# Experiment No. 5

**5.1 Aim:** Write a program to implement a word count program for MapReduce using either java or python **.**

**5.2 Course Outcome:** Apply Map reduce paradigm to solve problems.

**5.3 Learning Objectives:** To implement the mapper and reducer function for counting the words in a document and printing the same in part file of HDFS.

**5.4 Output:**

```
from io import TextIOWrapper import
threading


# Mapper Function
def mapper(block : list) :
  inter1 = open("intermediate1.txt","a")
  for line in block:      for
word in line.split() :
    inter1.write(word + ":1\n")
inter1.close()


# Combiner Function def
combiner():
  inter1 = open("intermediate1.txt","r")
inter2 = open("intermediate2.txt","a")
word_line_map = dict()     lines =
inter1.readlines()     for line in lines:
    word = line.split(':')[0]
    try:
      word_line_map[word].append(line)
    except:
      word_line_map[word] = [line]
values = list(word_line_map.values())
for i in range(len(values)):     for val in
values[i]:       inter2.write(val)     if
i!=len(values)-1:
      inter2.write("breakline\n")
inter1.close()   inter2.close()

# Reducer Function def reducer(block :
list,op : TextIOWrapper):
  op.write(block[0].split(":")[0] + ":" + str(len(block)) +"\n")




def MapParallely(m,b1,b2):
  thread1 = threading.Thread(target=m,args=[b1])
thread2 = threading.Thread(target=m,args=[b2])
  thread1.start()
thread2.start()
thread1.join()   thread2.join()

def ReduceParallely(r,blocks,op):   for block in blocks:
thread = threading.Thread(target=r,args=[block,op])
    thread.start()
```

```
# ****************  Driver Code  **********************

# File Opening file =
open("words.txt","r") lines
= file.readlines()



# Creating Blocks n =
len(lines) block1 =
lines[:n//2:] block2 =
lines[n//2+1::]



# Running Mapper Functions parallely
MapParallely(mapper,block1,block2)

# Combining the results so that same word pairs go to same reducer combiner()

# In intermediate2.txt output of combiner is stored inter2
= open("intermediate2.txt")

# The Main output file output =
open("output.txt","a")

# creating blocks of same word pairs
blocks = [] lines=[]
for line in
inter2.readlines():
if line ==
"breakline\n":
blocks.append(lin
es)
        lines=[]
continue     else:
        lines.append(line)

# Reducing all the blocks parallely
ReduceParallely(reducer,blocks,output)

# Printing the output output =
open("output.txt","r")
for line in output.readlines():
    print(line)



# Deleting the content of files so that for next run they are empty
inter1 = open("intermediate1.txt","w") inter2 =
open("intermediate2.txt","w") output = open("output.txt","w")
inter1.truncate() inter2.truncate() output.truncate() inter1.close()
inter2.close() output.close()
```

## Input_Words :

Maybe it's the way you say my name

Maybe it's the way you play your game

But it's so good I've never known anybody like you

But it's so good I've never dreamed of nobody like you

And I've heard of a love that comes once in a lifetime

And I'm pretty sure that you are that love of mine

Cause I'm in a field of dandelions

Wishing on every one that you'll be mine mine

And I see forever in your eyes

I feel okay when I see you smile smile

Wishing on dandelions all of the time

Praying to God that one day you'll be mine

Wishing on dandelions all of the time all of the time

I think that you are the one for me

Cause it gets so hard to breathe

When you're looking at me I've never felt so alive and free

When you're looking at me I've never felt so happy

And I've heard of a love that comes once in a lifetime

And I'm pretty sure that you are that love of mine

Cause I'm in a field of dandelions

Wishing on every one that you'll be mine mine

And I see forever in your eyes

I feel okay when I see you smile smile

Wishing on dandelions all of the time

Praying to God that one day you'll be mine

Wishing on dandelions all of the time all of the time

Dandelion into the wind you go

Won't you let my darling know

Dandelion into the wind you go

Won't you let my darling know that

I'm in a field of dandelions

Wishing on every one that you'll be mine mine

And I see forever in your eyes

I feel okay when I see you smile smile

Wishing on dandelions all of the time

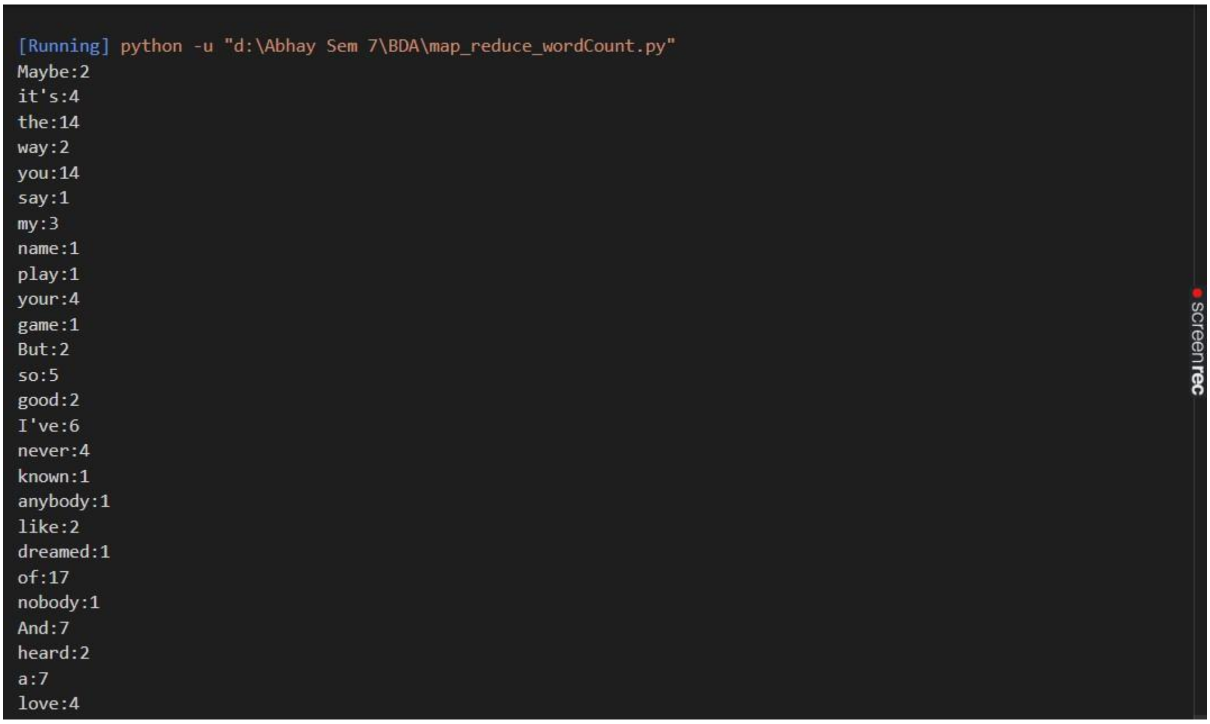
Praying to God that one day you'll be mine

Wishing on dandelions all of the time all of the time

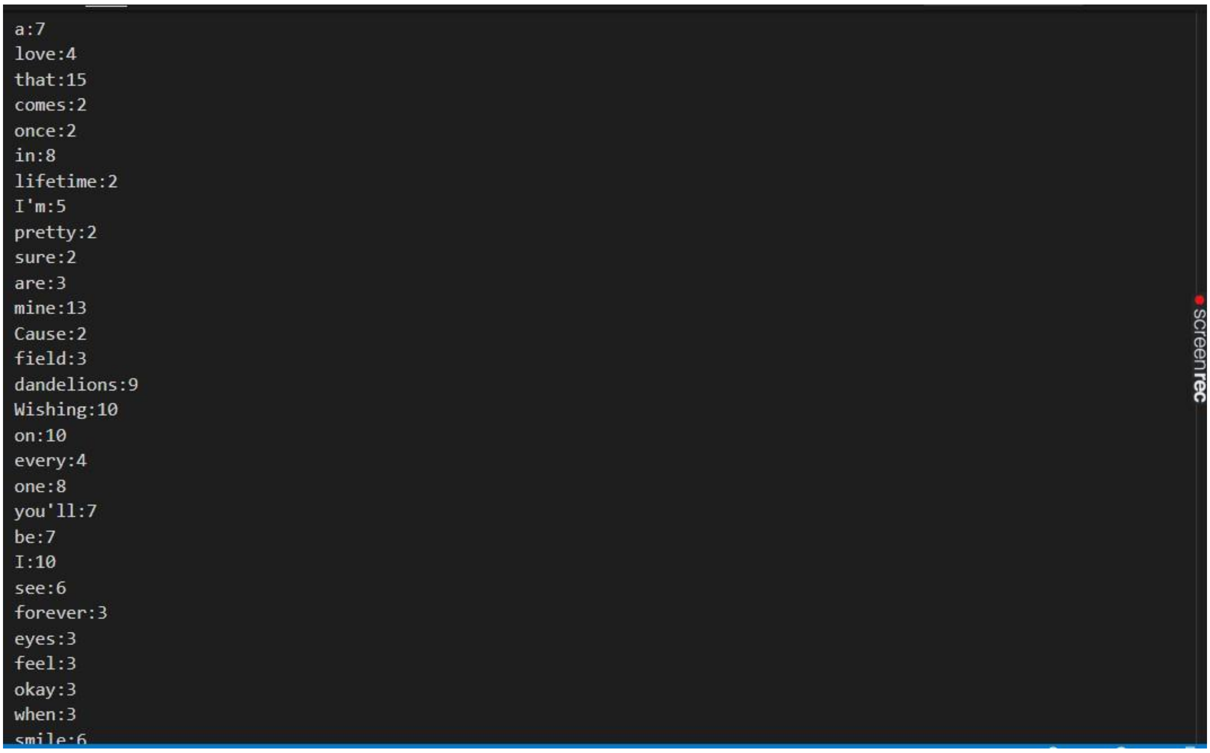
I'm in a field of dandelions

Wishing on every one that you'll be mine mine

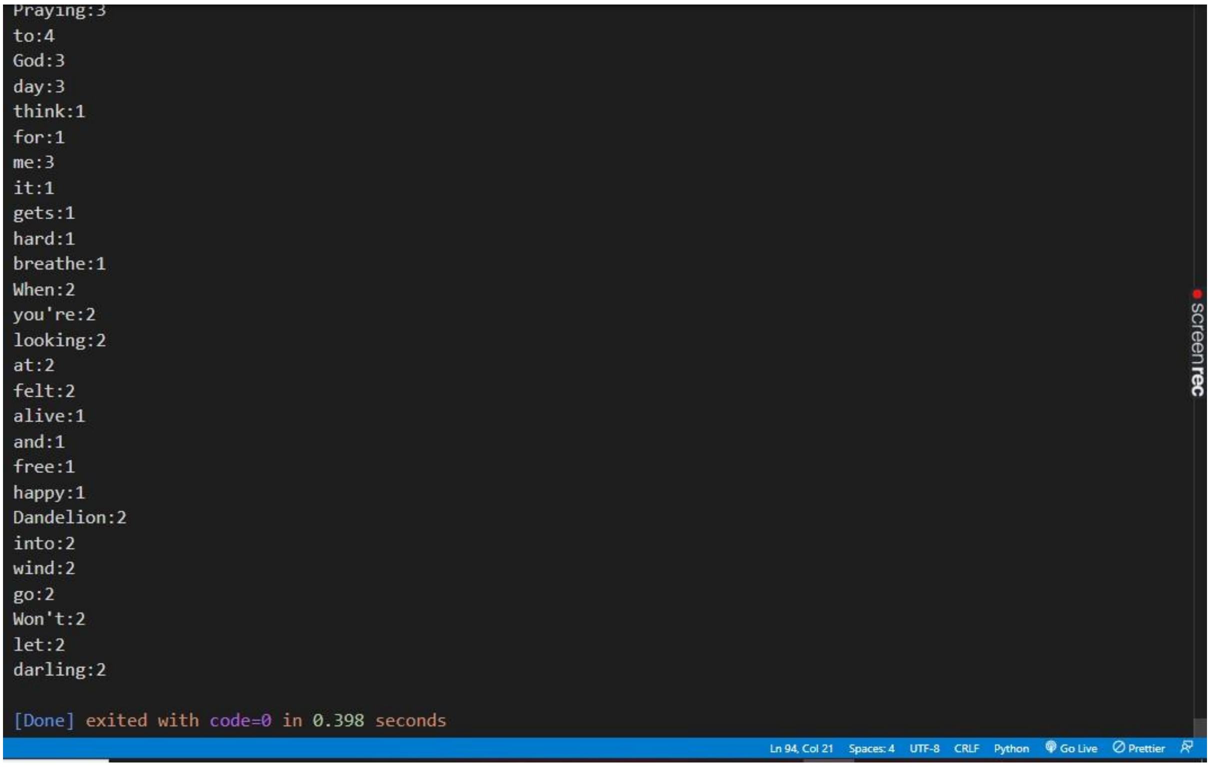
## Output :



```
[Running] python -u "d:\Abhay Sem 7\BDA\map_reduce_wordCount.py"
Maybe:2
it's:4
the:14
way:2
you:14
say:1
my:3
name:1
play:1
your:4
game:1
But:2
so:5
good:2
I've:6
never:4
known:1
anybody:1
like:2
dreamed:1
of:17
nobody:1
And:7
heard:2
a:7
love:4
```



```
a:7
love:4
that:15
comes:2
once:2
in:8
lifetime:2
I'm:5
pretty:2
sure:2
are:3
mine:13
Cause:2
field:3
dandelions:9
Wishing:10
on:10
every:4
one:8
you'll:7
be:7
I:10
see:6
forever:3
eyes:3
feel:3
okay:3
when:3
smile:6
```



```
Praying:3
to:4
God:3
day:3
think:1
for:1
me:3
it:1
gets:1
hard:1
breathe:1
When:2
you're:2
looking:2
at:2
felt:2
alive:1
and:1
free:1
happy:1
Dandelion:2
into:2
wind:2
go:2
Won't:2
let:2
darling:2

[Done] exited with code=0 in 0.398 seconds
```