1. 优化方式一：We uses a hash table “combine” to maintain the words and counts per line in each call to the map function. After processing each line it emits the counts for this line. 这个方法会将每行中相同的word合并计数，例如原来我们需要两次print(f"{word}\t1,1")，使用这个方法后，我们只需要一次print(f"{word}\t1,{count}")，也就是print(f"{word}\t1,2")
2. 优化方法二：In this approach, each mapper will use a hash table to maintain the words and counts across all lines in a single mapper. By aggregating tuples across map tasks, this will obviously decrease the amount of data exchanged between the mappers and reducers for this program.

2.

输出只保留了正确输出的一部分（从24行减少至15行），并且在运行了三次之后，每次运行的输出结果都相同