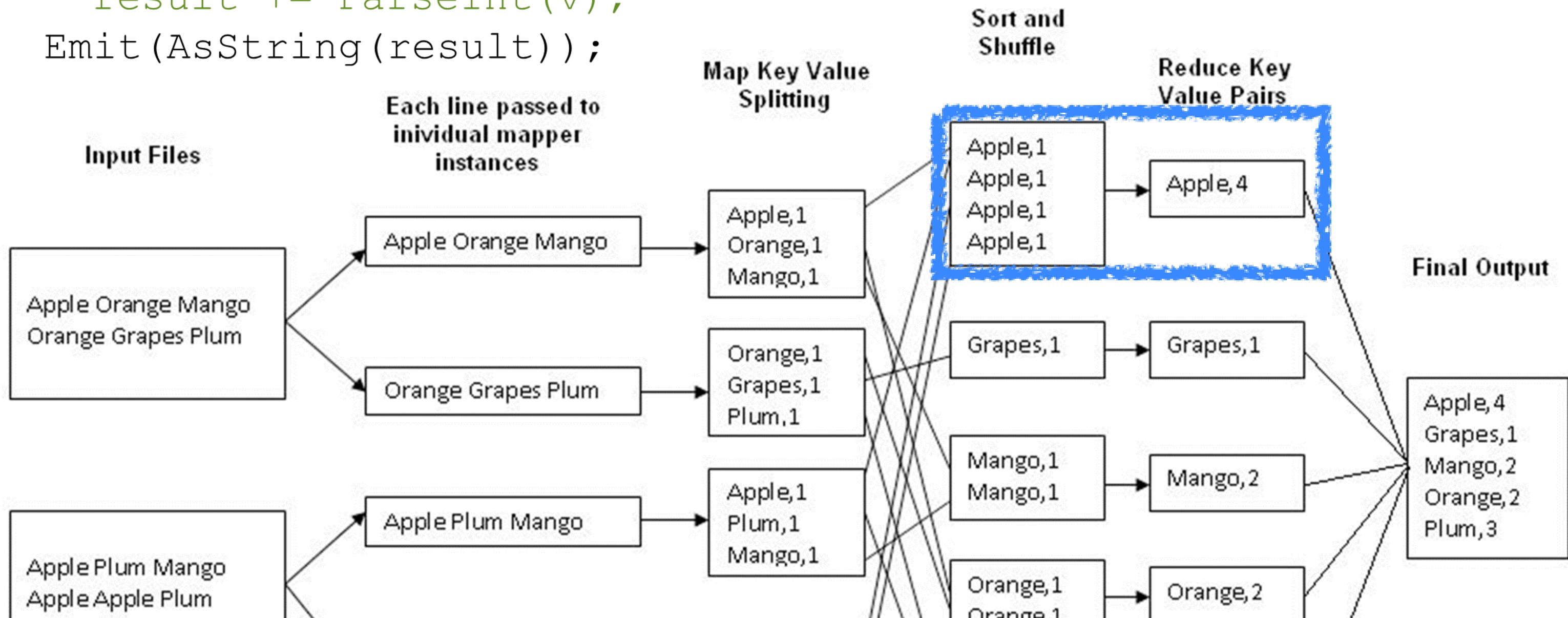
How to Implement This?

```
map(String key, String value):  
    // key: document id  
    // value: document contents  
    for each word w in value:  
        emit(w, "1");
```



How to Implement This?

```
reduce(String key, Iterator values):  
    // key: a word  
    // values: a list of counts  
    int result = 0;  
    for each v in values:  
        result += ParseInt(v);  
    Emit(AsString(result));
```





DEEP
LEARNING
INSTITUTE



PRAIRIE VIEW
A&M UNIVERSITY

DLI Accelerated Data Science Teaching Kit

Thank You