wordcount_mapper.py

```
#!/usr/bin/python
import sys
for line in sys.stdin:
        words = line.split() #default split() arument is space
        for word in words:
                print ('{0}\t{1}'.format(word,1))
```

wordcount_reducer.py

```
#!/usr/bin/python
import sys
curr_word = None
curr_count = 0

# Process each key-value pair from the mapper
for line in sys.stdin:
  # Get the key and value from the current line
  word, count = line.split('\t')
  # Convert the count to an int
  count = int(count)
  # If the current word is the same as the previous word,
  # increment its count, otherwise print the words count
  # to stdout
  if word == curr_word:
    curr_count += count
  else:
    # Write word and its number of occurrences as a key-value
    # pair to stdout
    if curr_word:
        print ('{0}\t{1}'.format(curr_word, curr_count))
    curr_word = word
    curr_count = count
# Output the count for the last word
if curr_word == word:
        print ('{0}\t{1}'.format(curr_word, curr_count))
```

Command

```
echo "big data is big" | sort | ./wordcount_mapper.py | ./wordcount_reducer.py
```