

Code: python file called wordcount_2.py

```
from mrjob.job import MRJob
from mrjob.step import MRStep

class MRWordCount(MRJob):

    def mapper(self, _, line):
        # Split the line into words
        for word in line.split():
            # Yield each word in lowercase with a count of 1
            yield (word.lower(), 1)

    def reducer(self, word, counts):
        # Sum the counts for each word
        yield (word, sum(counts))

    def steps(self):
        return [
            MRStep(mapper=self.mapper,
                  reducer=self.reducer)
        ]

if __name__ == '__main__':
    MRWordCount.run()
```

```
$ python wordcount_2.py file1.txt
```

No configs found; falling back on auto-configuration

No configs specified for inline runner

Creating temp directory

C:\Users\naaale\AppData\Local\Temp\wordcount_2.naaale.20241003.210433.775539

Running step 1 of 1...

job output is in

C:\Users\naaale\AppData\Local\Temp\wordcount_2.naaale.20241003.210433.775539\output

Streaming final output from

C:\Users\naaale\AppData\Local\Temp\wordcount_2.naaale.20241003.210433.775539\output...

"(or" 1

"(or," 1

"10" 1

"11" 1

"12" 1

"7" 1

"8" 1

"9" 1

"?\u0094" 2

"\u0091back\u0092" 1

"\u0093" 1

"\u0093a" 1

"\u0093and" 1

"\u0093back?" 1

"\u0093but" 3

"\u0093but,\u0094" 1

"\u0093he-who-must-not-be-named!\u0094" 1

"\u0093hogwarts,\u0094" 1

"\u0093how" 1

"\u0093i" 2

"\u0093i\u0092m" 1
"\u0093involved,\u0094" 1
"\u0093it\u0092s" 1
"\u0093lord" 1
"\u0093middle" 1
"\u0093my" 1
"\u0093no" 1
"\u0093no," 1
"\u0093no,\u0094" 1
"\u0093not" 1
"\u0093now," 1
"\u0093oh," 1
"\u0093prime" 1
"\u0093serious\u0094" 1
"\u0093so" 1
"\u0093that" 1
"\u0093the" 2
"\u0093they\u0092ve" 1
"\u0093we" 2
"\u0093we\u0092ll" 1
"\u0093we\u0092re" 1
"\u0093well," 2
"\u0093what" 1
"\u0093why" 1
"\u0093yes," 2
"\u0093you\u0092re" 1
"\u0094" 2
"\u0096" 6
"\u0097" 15

"\u0097\u0094" 3

"a" 48

"abandoned" 1

"about" 4

"about." 1

"above" 1

"actually" 1

"added." 1

"affect" 1

"afraid" 2

"again" 2

"again," 2

"again." 2

"ago," 1

"air" 1

"air," 1

"alive" 1

"alive," 1

"alive,\u0094" 1

"alive.\u0094" 1

"alive?" 1

"all" 5

"almost" 1

"alone" 2

"aloud" 1

"already" 1

"always" 2

"am" 1

"amber" 1

"amount" 1
"an" 3
"and" 40
"announced" 2
"any" 1
"anybody?\u0094" 1
"anything" 1
"anyway," 2
"appear" 1
"appeared" 2
"are" 3
"arrangements" 1
"arrival" 1
"arrival." 1
"art" 1
"as" 5
"ask" 2
"assurances" 1
"at" 11
"attempt" 2
"axminster," 1
"azkaban,\u0094" 1
"azkaban." 1
"back" 1
"back." 1
"back.\u0094" 1
"be" 7
"becoming" 1
"been" 7

"before" 3
"before." 1
"behind." 1
"being" 1
"believe" 1
"best" 1
"better" 2
"beyond" 1
"black" 1
"black," 1
"black\u0092s" 1
"bleated" 1
"blood" 6
"blustered" 1
"body" 1
"book" 1
"bother" 2
"bowler" 1
"boy" 1
"breakout" 2
"breakout?\u0094" 1
"breath" 1
"breathlessly," 1
"bridge" 1
"bringing" 1
"brockdale" 1
"brought" 1
"builder" 1
"burst" 1

"but" 10
"by" 2
"cabinet" 1
"called" 2
"came" 1
"campaign." 1
"can\u0092t" 1
"canvas" 1
"carpenters," 1
"cast" 1
"certain" 1
"certainly" 1
"chair." 1
"chancellor" 1
"chewing" 1
"chin," 1
"chorley\u0092s" 1
"chortling," 1
"cloak" 1
"come" 2
"committed" 1
"community?\u0094" 1
"concerns" 1
"concerns,\u0094" 1
"conjured" 1
"considerable" 1
"control" 1
"conversation" 1
"convince" 1

"corner" 2
"could" 4
"country" 1
"country.\u0094" 1
"course" 1
"course," 1
"creatures" 2
"crimes" 1
"cup" 1
"currently" 1
"dangerous" 2
"dare" 1
"day" 1
"day,\u0094" 1
"dealing" 1
"dear" 1
"deep" 1
"delighted" 1
"dementors" 1
"department" 1
"desperate" 1
"despite" 1
"details" 1
"did" 2
"didn\u0092t" 1
"disappearance" 1
"discussion," 1
"disheveled" 1
"dismay," 1

"do" 1
"don\u0092t" 4
"down" 4
"down," 2
"dragons" 2
"dragons?\u0094" 1
"drew" 1
"dripping" 1
"dumbledore" 2
"during" 1
"e" 6
"each" 4
"earlier." 1
"effected" 1
"election" 1
"else" 1
"emerald" 1
"encounter" 1
"encounter," 1
"end," 1
"enough" 1
"er" 1
"erupted" 1
"escaped" 1
"even" 1
"even," 1
"ever" 1
"exactly" 1
"excellent." 1

"exchequer" 1
"explain" 1
"extremely" 1
"eye" 1
"eyes" 1
"fact" 1
"family" 1
"fastened" 1
"fatherly" 1
"fear" 1
"feared" 1
"fifteen" 1
"fill" 1
"find" 1
"fire" 2
"fireplace," 2
"firmly" 1
"first" 1
"flames," 1
"flames." 1
"flight" 1
"flustered" 1
"foolish" 1
"foot" 1
"for" 10
"foreign" 1
"forgot,\u0094" 1
"former" 1
"found" 2

"frame" 1
"fretful" 1
"from" 3
"fudge" 21
"fudge," 4
"fudge." 3
"fudge\u0092s" 2
"full" 1
"g" 6
"gazed" 1
"gently." 1
"gerbil" 1
"gerbil." 1
"given" 1
"glasses" 1
"gloomy" 1
"going" 3
"good" 2
"got" 2
"graver" 1
"great," 1
"green" 1
"groped" 1
"grueling" 1
"habit" 1
"had" 47
"half" 6
"hallucination" 1
"hand," 1

"hand." 2
"happened" 1
"happened." 1
"harassed-looking" 1
"harry" 7
"has" 1
"hasn\u0092t" 1
"hat" 1
"have" 8
"he" 23
"he-who-must-not-be-named" 2
"he\u0092s" 6
"head," 1
"heard" 1
"help" 1
"herbert" 1
"herself" 1
"highly" 1
"him" 4
"him-" 1
"him." 2
"him?" 1
"himself" 4
"his" 27
"historian," 1
"hoarsely." 1
"hoax" 1
"hoax," 1
"hope" 3

"hope." 1
"hoped" 1
"hopelessly" 1
"horrible" 1
"hospital" 1
"hour." 1
"however," 2
"hurricane." 1
"i" 13
"i\u0092d" 1
"i\u0092ll" 1
"i\u0092m" 3
"i\u0092ve" 1
"if" 4
"imminent" 1
"importing" 1
"impossible" 1
"in" 17
"in." 1
"incident," 1
"indeed" 1
"inform" 1
"injuries." 1
"inside" 1
"instead" 1
"instructed" 1
"interrupted." 1
"into" 7
"is!\u0094" 1

"is" 5
"is,\u0094" 1
"isolated" 1
"it" 9
"it," 2
"it\u0092s" 3
"j.k." 6
"jacket." 1
"just" 3
"killed?" 1
"killer" 1
"killing," 1
"know!\u0094" 1
"know" 5
"know," 1
"known" 1
"kwidditch" 1
"lack" 1
"large" 2
"last" 1
"last," 1
"last." 1
"later" 1
"later," 1
"laughed." 1
"leave," 1
"leaving" 1
"left" 1
"less" 2

"let" 2
"liaison" 1
"like" 3
"like)" 1
"liked" 1
"likely" 1
"liquid" 1
"little" 3
"live" 1
"live." 1
"lived," 1
"living" 1
"long" 1
"look" 1
"look.\u0094" 1
"looking" 1
"lot" 2
"made" 2
"magic" 2
"magical" 2
"maintain" 1
"making" 1
"maladies" 1
"man" 3
"man." 1
"mark" 1
"mass" 2
"may" 1
"me" 5

"me," 1
"mean" 1
"meant" 1
"meeting," 1
"memory" 2
"mention" 1
"might" 1
"minister!" 1
"minister" 25
"minister)," 1
"minister," 3
"minister." 4
"minister\u0092s" 4
"modifications" 1
"moment!\u0094" 1
"moment" 1
"moment," 1
"more" 2
"more," 1
"motionless" 1
"motionless," 1
"move" 1
"much," 1
"muddy-brown" 1
"muggle" 3
"muggles" 2
"muggles." 1
"mungo\u0092s" 1
"murders" 1

"must" 1
"mysterious" 1
"name" 2
"named" 1
"nasty" 1
"need" 2
"never" 4
"never," 1
"nevertheless." 1
"news" 2
"next" 2
"niece" 1
"night" 1
"night.\u0094" 1
"no" 2
"no." 1
"nobody" 1
"non-magical" 1
"none" 1
"nor" 1
"north" 1
"nose;" 1
"not" 10
"not." 1
"nothing" 1
"nothing;" 1
"notice" 1
"notify" 1
"now" 1

"now." 1
"occasionally" 1
"occupant" 1
"odds-on" 1
"of" 44
"of," 1
"off." 1
"offered" 1
"office" 2
"office," 1
"office." 1
"on" 8
"once" 3
"one" 4
"only" 2
"opposition" 1
"opposition.\u0094" 1
"or" 4
"other" 3
"other," 1
"others," 1
"otherwise," 1
"ought" 2
"our" 2
"out" 11
"out." 1
"over" 1
"own" 2
"p" 6

"painting" 1
"panic." 1
"panted," 1
"parchment," 1
"persistent" 1
"picture" 1
"piece" 1
"pin-striped" 1
"planned" 1
"planning" 1
"playing" 1
"pocket." 1
"point," 1
"poking" 1
"population," 1
"portrait" 3
"potter" 6
"potter," 1
"powder" 1
"predecessor." 1
"press" 1
"previously," 1
"prime" 34
"prince" 6
"prise" 1
"prison" 1
"private" 1
"properly" 1
"proved" 1

"pulled" 1
"purposes" 1
"pushed" 1
"put" 1
"quite" 3
"ranting" 1
"rather" 2
"realized" 1
"really" 3
"refused" 1
"regulation" 1
"rejoin" 1
"remained" 1
"reminders" 1
"remotest" 1
"remove." 1
"repeated" 1
"resented" 1
"resolved" 1
"rest" 1
"reveals" 1
"rid" 1
"rim" 1
"room" 1
"rounded" 1
"routine," 1
"rowing" 6
"rule" 1
"run" 1

"safer" 1
"said" 6
"said," 3
"same" 1
"sat" 1
"saw" 1
"say" 3
"say," 2
"say." 1
"says" 1
"scratching" 1
"sea," 1
"secrecy.\u0094" 1
"secretary" 1
"see" 2
"seeing" 1
"seen" 2
"sense" 1
"serious" 1
"several" 2
"shock" 1
"should" 3
"shout," 1
"shouted" 1
"shower" 1
"shuddered" 1
"sight," 1
"silent" 1
"simply" 2

"sit" 3
"sleep" 1
"snapped" 1
"snarled" 1
"so" 2
"some" 1
"something" 3
"sopping" 1
"sorry" 1
"sorry." 1
"sort" 1
"soul," 1
"sound." 1
"sounded" 2
"sparks." 1
"speech," 1
"sphinx" 1
"sphinx." 1
"sphinxes" 1
"spluttered" 1
"spoke." 1
"spot" 1
"squinted" 1
"st." 1
"started" 1
"state" 1
"stepped" 1
"stepping" 1
"sternly" 1

"still" 2
"still." 1
"stood" 3
"stretch" 1
"subject" 1
"support," 1
"suppose," 1
"sure" 1
"surprised" 1
"sworn" 1
"take" 1
"taken" 1
"taking" 1
"talked" 1
"talking" 1
"teacup" 2
"tell" 3
"tells" 1
"term" 1
"terrible" 1
"than" 4
"that" 25
"that," 1
"that\u0092s" 1
"the" 101
"their" 1
"them" 2
"then" 1
"then," 1

"then,\u0094" 1
"then?" 1
"then?\u0094" 2
"there" 2
"there," 2
"there\u0092s" 1
"therefore," 1
"they" 3
"thin" 2
"thing" 2
"think" 3
"this" 5
"this," 3
"those" 1
"though" 1
"thought" 2
"thousand" 1
"three" 3
"three,\u0094" 1
"throw" 1
"thrown" 1
"thrust" 1
"time" 3
"time," 1
"tipping" 1
"to" 42
"told" 3
"tonight," 1
"tonight.\u0094" 1

"too." 1
"took" 1
"tournament," 1
"trained" 1
"transferred" 1
"tricks" 1
"tried" 3
"triwizard" 1
"turned" 1
"twice," 1
"two" 2
"two," 1
"ugly" 1
"uncomfortable" 1
"under" 1
"understand" 1
"unless" 1
"unsuccessfully" 1
"up" 5
"uproar\u0094" 1
"vain" 1
"vanished" 2
"very" 3
"visit." 1
"voice" 1
"vol" 1
"wait" 1
"walking" 2
"wall," 1

"wand" 1
"wand," 1
"warned" 1
"warning," 1
"was" 12
"wasn\u0092t" 1
"watching" 1
"water" 1
"way" 1
"way." 1
"we" 4
"we\u0092re" 1
"we\u0092ve" 1
"wear" 2
"week." 1
"well," 2
"well-informed" 1
"were" 3
"wet" 1
"what" 3
"what?\u0094" 1
"what\u0092s" 1
"whatever" 1
"when" 6
"which" 2
"while" 1
"whiskey" 1
"whiskey," 1
"whiskey-free" 1

"who" 6
"whooshing" 1
"why" 4
"wide" 1
"window," 1
"wishing" 1
"with" 6
"without" 2
"wizard" 2
"wizarding" 1
"won\u0092t" 1
"work" 1
"world" 2
"worry!\u0094" 1
"worry," 2
"worry,\u0094" 1
"worrying" 1
"worst" 2
"would" 5
"wrote" 1
"yawning," 1
"year" 1
"years" 4
"yes," 1
"yet" 1
"you" 13
"you," 1
"you-know-who" 1
"you-know-who." 1

"you-know-who\u0092s" 1

"you.\u0094" 1

"you\u0092ll" 2

"you\u0092re" 2

"your" 1

"|" 6

Code: python File called: not-english-words.py

```
from spellchecker import SpellChecker
from mrjob.job import MRJob
from mrjob.step import MRStep

class MRNonEnglishWordCount(MRJob):

    def mapper(self, _, line):
        spell = SpellChecker()
        words = line.lower().split()
        for word in words:
            if word not in spell: # If the word is not recognized
                yield (word, 1)

    def reducer(self, word, counts):
        yield (word, sum(counts))

    def steps(self):
        return [
            MRStep(mapper=self.mapper,
                    reducer=self.reducer)
        ]

if __name__ == '__main__':
    MRNonEnglishWordCount.run()
```

```
$ python not_english-words.py file2.txt
```

No configs found; falling back on auto-configuration

No configs specified for inline runner

Creating temp directory C:\Users\naaale\AppData\Local\Temp\not_english-words.naaale.20241003.210941.595869

Running step 1 of 1...

job output is in C:\Users\naaale\AppData\Local\Temp\not_english-words.naaale.20241003.210941.595869\output

Streaming final output from C:\Users\naaale\AppData\Local\Temp\not_english-words.naaale.20241003.210941.595869\output...

"90" 1

"91" 1

"92" 1

"93" 1

"94" 1

"95" 1

"?\u0094" 2

"\u0092night" 1

"\u0093" 1

"\u0093all" 1

"\u0093arthur," 2

"\u0093both" 1

"\u0093bread," 1

"\u0093but" 1

"\u0093correct,\u0094" 1

"\u0093dear," 1

"\u0093ever" 1

"\u0093everyone\u0092s" 1

"\u0093he" 2

"\u0093he\u0092s" 1

"\u0093hello," 1
"\u0093hi," 1
"\u0093i" 5
"\u0093i\u0092d" 1
"\u0093is" 1
"\u0093it\u0092s" 1
"\u0093molly!" 1
"\u0093molly!\u0094" 1
"\u0093mollywobbles,\u0094" 1
"\u0093mortal" 1
"\u0093no," 2
"\u0093now" 1
"\u0093of" 1
"\u0093oh" 1
"\u0093oh," 1
"\u0093please" 1
"\u0093sit" 1
"\u0093slughorn" 1
"\u0093so" 2
"\u0093take" 1
"\u0093thanks" 1
"\u0093thanks," 1
"\u0093that\u0092s" 1
"\u0093to" 1
"\u0093traveling.\u0094" 1
"\u0093we" 1
"\u0093well," 2
"\u0093what" 2
"\u0093wotcher," 1

"\u0093yeah," 1
"\u0093yes," 2
"\u0093yes,\u0094" 1
"\u0093you" 1
"\u0093you\u0092re" 1
"\u0094" 3
"\u0096" 6
"\u0097" 6
"\u0097\u0094" 2
"account,\u0094" 1
"added," 1
"again;" 1
"ages," 1
"ah," 1
"air," 1
"air." 1
"alone," 1
"am,\u0094" 1
"ambition?\u0094" 1
"appearance." 1
"arthur" 3
"arthur\u0092s" 3
"asked." 1
"balding," 1
"be," 1
"bed," 1
"blistering." 1
"bowl." 1
"bread," 1

"bubble-gum-pink" 1
"bubotuber" 1
"can\u0092t" 1
"check." 1
"cloak." 2
"clock\u0092s" 1
"coming!\u0094" 1
"course," 1
"course." 1
"courteously," 1
"crookshanks" 1
"day\u0092s" 1
"dear," 2
"dear." 1
"dear?\u0094" 1
"didn\u0092t" 2
"doing," 1
"don\u0092t" 5
"door." 2
"doorknob," 1
"doorstep," 1
"down," 2
"down." 1
"drawn," 1
"dumbledore" 2
"dumbledore," 2
"dumbledore." 1
"dumbledore\u0092s" 1
"ears." 1

"eater," 1
"eater." 1
"eaters." 1
"enough," 1
"everybody\u0092s" 1
"everyone.\u0094" 1
"everywhere," 1
"expected." 1
"eyes." 1
"face." 1
"fact," 1
"favorites," 1
"first!\u0094" 1
"fletcher," 1
"gasped." 1
"going,\u0094" 1
"great!\u0094" 1
"hair." 1
"hands," 1
"hands." 1
"harry," 2
"harry,\u0094" 1
"harry.\u0094" 1
"harry?\u0094" 1
"harry\u0092s" 1
"he\u0092d" 1
"he\u0092s" 2
"head." 1
"heart-shaped" 1

"hello," 1
"her." 1
"here?" 1
"hermione"s 1
"highflier." 1
"him." 1
"him?" 1
"hogwarts" 1
"hogwarts," 1
"home," 1
"horace" 1
"horn-rimmed" 1
"hour." 1
"hours." 1
"hungry," 1
"husband," 1
"i" 1
"i've" 1
"ill," 1
"in." 1
"is," 1
"is." 1
"it," 1
"it"s 5
"j.k." 6
"job," 2
"job?" 1
"know," 1
"late." 1

"look," 1
"lucky,\u0094" 1
"mad-eye" 1
"me,\u0094" 1
"mean," 1
"mean,\u0094" 1
"member," 1
"midnight." 1
"ministry\u0092s" 1
"mistakes." 1
"molly," 1
"molly.\u0094" 1
"mr." 5
"mrs." 21
"much." 1
"muggle" 1
"mundungus" 1
"neck," 1
"news." 1
"night," 1
"no," 1
"nodded," 1
"now!\u0094" 1
"now,\u0094" 1
"now." 1
"nymphadora!\u0094" 1
"objects." 1
"off,\u0094" 1
"off." 1

"oh!\u0094" 1
"once." 1
"once:" 1
"open." 1
"over;" 1
"pale," 1
"peril.\u0094" 1
"plugs." 1
"procedure," 1
"professor,\u0094" 1
"promoted!\u0094" 1
"purring." 1
"question!\u0094" 1
"quickly," 1
"really," 2
"red-haired" 1
"red;" 1
"remus" 1
"right," 1
"robes." 1
"ron,\u0094" 1
"ron\u0092s" 2
"rowing" 6
"rubbish.\u0094" 1
"rufus" 2
"said." 1
"scrimgeour" 1
"scrimgeour.\u0094" 1
"see," 2

"servant.\u0094" 1
"shoulder;" 1
"shoulders." 1
"shut." 1
"side," 1
"sighed," 1
"silly." 1
"situation," 1
"slughorn" 2
"slughorn\u0092s" 1
"smile." 1
"sneakoscopes" 1
"so-called" 1
"softly," 1
"soup," 1
"soup." 1
"speak." 1
"spot." 1
"stay," 1
"stove," 1
"sweet,\u0094" 1
"sympathy," 1
"table." 1
"table;" 1
"there," 1
"thick," 1
"thin," 1
"think." 1
"this," 1

"this." 1
"threshold." 1
"together?" 1
"tonks" 1
"tonks," 2
"tonks." 1
"troubled." 1
"up." 1
"up." 2
"ups," 1
"voice," 1
"voice." 1
"wall," 1
"wand." 1
"was." 1
"we're" 1
"weasley" 16
"weasley," 5
"weasley." 2
"weasley." 1
"weasley's" 2
"weasleys" 1
"weekend," 1
"well," 2
"who's" 1
"who've" 1
"wisely." 1
"work?" 1
"yard;" 1

"yes," 1

"yesterday,\u0094" 1

"you!\u0094" 1

"you," 1

"you-know-who" 2

"you-know-who," 1

"you." 1

"you?\u0094" 1

"you\u0092ve" 1

"yourself." 1

"|" 6