

DATA 603 ASSIGNMENT 2

BHEEMISHETTY SAI BABA
US53017
saibabb1@umbc.edu

Chapter 1

Birth date: 30- Jan - 1999

Ans 1 :

Code :

```
# import the MRJob library
from mrjob.job import MRJob

# define a new MRJob class for counting words
class WordCount(MRJob):

    # define a mapper function
    def mapper(self, _, line):
        # split the line into words and emit each word as a key with value 1
        for word in line.strip().split():
            # yield each word as a key with value 1
            # use lower() to convert the word to lowercase
            yield word.lower(), 1

    # define a reducer function
    def reducer(self, word, counts):
        # sum the counts for each word and emit the result
        # the word argument is the key, and counts is an iterator over the values
        yield word, sum(counts)

#code runs the MRJob class
if __name__ == '__main__':
    WordCount.run()
```

Code Snap shot:

```
jupyter count_occurrences.py ✓ 21 hours ago
File Edit View Language

1 # import the MRJob library
2 from mrjob.job import MRJob
3
4 # define a new MRJob class for counting words
5 class WordCount(MRJob):
6
7     # define a mapper function
8     def mapper(self, _, line):
9         # split the line into words and emit each word as a key with value 1
10         for word in line.strip().split():
11             # yield each word as a key with value 1
12             # use lower() to convert the word to lowercase
13             yield word.lower(), 1
14
15     # define a reducer function
16     def reducer(self, word, counts):
17         # sum the counts for each word and emit the result
18         # the word argument is the key, and counts is an iterator over the values
19         yield word, sum(counts)
20
21 #code runs the MRJob class
22 if __name__ == '__main__':
23     WordCount.run()
24
```

Out put :

(Since I have done this in pycharm, I have executed the file in terminal and the output was taken as the snapshots)

```
ner.py", line 1133, in _check_input_paths
    self._check_input_path(path)
  File "C:\Users\babas\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.10_qbz5n2kfra8p0\LocalCache\local-packages\Python310\site-packages\mrjob\run
ner.py", line 1146, in _check_input_path
    raise IOError(
OSError: Input path Harry_Potter_assign does not exist!
PS C:\Users\babas> python count_occurrences.py file1.txt
No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory C:\Users\babas\AppData\Local\Temp\count_occurrences.babas.20230302.041240.911831
Running step 1 of 1...
job output is in C:\Users\babas\AppData\Local\Temp\count_occurrences.babas.20230302.041240.911831\output
Streaming final output from C:\Users\babas\AppData\Local\Temp\count_occurrences.babas.20230302.041240.911831\output...
"18" 1
"19" 1
"20" 1
"21" 1
"22" 1
"23" 1
"\u0092atta" 1
"\u0093all" 2
"\u0093and" 2
"\u0093are" 1
"\u0093bad" 1
"\u0093comb" 1
"\u0093darling," 1
"\u0093good" 1
"\u0093i" 1
"\u0093i\u0092d" 1
"\u0093in" 1
"\u0093little" 1
"\u0093mrs." 1
"\u0093nearly,\u0094" 1
"\u0093nothing," 1
"\u0093now" 1
"\u0093oh,\u0094" 1
"\u0093so" 1
"\u0093that\u0092s" 2
"\u0093thirty-nine," 1
"\u0093thirty-six,\u0094" 1
```

"\u0093to"	1	
"\u0093up!"	1	
"\u0093up!\u0094"		1
"\u0093well,"	1	
"\u0093well,\u0094"		1
"\u0093what"	1	
"\u0093yeah,\u0094"		1
"\u0094"	2	
"\u0096"	6	
"\u0097"	8	
"a"	49	
"about"	4	
"adventure"	1	
"after"	2	
"again."	1	
"age."	1	
"ago,"	1	
"air"	1	
"airplane,"		1
"all"	8	
"all."	1	
"almost"	2	
"always"	1	
"and"	46	
"and,"	1	
"angel"	1	
"angry"	1	
"another"	2	
"answer"	1	
"appearance"	1	
"around"	1	
"arrived"	1	
"as"	11	
"ask"	2	
"asking"	1	
"asleep"	1	
"astonishing"	1	
"at"	10	
"aunt"	11	
"auntie"	1	

```

"which" 3
"while" 2
"who" 3
"whole" 1
"why" 1
"wig." 1
"wiping" 1
"with" 9
"without" 1
"woke" 1
"woken" 2
"wolfing" 1
"wore" 1
"work." 1
"worried." 1
"worth," 1
"would" 3
"wristwatch" 1
"www.ztcprep.com" 10
"year" 1
"year," 1
"year.\u0094" 1
"years" 2
"yet" 1
"yet?\u0094" 1
"you" 8
"your" 2
"|" 6
Removing temp directory C:\Users\babas\AppData\Local\Temp\count_occurrences.babas.20230302.041240.911831...
PS C:\Users\babas> |

```

Ans 2 :

Code:

#Birthdate: 30-jan-1999

importing the necessary libraries
from functools import reduce

choosing the file path
file_path = "file2.txt"

function to split the string into words
def split_words(string):
 return string.split()

Read the file and split it into words
with open(file_path, "r") as file:
 words = file.read().split()

Use map to create a list of tuples (word, 1)
word_count_pairs = list(map(lambda x: (x, 1), words))

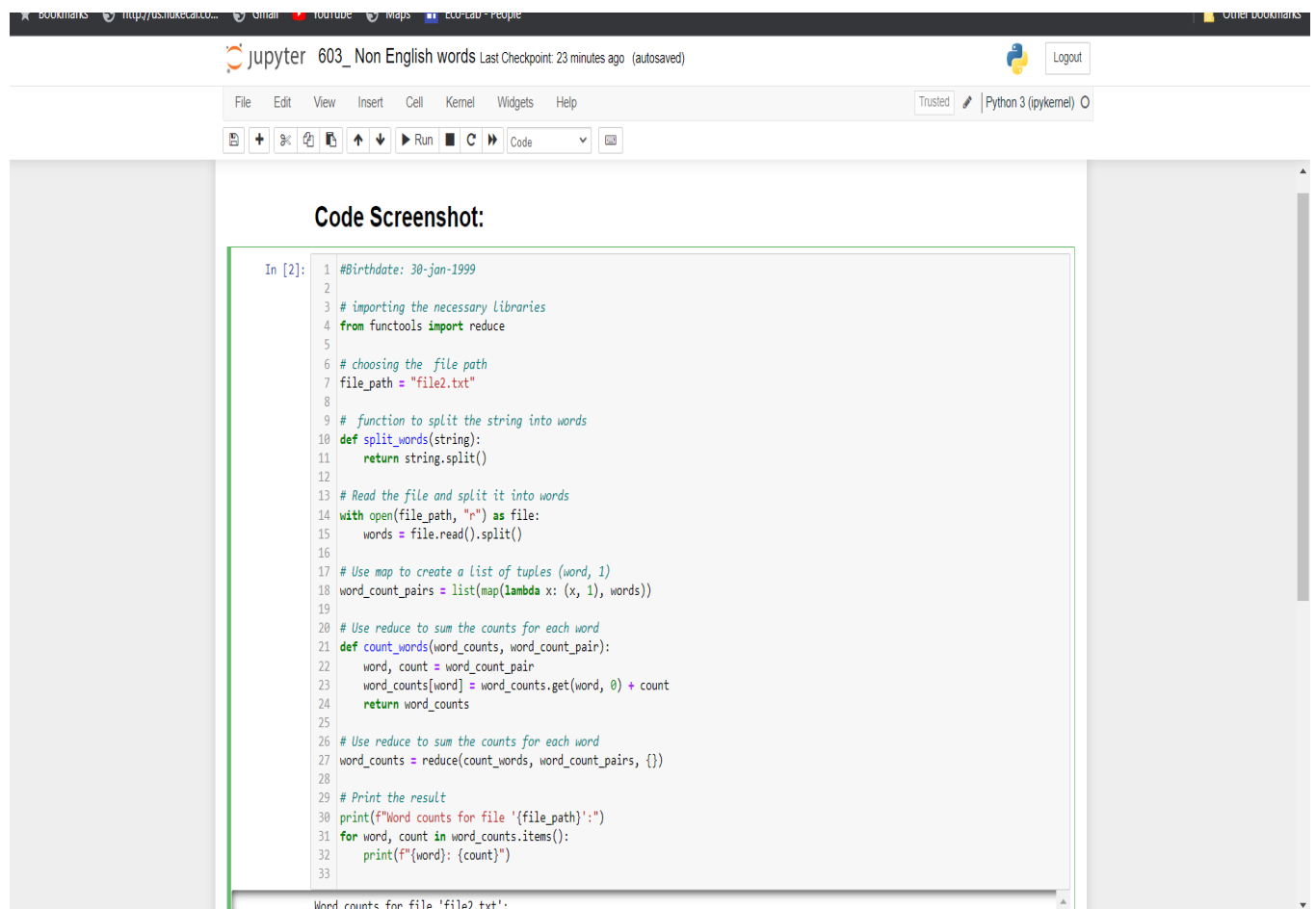
Use reduce to sum the counts for each word
def count_words(word_counts, word_count_pair):
 word, count = word_count_pair

```
word_counts[word] = word_counts.get(word, 0) + count
return word_counts
```

```
# Use reduce to sum the counts for each word
word_counts = reduce(count_words, word_count_pairs, {})
```

```
# Print the result
print(f"Word counts for file '{file_path}':")
for word, count in word_counts.items():
    print(f"{word}: {count}")
```

Code Snapshot :



Code Screenshot:

```
In [2]: 1 #Birthdate: 30-Jan-1999
2
3 # importing the necessary libraries
4 from functools import reduce
5
6 # choosing the file path
7 file_path = "file2.txt"
8
9 # function to split the string into words
10 def split_words(string):
11     return string.split()
12
13 # Read the file and split it into words
14 with open(file_path, "r") as file:
15     words = file.read().split()
16
17 # Use map to create a list of tuples (word, 1)
18 word_count_pairs = list(map(lambda x: (x, 1), words))
19
20 # Use reduce to sum the counts for each word
21 def count_words(word_counts, word_count_pair):
22     word, count = word_count_pair
23     word_counts[word] = word_counts.get(word, 0) + count
24     return word_counts
25
26 # Use reduce to sum the counts for each word
27 word_counts = reduce(count_words, word_count_pairs, {})
28
29 # Print the result
30 print(f"Word counts for file '{file_path}':")
31 for word, count in word_counts.items():
32     print(f"{word}: {count}")
33
```

Word counts for file 'file2.txt':

Out put :

Word counts for file 'file2.txt':

Harry: 25

Potter: 9

and: 46

the: 55
Philosophers: 9
Stone: 9
—: 9
J.K.: 9
Rowling: 9
nothing: 2
a: 48
good: 2
beating: 1
wouldn't: 1
have: 3
cured: 1
—: 29
as: 4
for: 5
all: 6
this: 3
about: 5
your: 1
parents,: 1
well,: 1
they: 4
were: 2
weirdos,: 1
no: 3
denying: 1
it,: 2
world's: 1
better: 1
off: 5
without: 3
them: 1
in: 17
my: 4
opinion: 1
asked: 2
got,: 1
getting: 2
mixed: 1
up: 5
with: 6
these: 1
wizarding: 1
types: 1
just: 2

what: 3
I: 17
expected,: 1
always: 2
knew: 1
they'd: 2
come: 2
to: 26
sticky: 1
end: 2
—": 6
But: 5
at: 15
that: 7
moment,: 1
Hagrid: 13
leapt: 1
from: 2
sofa: 1
drew: 1
battered: 1
pink: 1
umbrella: 5
inside: 3
his: 28
coat.: 3
Pointing: 1
Uncle: 7
Vernon: 4
like: 5
sword,: 1
he: 38
said,: 1
"I'm: 3
warning: 2
you,: 1
Dursley: 1
I'm: 1
you: 11
one: 5
more: 4
word: 1
...": 1
In: 2
danger: 1
of: 20

being: 2
speared: 1
on: 16
an: 3
by: 5
bearded: 1
giant,: 1
Vernon's: 1
courage: 1
failed: 1
again,: 1
flattened: 1
himself: 4
against: 1
wall: 1
fell: 2
silent.: 1
"That's: 1
better,": 1
said: 9
Hagrid,: 5
breathing: 1
heavily: 1
sitting: 1
back: 5
down: 5
sofa,: 2
which: 1
time: 3
sagged: 1
right: 2
floor.: 1
Harry,: 4
meanwhile,: 1
still: 5
had: 13
questions: 1
ask,: 1
www.ztcprep.com: 11
hundreds: 1
them.: 3
"But: 2
happened: 2
Vol-,: 1
sorry: 1
mean,: 1

You-: 1
Know-Who?": 1
"Good: 1
question,: 1
Harry.: 4
Disappeared.: 1
Vanished.: 1
Same: 1
night: 3
tried: 3
ter: 18
kill: 1
you.: 1
Makes: 1
yeh: 7
even: 2
famous.: 1
That's: 1
biggest: 1
myst'ry,: 1
see: 1
...: 7
was: 28
gettin': 2
an': 10
powerful: 1
why'd: 1
go?: 1
"Some: 1
say: 3
died.: 1
Codswallop,: 1
opinion.: 1
Dunno: 1
if: 6
enough: 1
human: 1
left: 2
him: 9
die.: 1
Some: 2
he's: 3
out: 7
there,: 1
bidin': 1
time,: 1

like,: 1
but: 8
don': 1
believe: 1
it.: 3
People: 1
who: 3
side: 1
came: 4
P: 8
g: 8
e: 8
|: 8
63: 1
ours.: 1
'em: 1
outta: 1
kinda: 1
trances.: 1
Don': 1
reckon: 2
could've: 1
done: 1
comin': 1
back.: 1
"Most: 1
us: 2
there: 6
somewhere: 1
lost: 2
powers.: 1
Too: 1
weak: 1
carry: 1
on.: 1
'Cause: 1
somethin': 3
finished: 1
him,: 3
There: 2
goin': 2
hadn't: 3
counted: 2
dunno: 1
it: 10
was,: 1

does: 1
stumped: 1
right.”: 1
looked: 4
warmth: 1
respect: 1
blazing: 1
eyes,: 1
instead: 1
feeling: 1
pleased: 1
proud,: 1
felt: 2
quite: 1
sure: 1
been: 6
horrible: 1
mistake.: 2
A: 1
wizard?: 1
Him?: 1
How: 1
could: 3
possibly: 1
be?: 1
He’d: 1
spent: 1
life: 1
clouted: 1
Dudley,: 1
bullied: 1
Aunt: 3
Petunia: 3
Vernon,: 1
really: 1
wizard,: 2
why: 1
turned: 2
into: 6
warty: 1
toads: 1
every: 2
lock: 1
cupboard?: 1
If: 1
he’d: 2

once: 1
defeated: 1
greatest: 2
sorcerer: 1
world,: 1
how: 1
Dudley: 5
able: 1
kick: 1
around: 1
football?: 1
"Hagrid,: 1
quietly,: 1
"I: 4
think: 5
must: 1
made: 5
don't: 1
can: 2
be: 9
wizard.": 1
To: 1
surprise,: 1
chuckled.: 1
"Not: 1
eh?: 1
Never: 1
things: 1
happen: 1
when: 2
scared: 1
or: 2
angry?": 1
fire.: 1
Now: 1
odd: 1
thing: 1
ever: 3
aunt: 1
uncle: 1
furious: 1
he,: 1
upset: 1
angry: 1
chased: 1
Dudley's: 1

gang,: 1
somehow: 1
found: 1
their: 3
reach: 1
dreading: 1
going: 4
school: 3
ridiculous: 1
haircut,: 1
managed: 1
make: 1
grow: 1
very: 1
last: 3
hit: 1
64: 1
got: 6
revenge,: 1
realizing: 1
doing: 1
it?: 1
Hadn't: 1
set: 1
boa: 1
constrictor: 1
him?: 1
smiling,: 1
saw: 2
positively: 1
beaming: 1
him.: 3
"See?": 1
Hagrid.: 2
"Harry: 1
Potter,: 1
not: 4
wizard: 2
wait,: 1
you'll: 1
famous: 1
Hogwarts.": 1
wasn't: 2
give: 1
fight.: 1
"Haven't: 1

told: 2
going?": 1
hissed.: 1
"He's: 1
Stonewall: 1
High: 1
he'll: 2
grateful: 2
I've: 1
read: 1
those: 1
letters: 2
needs: 1
sorts: 1
rubbish: 1
spell: 1
books: 2
wands: 1
"If: 1
wants: 2
go,: 1
great: 1
Muggle: 1
won't: 3
stop: 2
him,": 2
growled: 1
"Stop: 1
Lily: 1
James: 1
Potter's: 1
son: 1
Hogwarts!: 1
Yer: 1
mad.: 1
His: 1
name's: 1
since: 1
born.: 1
He's: 1
finest: 1
witchcraft: 1
wizardry: 1
world.: 1
Seven: 1
years: 1

know: 1
himself.: 1
He'll: 1
youngsters: 1
own: 1
sort,: 1
fer: 4
change,: 1
under: 3
headmaster: 1
Hogwarts: 2
had,: 1
Albus: 1
Dumbled—": 1
AM: 1
NOT: 1
PAYING: 1
FOR: 1
SOME: 1
CRACKPOT: 1
OLD: 1
FOOL: 1
TO: 1
TEACH: 1
HIM: 1
MAGIC: 1
TRICKS!": 1
yelled: 1
Vernon.: 1
finally: 1
gone: 1
too: 1
far.: 1
seized: 1
whirled: 1
over: 3
head,: 1
"NEVER: 1
thundered,: 1
"—: 1
INSULT: 1
ALBUS: 1
DUMBLEDORE: 1
IN: 1
FRONT: 1
OF: 1

ME!": 1
He: 6
brought: 1
swishing: 1
through: 3
air: 1
point: 1
flash: 1
violet: 1
light,: 1
sound: 1
firecracker,: 1
sharp: 1
squeal,: 1
65: 1
next: 2
second,: 1
dancing: 1
spot: 1
hands: 1
clasped: 1
fat: 1
bottom,: 1
howling: 1
pain.: 1
When: 2
them,: 1
curly: 1
pig's: 1
tail: 1
poking: 1
hole: 1
trousers.: 1
roared.: 1
Pulling: 1
other: 1
room,: 1
cast: 2
terrified: 1
look: 2
slammed: 1
door: 1
behind: 1
stroked: 1
beard.: 1
"Shouldn'ta: 1

me: 6
temper,: 1
ruefully,: 1
"but: 1
didn't: 5
work: 1
anyway.: 1
Meant: 1
turn: 1
pig,: 1
suppose: 1
so: 5
much: 2
pig: 1
anyway: 1
do.": 1
sideways: 1
bushy: 1
eyebrows.: 1
"Be: 1
mention: 1
anyone: 1
Hogwarts,: 1
said.: 2
er: 2
supposed: 2
do: 6
magic,: 1
strictly: 1
speakin'.: 1
allowed: 1
bit: 3
follow: 1
get: 4
yer: 5
stuff: 2
o': 4
reasons: 1
keen: 1
take: 1
job: 1
"Why: 2
aren't: 1
magic?": 1
"Oh,: 1
well: 1

meself: 1
expelled,: 1
tell: 4
truth.: 1
third: 1
year.: 1
They: 1
snapped: 2
wand: 1
half: 1
everything.: 1
Dumbledore: 1
let: 1
stay: 1
gamekeeper.: 1
Great: 1
man,: 1
Dumbledore.”: 1
expelled?”: 1
“It’s: 1
late: 1
we’ve: 1
lots: 2
tomorrow,”: 1
loudly.: 2
“Gotta: 1
town,: 1
that.”: 2
66: 1
took: 1
thick: 1
black: 1
coat: 3
threw: 1
“You: 1
kip: 1
that,”: 2
“Don’: 1
mind: 1
wiggles: 1
bit,: 1
couple: 1
dormice: 1
pockets.”: 2
67: 1
DIAGON: 1

ALLY: 1
woke: 1
early: 1
morning.: 1
Although: 1
daylight,: 1
kept: 1
eyes: 2
shut: 1
tight.: 1
"It: 1
dream,": 1
firmly.: 1
dreamed: 1
giant: 1
called: 1
wizards.: 1
open: 3
I'll: 2
home: 1
cupboard.": 1
suddenly: 1
loud: 1
tapping: 1
noise.: 1
And: 1
there': 1
s: 1
knocking: 1
door,: 1
thought,: 1
heart: 1
sinking.: 1
eyes.: 1
It: 1
such: 1
dream.: 1
Tap.: 3
"All: 1
right,": 1
mumbled,: 1
up.": 1
sat: 2
Hagrid's: 3
heavy: 1
The: 3

hut: 1
full: 1
sunlight,: 1
storm: 1
over,: 1
asleep: 1
collapsed: 1
68: 1
owl: 6
rapping: 1
its: 3
claw: 1
window,: 1
newspaper: 2
held: 2
beak.: 1
scrambled: 1
feet,: 1
happy: 2
though: 2
large: 1
balloon: 2
swelling: 1
went: 1
straight: 1
window: 1
jerked: 1
open.: 1
swooped: 1
dropped: 2
top: 1
wake: 1
up.: 1
then: 1
fluttered: 1
onto: 1
floor: 1
began: 1
attack: 1
"Don't: 2
wave: 1
way,: 1
beak: 1
fiercely: 1
carried: 1
savaging: 1

“Hagrid!”: 1
“There’s: 1
“Pay: 1
grunted: 1
sofa.: 1
“What?”: 1
“He: 1
payin’: 1
deliverin’: 1
paper.: 1
Look: 1
seemed: 1
pockets: 1
bunches: 1
keys,: 1
slug: 1
pellets,: 1
balls: 1
string,: 1
peppermint: 1
humbugs,: 1
teabags: 1
finally,: 1
pulled: 1
handful: 1
strange-looking: 1
coins.: 1
“Give: 1
five: 2
Knuts,”: 1
sleepily.: 1
“Knuts?”: 1
“The: 1
little: 2
bronze: 2
ones.”: 1
coins,: 1
leg: 1
put: 1
money: 2
69: 1
small: 1
leather: 1
pouch: 1
tied: 1
Then: 1

flew: 1
window.: 1
yawned: 1
loudly,: 1
up,: 1
stretched.: 1
"Best: 1
off,: 1
today,: 1
gotta: 1
London: 1
buy: 1
school.": 1
turning: 1
coins: 1
looking: 1
thought: 1
something: 1
feel: 1
puncture.: 1
"Um: 1
Hagrid?": 1
"Mm?": 1
pulling: 1
huge: 1
boots.: 1
haven't: 1
any: 1
heard: 1
pay: 1
go: 1
learn: 1
magic.": 1
worry: 1
standing: 1
scratching: 1
head.: 1
"D'yeh: 1
parents: 1
leave: 1
anything?": 1
house: 1
destroyed: 1
"They: 1
didn': 1
keep: 1

gold: 1
house,: 1
boy!: 1
Nah,: 1
first: 1
is: 1
Gringotts.: 2
Wizards': 1
bank.: 1
Have: 1
sausage,: 1
they're: 1
bad: 1
cold: 1
wouldn': 1
teh: 1
birthday: 1
cake,: 1
neither." : 1
"Wizards: 1
banks?": 1
"Just: 1
one.: 1
Run: 1
goblins." : 1
sausage: 1
holding.: 1
70: 1
"Goblins?": 1
"Yeah: 1
yeh'd: 1
mad: 1
try: 1
rob: 1

In []: