

#BIRTHDAY 14-10-1999

1. Write Python code and use MapReduce to count occurrences of each word in the first text file (file.txt). How many times each word is repeated?

```
from collections import defaultdict
import re
from functools import reduce

#text from the file using UTF-8 encoding
with open('file1.txt', 'r', encoding='utf-8') as file:
    text = file.read()

# Tokenize the text into words
words = WORD_RE.findall(text.lower())

# Split text into words
words = text.split()

# Define a mapper function to emit key-value pairs (word, 1)
def mapper(word):
    return (word, 1)

# Define a reducer function to sum the counts for each word
def reducer(counts, word_count):
    word, count = word_count
    counts[word] += count
    return counts

# Split the words into smaller chunks for mapping (e.g., by line)
chunk_size = len(words) // 4
word_chunks = [words[i:i + chunk_size] for i in range(0, len(words),
chunk_size)]

# Map
mapped_results = map(map_words, word_chunks)

# Reduce using functools.reduce
final_result = reduce(reduce_counts, mapped_results)

# Print the word counts
for word, count in final_result.items():
    print(f"{word}: {count}")

P: 11
a: 39
g: 11
e: 11
|: 11
```

14: 1  
Harry: 49  
Potter: 11  
and: 52  
the: 111  
Order: 11  
of: 43  
Phoenix: 11  
—: 11  
J.K.: 11  
Rowling: 11  
frustration: 2  
on: 16  
boys: 2  
who: 3  
had: 25  
once: 2  
made: 4  
his: 52  
life: 1  
hell: 1  
—: 28  
But: 5  
they: 6  
did: 3  
not: 10  
turn: 2  
around,: 3  
see: 5  
him,: 4  
were: 9  
almost: 2  
at: 13  
railings.: 1  
mastered: 1  
impulse: 1  
to: 43  
call: 2  
after: 5  
them...: 2  
Seeking: 1  
fight: 1  
was: 43  
smart: 1  
move...: 1  
He: 18  
must: 1  
use: 2  
magic...: 1

would: 4  
be: 8  
risking: 1  
expulsion: 1  
again...: 1  
Dudley's: 9  
gang's: 1  
voices: 2  
died;: 1  
out: 7  
sight,: 2  
heading: 1  
along: 3  
Magnolia: 7  
Road.: 1  
There: 5  
you: 21  
go,: 1  
Sirius,: 1  
thought: 2  
dully.: 1  
Nothing: 1  
rash.: 1  
Kept: 1  
my: 3  
nose: 1  
clean.: 1  
Exactly: 1  
opposite: 1  
what: 5  
you'd: 1  
have: 7  
done: 2  
...: 5  
got: 2  
feet: 3  
stretched.: 1  
Aunt: 1  
Petunia: 1  
Uncle: 3  
Vernon: 2  
seemed: 2  
feel: 2  
that: 17  
whenever: 1  
Dudley: 20  
turned: 5  
up: 9  
right: 4

time: 3  
home,: 1  
anytime: 1  
much: 2  
too: 1  
late.: 1  
threatened: 1  
lock: 1  
in: 28  
shed: 1  
if: 3  
he: 41  
came: 5  
home: 2  
again,: 3  
so,: 1  
stifling: 1  
yawn,: 1  
still: 1  
scowling,: 1  
set: 1  
off: 5  
toward: 3  
park: 1  
gate.: 1  
Road,: 1  
like: 8  
Privet: 1  
Drive,: 1  
full: 1  
large,: 2  
square: 2  
houses: 1  
with: 6  
perfectly: 1  
manicured: 1  
lawns,: 1  
all: 5  
owned: 1  
by: 4  
owners: 1  
drove: 1  
very: 2  
clean: 1  
cars: 3  
similar: 1  
Vernon's.: 1  
preferred: 1  
Little: 3

Whinging: 3  
night,: 3  
when: 6  
curtained: 1  
windows: 1  
patches: 1  
jewel-bright: 1  
colors: 1  
darkness: 2  
ran: 1  
no: 4  
danger: 1  
hearing: 1  
disapproving: 1  
mutters: 1  
about: 5  
"delinquent": 1  
appearance: 1  
passed: 1  
householders.: 1  
walked: 1  
quickly,: 1  
so: 3  
halfway: 2  
Road: 1  
gang: 2  
into: 13  
view: 1  
again;: 1  
saying: 1  
their: 5  
farewells: 1  
entrance: 1  
Crescent.: 1  
stepped: 1  
shadow: 1  
large: 2  
lilac: 1  
tree: 1  
waited.: 1  
15: 1  
"...: 1  
squealed: 1  
pig,: 1  
didn't: 3  
he?": 1  
Malcolm: 1  
saying,: 1  
guffaws: 1

from: 12  
others.: 1  
"Nice: 1  
hook,: 1  
Big: 4  
D,": 1  
said: 18  
Piers.: 1  
"Same: 1  
tomorrow?": 1  
Dudley.: 4  
"Round: 1  
place,: 1  
parents: 1  
are: 8  
out,": 1  
Gordon.: 1  
"See: 2  
then,": 1  
"Bye: 1  
Dud!": 1  
ya,: 1  
D!": 2  
waited: 1  
for: 8  
rest: 1  
move: 1  
before: 3  
setting: 1  
again.: 4  
When: 1  
faded: 1  
more: 3  
headed: 1  
around: 2  
corner: 1  
Crescent: 3  
walking: 1  
quickly: 1  
soon: 1  
within: 1  
hailing: 1  
distance: 1  
Dudley,: 6  
strolling: 1  
ease,: 1  
humming: 1  
tunelessly.: 1  
"Hey,: 1

turned.: 1  
"Oh,: 1  
grunted.: 1  
"It's: 1  
you.": 1  
"How: 2  
long: 1  
been: 7  
'Big: 1  
D': 1  
then?": 2  
Harry.: 1  
"Shut: 4  
it,: 2  
snarled: 3  
turning: 2  
away: 2  
"Cool: 1  
name,: 1  
Harry,: 1  
grinning: 1  
falling: 1  
step: 1  
beside: 1  
cousin.: 2  
"But: 1  
you'll: 1  
always: 1  
Ickle: 1  
Diddykins: 1  
me.": 2  
"I: 8  
said,: 2  
SHUT: 1  
IT!": 3  
whose: 1  
ham-like: 1  
hands: 4  
curled: 3  
fists.: 1  
16: 1  
"Don't: 3  
know: 7  
that's: 4  
your: 6  
mum: 1  
calls: 1  
you?": 4  
face.": 1

"You: 3  
don't: 3  
tell: 2  
her: 5  
shut: 4  
face.: 1  
What: 2  
'popkin': 1  
'Dinky: 1  
Diddydums,: 1  
can: 2  
I: 10  
them: 4  
nothing.: 1  
The: 8  
effort: 1  
keeping: 1  
himself: 1  
hitting: 2  
demanding: 1  
self-control.: 1  
"So: 1  
who've: 1  
beating: 1  
tonight?": 1  
asked,: 1  
grin: 1  
fading.: 1  
"Another: 1  
ten-year-old?: 1  
Mark: 1  
Evans: 1  
two: 1  
nights: 1  
ago: 1  
-": 7  
"He: 3  
asking: 1  
"Oh: 1  
yeah?": 1  
cheeked: 1  
"Yeah?: 1  
Did: 1  
say: 1  
look: 2  
pig: 1  
taught: 1  
walk: 2  
its: 7



hind: 1  
legs?: 1  
'Cause: 1  
cheek,: 1  
Dud,: 1  
true: 1  
...": 1  
A: 7  
muscle: 1  
twitching: 1  
jaw.: 1  
It: 4  
gave: 3  
enormous: 2  
satisfaction: 1  
how: 2  
furious: 1  
making: 1  
Dudley;; 1  
felt: 4  
as: 18  
though: 6  
siphoning: 1  
own: 3  
cousin,: 1  
only: 2  
outlet: 1  
had.: 1  
They: 3  
down: 4  
narrow: 1  
alleyway: 3  
where: 2  
first: 1  
seen: 1  
Sirius: 1  
which: 1  
formed: 1  
shortcut: 1  
between: 2  
Wisteria: 1  
Walk.: 1  
empty: 1  
darker: 1  
than: 3  
streets: 1  
it: 11  
linked: 1  
because: 1

there: 4  
streetlamps.: 1  
Their: 1  
17: 1  
footsteps: 3  
muffled: 1  
garage: 1  
walls: 1  
one: 2  
side: 2  
high: 1  
fence: 1  
other.“Think: 1  
you’re: 5  
big: 1  
man: 1  
carrying: 1  
thing,: 2  
few: 1  
seconds.: 1  
“What: 5  
thing?”: 1  
“That: 1  
thing: 6  
hiding.”: 1  
grinned: 1  
“Not: 2  
stupid: 1  
look,: 1  
you,: 2  
Dud?: 1  
s’pose: 1  
were,: 1  
wouldn’t: 2  
able: 1  
talk: 2  
same: 2  
time...”: 1  
pulled: 1  
wand.: 2  
saw: 2  
sideways: 1  
it.: 1  
“You’re: 1  
allowed,”: 1  
once.: 1  
not.: 1  
You’d: 1  
get: 1

expelled: 1  
freak: 1  
school: 1  
go: 1  
to." : 1  
d'you: 3  
haven't: 2  
changed: 1  
rules,: 1  
D?": 1  
"They: 1  
haven't," : 1  
sound: 1  
completely: 2  
convinced.: 1  
laughed: 1  
softly.: 1  
guts: 1  
take: 1  
me: 1  
without: 2  
snarled.: 3  
"Whereas: 1  
just: 4  
need: 1  
four: 1  
mates: 1  
behind: 3  
beat: 1  
ten-year-old.: 1  
You: 2  
boxing: 2  
title: 1  
keep: 1  
banging: 1  
about?: 1  
How: 1  
old: 2  
opponent?: 1  
Seven?: 1  
Eight?": 1  
sixteen: 1  
information," : 1  
"and: 1  
cold: 2  
twenty: 1  
minutes: 1  
18: 1  
I'd: 1  
finished: 1

him: 4  
twice: 1  
heavy: 1  
you.: 1  
wait: 1  
till: 1  
Dad: 1  
"Running: 1  
Daddy: 1  
now,: 2  
you?: 2  
Is: 1  
ickle: 1  
champ: 1  
frightened: 2  
nasty: 1  
Harry's: 6  
wand?": 1  
this: 2  
brave: 2  
sneered: 1  
"This: 1  
is: 1  
Diddykins.: 1  
That's: 1  
we: 1  
goes: 1  
dark: 1  
this." : 1  
mean: 2  
bed!": 1  
stopped: 2  
walking.: 1  
too,: 1  
staring: 2  
From: 1  
little: 1  
could: 6  
face,: 1  
wearing: 1  
strangely: 1  
triumphant: 1  
look.: 1  
mean,: 1  
I'm: 3  
bed?": 1  
nonplussed.: 1  
am: 1  
supposed: 1

pillows: 1  
or: 2  
something?": 1  
heard: 4  
last: 2  
night,": 1  
breathlessly.: 1  
"Talking: 1  
sleep.: 1  
Moaning.": 1  
mean?": 1  
but: 5  
cold,: 1  
plunging: 1  
sensation: 1  
stomach.: 1  
revisited: 1  
graveyard: 1  
night: 2  
dreams.: 1  
harsh: 1  
bark: 1  
laughter: 3  
then: 3  
adopted: 1  
high-pitched,: 1  
whimpering: 2  
voice.: 1  
": 3  
'Don't: 1  
kill: 4  
Cedric!: 2  
Don't: 2  
Cedric!': 1  
Who's: 1  
Cedric: 1  
boyfriend?": 1  
lying: 2  
automatically.: 1  
mouth: 1  
gone: 1  
dry.: 1  
knew: 1  
wasn't: 1  
else: 1  
Cedric?: 1  
19: 1  
'Dad!: 1  
Help: 1

me,: 3  
Dad!: 3  
He's: 3  
going: 4  
Boohoo!': 1  
": 2  
up,": 1  
quietly.: 1  
up,: 3  
warning: 1  
you!": 1  
'Come: 1  
help: 4  
Mum,: 1  
come: 2  
me!: 2  
killed: 1  
Dad,: 1  
-': 1  
point: 1  
me!": 1  
backed: 1  
alley: 3  
wall.: 1  
pointing: 1  
wand: 9  
directly: 1  
heart.: 1  
fourteen: 1  
years': 1  
hatred: 1  
pounding: 1  
veins: 1  
give: 1  
strike: 1  
jinx: 1  
thoroughly: 1  
he'd: 2  
crawl: 1  
an: 4  
insect,: 1  
struck: 1  
dumb,: 1  
sprouting: 1  
feelers: 1  
ever: 1  
again,": 1  
"D'you: 1  
understand: 2

me?": 2  
"Point: 2  
somewhere: 2  
else!": 2  
do: 2  
"DO: 1  
YOU: 2  
UNDERSTAND: 1  
ME?": 1  
"GET: 2  
THAT: 1  
THING: 1  
AWAY: 1  
FROM: 1  
odd,: 1  
shuddering: 1  
gasp,: 1  
doused: 1  
icy: 3  
water.: 1  
Something: 1  
happened: 1  
night.: 1  
starstrewn: 1  
indigo: 1  
sky: 1  
suddenly: 2  
pitch-black: 1  
lightless: 1  
stars,: 2  
moon,: 1  
misty: 1  
streetlamps: 2  
either: 1  
end: 2  
vanished.: 1  
distant: 1  
grumble: 1  
whisper: 1  
trees: 1  
gone.: 1  
balmy: 1  
evening: 1  
piercingly,: 1  
bitingly: 1  
20: 1  
cold.: 1  
surrounded: 1  
total,: 1

impenetrable,: 1  
silent: 1  
darkness,: 2  
some: 1  
giant: 1  
hand: 2  
dropped: 1  
thick,: 1  
mantle: 1  
over: 6  
entire: 1  
alleyway,: 2  
blinding: 1  
them.: 1  
For: 1  
split: 1  
second: 3  
magic: 1  
meaning: 1  
to,: 1  
despite: 1  
fact: 1  
resisting: 1  
hard: 2  
reason: 1  
caught: 3  
senses: 2  
power: 1  
stars.: 1  
head: 4  
way: 1  
that,: 1  
trying: 2  
something,: 1  
pressed: 1  
eyes: 4  
weightless: 2  
veil.: 1  
terrified: 1  
voice: 3  
broke: 1  
ear.: 1  
"W-what: 1  
d-doing?: 1  
St-stop: 1  
it!": 1  
"I'm: 2  
doing: 2  
anything!: 1



Shut: 1  
move!": 1  
c-can't: 1  
see!: 1  
I've: 1  
g-gone: 1  
blind!: 1  
up!": 1  
stood: 3  
stock-still,: 1  
sightless: 1  
left: 1  
right.: 1  
intense: 1  
shivering: 1  
over;: 1  
goose: 1  
bumps: 1  
erupted: 2  
arms,: 1  
hairs: 1  
back: 3  
neck: 1  
standing: 1  
opened: 1  
fullest: 1  
extent,: 1  
blankly: 1  
unseeing: 1  
impossible...: 1  
couldn't: 3  
here...: 1  
Not: 1  
strained: 1  
ears...: 1  
hear: 1  
"I'll: 1  
t-tell: 1  
Dad!": 1  
whimpered.: 1  
"W-where: 1  
d-do: 1  
?": 1  
"Will: 1  
up?": 1  
hissed,: 1  
lis: 1  
21: 1  
fell: 1

silent.: 1  
dreading.: 1  
something: 3  
apart: 1  
themselves,: 1  
drawing: 1  
long,: 1  
hoarse,: 1  
rattling: 1  
breaths.: 1  
horrible: 2  
jolt: 1  
dread: 1  
trembling: 1  
freezing: 1  
air.: 1  
"C-cut: 1  
out!: 1  
Stop: 1  
it!: 1  
I'll: 1  
h-hit: 1  
swear: 1  
will!": 1  
"Dudley,: 1  
WHAM!: 1  
fist: 1  
contact: 1  
head,: 2  
lifting: 1  
feet.: 1  
Small: 1  
white: 1  
lights: 1  
popped: 1  
front: 1  
eyes;: 1  
hour: 1  
cleaved: 1  
two;: 1  
next: 1  
moment: 2  
landed: 1  
ground,: 4  
flown: 1  
hand.: 1  
moron,: 1  
Dudley!": 1  
yelled,: 1

watering: 1  
pain,: 1  
scrambled: 2  
knees,: 1  
now: 1  
feeling: 1  
frantically: 1  
blackness.: 1  
blundering: 1  
away,: 3  
fence,: 1  
stumbling.: 1  
"DUDLEY,: 2  
COME: 1  
BACK!: 1  
YOU'RE: 1  
RUNNING: 1  
RIGHT: 1  
AT: 1  
squealing: 1  
yell,: 1  
stopped.: 1  
At: 1  
moment,: 1  
creeping: 1  
chill: 1  
thing.: 1  
one.: 1  
KEEP: 2  
YOUR: 2  
MOUTH: 2  
SHUT!: 2  
WHATEVER: 1  
DO,: 1  
Wand!": 1  
22: 1  
muttered: 1  
frantically,: 1  
flying: 1  
ground: 1  
spiders.: 1  
"Where's: 1  
Lumos!": 1  
spell: 3  
automatically,: 1  
desperate: 1  
light: 2  
search: 1  
disbelieving: 1

relief,: 1  
flared: 1  
inches: 1  
tip: 3  
ignited.: 1  
snatched: 1  
feet,: 2  
around.: 1  
His: 2  
stomach: 1  
over.: 1  
towering,: 1  
hooded: 2  
figure: 1  
gliding: 1  
smoothly: 1  
hovering: 1  
face: 4  
visible: 1  
beneath: 1  
robes,: 2  
sucking: 1  
came.: 1  
Stumbling: 1  
backward,: 2  
raised: 1  
"Expecto: 2  
Patronum!": 2  
silvery: 1  
wisp: 2  
vapor: 1  
shot: 1  
dementor: 5  
slowed,: 1  
hadn't: 1  
worked: 1  
properly,: 1  
tripping: 1  
retreated: 1  
farther: 1  
bore: 1  
upon: 1  
panic: 1  
fogging: 1  
brain: 1  
concentrate: 1  
pair: 1  
gray,: 1  
slimy,: 1

scabbed: 1  
slid: 1  
inside: 3  
dementor's: 4  
reaching: 1  
him.: 2  
rushing: 1  
noise: 1  
filled: 2  
ears.: 1  
sounded: 1  
dim: 1  
distant...: 1  
Another: 1  
silver: 5  
smoke,: 1  
feebler: 1  
last,: 1  
drifted: 1  
anymore,: 1  
work: 1  
shrill,: 1  
highpitched: 2  
laughter...: 1  
smell: 1  
23: 1  
putrid,: 1  
death-cold: 1  
breath,: 1  
filling: 1  
lungs,: 1  
drowning: 1  
Think: 1  
happy...: 1  
happiness: 1  
him...: 2  
fingers: 1  
closing: 1  
throat: 1  
growing: 1  
louder: 1  
louder,: 1  
spoke: 1  
"Bow: 1  
death,: 1  
Harry...: 1  
might: 1  
even: 1  
painless...: 1

know...: 1  
never: 2  
died...: 1  
Ron: 1  
Hermione: 1  
again: 1  
And: 1  
faces: 1  
burst: 2  
clearly: 1  
mind: 1  
fought: 1  
breath: 1  
"EXPECTO: 1  
PATRONUM!": 1  
An: 1  
stag: 4  
wand;: 1  
antlers: 2  
place: 1  
heart: 1  
should: 1  
been;: 1  
thrown: 2  
charged,: 1  
swooped: 1  
batlike: 1  
defeated.: 1  
"THIS: 1  
WAY!": 1  
shouted: 1  
stag.: 1  
Wheeling: 1  
sprinted: 1  
holding: 1  
lit: 1  
aloft.: 1  
"DUDLEY?: 1  
DUDLEY!": 1  
run: 1  
barely: 2  
dozen: 1  
steps: 1  
reached: 1  
them:: 1  
arms: 1  
clamped: 1  
face;: 1  
crouching: 1  
low: 1

gripping: 1  
wrists: 1  
slimy: 1  
hands,: 1  
prizing: 1  
slowly,: 1  
lovingly: 1  
apart,: 1  
lowering: 1  
kiss: 1  
bellowed,: 1  
rushing,: 1  
roaring: 1  
sound,: 1  
conjured: 1  
galloping: 1  
past: 1  
eyeless: 1  
24: 1  
inch: 1  
it;: 1  
air: 2  
and,: 1  
fellow,: 1  
soared: 1  
absorbed: 1  
darkness.: 1  
cantered: 1  
dissolved: 1  
mist.: 1  
Moon,: 1  
life.: 1  
warm: 1  
breeze: 1  
swept: 1  
alleyway.: 1  
Trees: 1  
rustled: 1  
neighboring: 1  
gardens: 1  
mundane: 1  
rumble: 1  
quite: 1  
still,: 1  
vibrating,: 1  
taking: 1  
abrupt: 1  
return: 1  
normality.: 1

After: 1  
became: 1  
aware: 1  
T-shirt: 1  
sticking: 1  
him,: 2  
drenched: 1  
sweat.: 1  
believe: 1  
happened.: 1  
Dementors: 1  
here,: 1  
lay: 1  
shaking.: 1  
bent: 1  
whether: 1  
fit: 1  
state: 1  
stand: 1  
loud,: 1  
running: 1  
instinctively: 1  
raising: 1  
spun: 1  
heel: 1  
newcomer.: 1  
Mrs.: 1  
Figg,: 1  
batty: 1  
neighbor,: 1  
panting: 1  
sight.: 1  
Her: 1  
grizzled: 1  
gray: 1  
hair: 1  
escaping: 1  
hairnet,: 1  
clanking: 1  
string: 1  
shopping: 1  
bag: 1  
swinging: 1  
wrist,: 1  
tartan: 1  
carpet: 1  
slippers.: 1  
stow: 1  
hurriedly: 1



```
put: 1
idiot: 1
boy!": 1
she: 1
shrieked.: 1
around?: 1
Oh,: 1
Mundungus: 1
Fletcher!": 1
```

2. From the second text file (file2.txt), write Python code and use MapReduce to count how many times non-English words (names, places, spells etc.) were used. List those words and how many times each was repeated. There are multiple ways of doing this. You can use pyenchant (<https://pypi.org/project/pyenchant/>), pyspellchecker (<https://pyspellchecker.readthedocs.io/en/latest/>) or just download a list of words (<http://www.gwicks.net/dictionaries.htm>) and search through them.

```
from spellchecker import SpellChecker
import re

# Initialize the spell checker
spell = SpellChecker()

# Define a regular expression pattern to match words
WORD_RE = re.compile(r'\w+')

# Read the text from the file using UTF-8 encoding
with open('file2.txt', 'r', encoding='utf-8') as file:
    text = file.read()

# Define a function to preprocess words (remove punctuation)
def preprocess_word(word):
    return re.sub(r'[^w\s]', '', word) # Removes all non-
    alphanumeric characters except spaces

# Split text into words
words = text.split()

# Define a mapper function to emit non-English words
def mapper(word):
    word = preprocess_word(word)
    correction = spell.correction(word)
    if correction != word:
        return (word, 1)
    return (word, 0)

# Define a reducer function to count non-English words
def reducer(counts, word_count):
```

```

    word, count = word_count
    counts[word] += count
    return counts

# Map and reduce to get counts of non-English words
non_english_counts = reduce(reducer, map mapper, words), Counter())

# Print non-English words and their counts
for word, count in non_english_counts.items():
    if count > 0:
        print(f'{word}: {count}')

```

```

P: 11
g: 11
: 51
JK: 11
Rowling: 11
Cmon: 1
houseelf: 2
Mrs: 19
Weasley: 22
Tonks: 9
exasperatedly: 1
earsplitting: 1
motheaten: 1
lifesize: 1
Halfbreeds: 1
Yooooou: 1
Harrys: 4
Hasnt: 1
Sirius: 4
Mr: 4
redhaired: 1
hornrimmed: 1
MadEye: 1
didnt: 4
Greenland: 1
Weasleys: 1
Evanesco: 1
Mundungus: 13
Somen: 1
m: 1
sleepily: 1
gree: 1
balefully: 1
arry: 1
pology: 1
Crookshanks: 2
Hermiones: 1
bandylegged: 1

```

absentmindedly: 1  
grimfaced: 1  
youre: 2  
incredulously: 1  
Animagus: 1  
Wormtail: 1  
Dumbledores: 1  
bracingly: 1  
Snapes: 2  
cleanings: 1  
hasnt: 1  
fifteenthcentury: 1  
goblinwrought: 1  
Thatd: 1  
butterbeer: 3  
breadboard: 1  
YOURE: 1  
Alastor: 1  
doxies: 1  
beaklike: 1